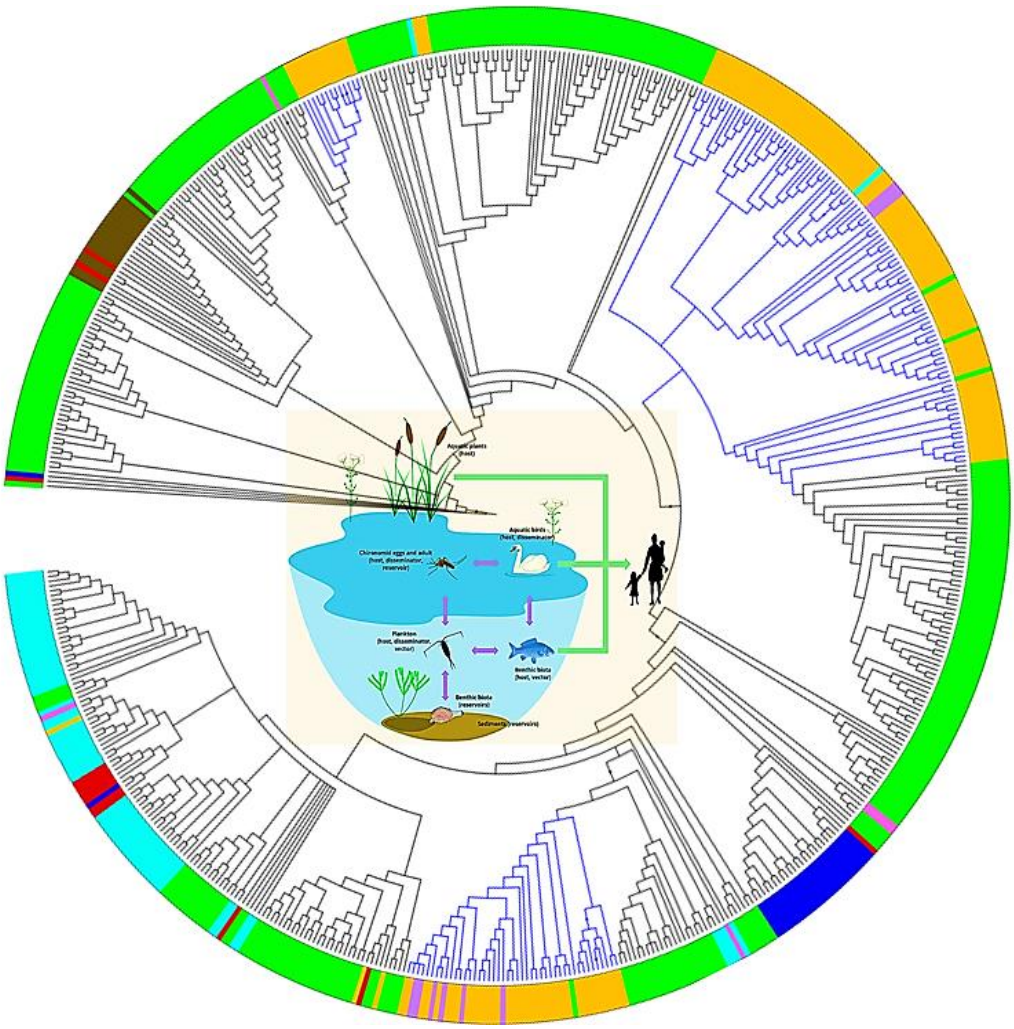
UNIVERSITY OF COPENHAGEN

FACULTY OF HEALTH AND MEDICAL SCIENCES

Genomic Epidemiology and Aquatic Reservoirs of the Seventh Pandemic *Vibrio cholerae* in Tanzania



PhD Thesis 2019 – Yaovi Mahuton Gildas Hounmanou

This Thesis has been submitted to the Graduate School of Health and Medical Sciences, University of Copenhagen on 08/10/2019.

University of Copenhagen

Faculty of Health and Medical Sciences

Genomic Epidemiology and Aquatic Reservoirs of the Seventh Pandemic *Vibrio cholerae* in Tanzania

PhD Thesis 2019

Yaovi Mahuton Gildas Hounmanou

Department of Veterinary and Animal Sciences

Faculty of Health and Medical Sciences

University of Copenhagen

This Thesis has been submitted to the Graduate School of Health and Medical Sciences at the University of Copenhagen on 08/10/2019

Genomic Epidemiology and Aquatic Reservoirs of the Seventh Pandemic *Vibrio cholerae* in Tanzania

Author: Yaovi Mahuton Gildas Hounmanou
Department of Veterinary and Animal Sciences
Faculty of Health and Medical Sciences
University of Copenhagen, Denmark

Principal supervisor: Professor Anders Dalsgaard
Department of Veterinary and Animal Sciences
Faculty of Health and Medical Sciences
University of Copenhagen, Denmark

Primary co-supervisor: Professor John Elmerdahl Olsen
Department of Veterinary and Animal Sciences
Faculty of Health and Medical Sciences
University of Copenhagen, Denmark

Co-supervisor: Professor Robinson Mdegela
Department of Veterinary Medicine and Public Health,
College of Veterinary Medicine and Biomedical Sciences,
Sokoine University of Agriculture, Tanzania

Assessment committee

Professor Hanne Ingmer
Department of Veterinary and Animal Sciences
Faculty of Health and Medical Sciences
University of Copenhagen, Denmark

Professor Iruka N. Okeke
Department of Pharmaceutical Microbiology
Faculty of Pharmacy
University of Ibadan, Nigeria

Associate Professor François-Xavier Weill
Enteric Bacterial Pathogens Unit
Institut Pasteur, Paris, France

Preface

For cholera, one of the world's oldest life-threatening human diseases, there is a global estimate of 1.4 to 4.3 million cases with about 21,000–143,000 deaths per year and huge economic consequences on affected countries. Its etiologic agent, *Vibrio cholerae,* is subsequently one of the well-described human pathogens in scientific research. Nevertheless, there is much to be investigated regarding transmission and evolution of cholera and the world is still unable to precisely predict the occurrence of cholera outbreaks to proactively propose preventive measures. By analogy, the interplay between person-to-person and environmental transmission still needs to be clarified, especially in the context of the evolutionary response of *V. cholerae* to environmental and host-driven selective pressures. Like many other countries in sub-Saharan Africa, Tanzania has consistently been affected by cholera and constitutes one of the seven African countries reporting the highest number of cholera cases in the past decade, therefore, calling for meticulous country-level characterization of cholera and *V. cholerae*. In this context, the studies described in this Thesis became part of a research project titled "Innovations and Markets for Lake Victoria Fisheries" (IMLAF). The IMLAF project was a DANIDA (Danish Development Assistance) funded project, carried out in the Lake Victoria basin of Tanzania. Its main objective was to catalyze socio-economic growth through expansion of markets for quality fish products and increased employment opportunities in the fisheries industry. The work package of the project that focused on characterization of microbial hazards in fish products encompassed pathogens such as *Salmonella, E. coli* and *V. cholerae,* with the latter thoroughly described in the present Thesis. This Thesis is structured around four manuscripts, each one representing one of the four specific objectives, all summarized in one general discussion. It ends with recommendations for policy to guide cholera control and future research needs.

Acknowledgements

I am grateful to the Danish Development Assistance for sponsoring me with grant DFC File No. 14-P01-TAN through the "Innovations and Markets for Lake Victoria Fisheries" project. I would also like to acknowledge the International Foundation for Sciences (IFS) for the individual grant number 1-2-A-6100-1 allocated to me for completion of the study. My gratitude also goes to the Research Group of Professor Anders Dalsgaard in the Section of Food Safety and Zoonosis for funding my enrollment at the University of Copenhagen to carry out this study.

My profound gratitude to my principal supervisor and friend, Anders Dalsgaard, for his mentorship and fatherhood and for our unique cooperation, always creating an interesting balance between professional relationship and social life. I hereby extend my recognition to his family for being my Danish family away from home. *Mange tak!*

I sincerely thank my co-supervisor, John Olsen, for his advice, guidance and fatherly attention throughout the realization of this study. I also extend my gratitude to Rene Hendriksen, Shinny Leekitcharoenphon, Jeffrey Withey, Henry Madsen, Kåre Mølbak and Victorien Dougnon for their guidance during the experiments. My stays in Tanzania would not have went seamlessly without my local co-supervisor Robinson Mdegela, to whom I would like to say *Asante sana!*

To my mother and brothers in Benin and my late father, I cannot thank you enough for being part of my life. We are in this world to thrive and together, we are getting there. *Akpé kaka!*

I would also like to express my gratitude to my friends, my colleagues of the IMLAF project in Tanzania, my colleagues from the Department of Veterinary and Animal Sciences at Copenhagen University, and all my study contributors from the Ministry of Health in Tanzania, Sokoine University of Agriculture and the National Fish Quality Control Laboratory in Mwanza. I thank you all!

Table of contents

Summary

With the currently available scientific knowledge and tools to guide prevention and treatment, every death due to cholera can be avoided. Cholera, however, continues to take a heavy toll in developing countries especially in Sub-Saharan Africa, where it has become recurrent for the past forty years favored by poverty, poor hygiene and the vulnerability of populations living around lakes and those affected by conflicts or natural disasters. To eliminate cholera from 20 of the 48 endemic countries in the world, the World Health Organization (WHO) has established a global roadmap to 2030 stipulating a rigorous focus on surveillance and source tracking of cholera outbreaks at country level.

Tanzania has been reporting cholera almost every year since 1974 and is one of the endemic countries on the list of the WHO. This Thesis intends to contribute to the pool of knowledge that will enable Tanzania to get rid of the cholera epidemics by 2030. The study was carried out using cholera surveillance data, clinical *V. cholerae* isolates from cholera patients and toxigenic *V. cholerae* recovered from fish (*Rastrineobola argentea*), phytoplankton and water from Lake Victoria in both dry and rainy seasons. The study also involved aquaria experiments determining if tilapia (*Orechromis nilioticus*) is a reservoir host and plays a role in the transmission of *V. cholerae*. We applied genomic and epidemiological tools as well as microbial ecology approaches regarding the key element of the global roadmap, which is to investigate how cholera emerges, evolves and reemerges in the Tanzanian context and how *V. cholerae* strains persist during and between outbreaks.

Overall, spatial analyses revealed that cholera hotspots in Tanzania include populations living near lakes and central regions, being areas where interventions should be prioritized for long-term cholera control. Moreover, we demonstrated that the endemicity of cholera in Tanzania is due to the existent rampant outbreak foci in the Eastern African region as a result of cross-border transmission of toxigenic *V. cholerae* O1 of three different sub-lineages of genotypes *ctx*B3 (T5), *ctx*B1 (T10), and *ctx*B7 (T13)

emanating from the same 7PET lineage. Toxigenic *V. cholerae* O1 recovered from Lake Victoria were phylogenetically related to pandemic strains causing outbreaks in Tanzania and other East-African countries with as low as three SNPs and identical accessory genome contents were observed between strains from the two niches. We further report the existence of environmental reservoirs such as fish and phytoplankton favoring survival and transmission of pandemic *V. cholerae* O1 in Lake Victoria, which offers optimum temperature and alkaline pH that enhance persistence and resurgence of the pathogen throughout the year. Antimicrobial resistance was surprisingly limited in the studied strains, especially among the T13 sub-lineages causing recent epidemics. These were mostly susceptible, with 10-kb nucleotide deletions on the SXT/ICE that normally encodes resistance to many antibiotics, including sulfamethoxazole and trimethoprim. The environmental strains, however, harbored genes for resistance to heavy metals, a phenotype that could support their growth and survival in aquatic environments. The *V. cholerae* O1 strains also harbor autoinducers (AI-2 *Lux*P and *Lux*Q) involved in quorum sensing and biofilm formation for their environmental fitness until resurgence.

If Tanzania is to become free of cholera by 2030, following suggested actions in the global roadmap for cholera control, there is a lot to be done politically based on the scientific insights provided in this Thesis. Authorities need to mobilize more resources and demonstrate more political determination towards implementation of hygiene and sanitation programs, prioritizing high-risk areas including the lake zones. There is also a need for more investment in accurate and holistic reporting of cholera in order to provide reliable data for action. Most importantly, the genetic fingerprints of strains causing outbreaks in Tanzania and its neighboring countries reveal unequivocally a cross-border transmission of similar clones. This finding calls for coordinated collaborative control measures, such as early vaccinations in the entire African Great Lakes Region, when a neighboring country declares an outbreak. Furthermore, the presence of environmental reservoirs (like fish and phytoplankton) for *V. cholerae* favoring

resurgence of the pathogen between outbreaks in Tanzania should be considered when designing control strategies in the region. This is important because, despite the cross-border evidence of human-mediated transmission of *V. cholerae*, the inability of the pathogen to survive long-term in acid conditions inside the human gut, and especially the high infectious dose required for disease, continuous transmission by humans alone is likely impossible and therefore very dependent on environmental reservoirs that refeed the cycle.

Sammendrag (Danish Summary)

Vi har i dag tilstrækkelig videnskabelig baseret viden og redskaber til effektivt at forebygge og behandle kolera; mennesker bør således ikke dø af kolera. Trods dette så fortsætter kolera med at ramme de fattigste lande hårdt, især lande syd for Sahara, som løbende har været ramt af kolera gennem de sidste 40 år. Dette skyldes især fattigdom, dårlige sanitære forhold og modtagelige befolkninger, som lever omkring store søer eller er påvirket af konflikter og naturkatastrofer. Verdenssundhedsorganisationen, WHO, har lavet en global plan frem mod år 2030 til elimination af kolera i 20 af i alt 48 lande, hvor kolera optræder endemisk. Overvågning og smitteeftersporing af koleraudbrud på landeniveau udgør et hovedelement i denne plan.

Tanzania har rapporteret kolera næsten hvert år siden 1974 og kolera er endemisk i landet ifølge WHO´s landeliste. Denne ph.d. afhandling tilstræber at bidrage med viden, som skal gøre Tanzania i stand til at forhindre, at landet ikke længere bliver ramt af koleraepidemier efter år 2030. I undersøgelserne, der er beskrevet i afhandlingen, er der anvendt kolera overvågningsdata, kliniske *V. cholerae* isolater fra kolerapatienter og toksigene *V. cholerae* isolater indsamlet fra fisk (*Rastrineobola argentea*), fytoplankton og vand i Victoria søen i regnvejrs- og tørre perioder af året. Der blev også foretaget akvarieeksperimenter for at fastlægge om fisken tilapia (*Orechromis nilioticus*) kunne være en reservoirvært og spille en rolle i overførslen af *V. cholerae*. Der blev anvendt molekylærgenomiske og epidemiologiske værktøjer, samt mikrobiel økologiske undersøgelser, til at belyse hovedelementer i WHO´s koleraplan, som fokuserer på at forstå, hvordan kolera og kolerabakterien opstår, udvikler sig og genopstår med reference til tanzanianske forhold, samt en generering af viden om, hvordan *V. cholerae* bakterier persisterer under og mellem koleraudbrud.

De spatiale analyser viste, at hotspots for kolera i Tanzania inkluderer befolkningsgrupper, som lever nær søer og i de centrale områder af landet; områder som fremadrettet bør prioriteres i mere

langsigtede tiltag til kontrol af kolera. Vi påviste også, at den endemiske kolerasituation i Tanzania især skyldtes introduktion af sygdommen fra områder med koleraudbrud i flere øst-Afrikanske naboområder. På tværs af landegrænser forekommer der således en transmission af toksigene *V. cholerae* O1, som tilhører en af de tre forskellige undergrupperinger af genotyperne *ctx*B3 (T5), *ctx*B1(T10), og *ctx*B7 (T13), som oprinder fra den samme 7PET linje af stammer. Toksigene *V. cholerae* O1 indsamlet fra Victoria søen var fylogenetisk meget tæt relateret til pandemiske bakteriestammer, som har forårsaget koleraudbrud i Tanzania og andre øst-afrikanske lande, idet der kun blev fundet tre SNPs (”single nucleotide polymorphisms”) og identiske tilhørende indhold af genomer mellem isolater fra de to undergrupperinger af genotyper. Vi påviste også eksistensen af et miljøreservoir i fisk og fytoplankton, som fremmer overlevelse og overførsel af pandemisk *V. cholerae* O1 i Victoria søen; et akvatisk system, som repræsenterer en optimal temperatur og et alkalisk pH, som øger persistens af *V. cholerae* O1 i gennem hele året. Forekomsten af antimikrobiel resistens var forbavsende begrænset i kolerabakterierne, især i de bakterier, som tilhører T13 undergrupperingen, som har forårsaget koleraepidemier fornyelig. Disse kolerabakterier, som er følsomme overfor antimikrobielle stoffer mangler et 10-kb stort nukleotid på SXT/ICE genet, som koder for resistens overfor mange antimikrobielle stoffer inklusiv sulfamethoxazol og trimethoprim. *V. cholerae* O1 fra miljøet indeholdt gener, som koder for resistens overfor tungmetaller, en fænotypisk egenskab som kan være gavnlig for vækst og overlevelse i akvatiske miljøer. *V. cholerae* O1 bakteriestammerne indeholdt også gener kodende for auto inducere (AI-2 *Lux*P and *Lux*Q), som er involveret i quorum sensing og dannelse af biofilm og derved øger stammernes fitness i miljøet indtil det tidspunkt bakteriecellerne skal vokse.

Hvis Tanzania i år 2030 skal være fri for kolera, som foreslået i den globale handlingsplan for kolera, så skal der gøres et stort stykke arbejde politisk med udgangspunkt i den viden, som er generet i denne ph.d.-undersøgelsen. Det er nødvendigt, at myndighederne mobiliserer flere ressourcer og demonstrerer

politisk vilje til at fremme implementeringen af vand- og sanitetsprogrammer med prioritet af høj-risiko områder, eksempelvis områderne omkring de store afrikanske søer. Det er også nødvendigt at øge investeringerne i mere præcise og holistiske rapporteringer af kolera, så der kan genereres pålidelige data at agere på. Det er vigtigt at understrege, at de genetiske fingeraftryk af kolerabakterierne, som forårsager udbrud i Tanzania og dets nabolande enstemmigt viser, at de samme kloner af kolerabakterier spredes på tværs af landegrænser, hvilket understreger nødvendigheden af en koordineret kontrolindsats mellem lande i hele området omkring de store afrikanske søer. Desuden bør tilstedeværelsen af miljøreservoirer (som fisk og planteplankton) for *V. cholerae*, der favoriserer genopblussen af patogene mellem udbrud i Tanzania, tages i betragtning, når man udformer kontrolstrategier i regionen. Humanmedieret transmission af *V. cholerae* mellem lande er dokumenteret, men patogenes manglende evne til langvarig overlevelse under syrebetingelser i menneskets mave og især den høje infektiøse dosis, der kræves for at forårsage sygdom gør, at en kontinuerlige transmission i og mellem mennesker alene ikke er muligt. *V. cholerae* O1 overlevelse og transmission er således afhængig af et miljøreservoir.

List of abbreviations

WHO:	World Health Organization
AGLR:	African Great Lakes Region
DRC:	Democratic Republic of Congo
7PET:	Seventh Pandemic El Tor (*Vibrio cholerae* O1)
WGS:	Whole Genome Sequencing
PFGE:	Pulsed-Field Gel Electrophoresis
RAPD:	Random Amplification of Polymorphic DNA
AFLP:	Amplified Fragment Length Polymorphism
MLST:	Multi-Locus Sequence Typing
VNTRs:	Variable Number of Tandem Repeats
MLVA:	Multi-Locus Variable Tandem Repeat
MVLST:	Multi-Virulence Locus Sequencing Typing
PCR:	Polymerase Chain Reaction
SNP:	Single Nucleotide Polymorphism
AMR:	Antimicrobial Resistance

Articles included in the Thesis

Manuscript A: "Cholera hotspots and surveillance constraints contributing to recurrent epidemics in Tanzania". **Yaovi M.G. Hounmanou**, Kåre Mølbak, Jonas Kähler, Robinson H. Mdegela, John E. Olsen, Anders Dalsgaard (**2019**). Article published as "short communication" in ***BMC Research Notes* 12,** 664 doi:10.1186/s13104-019-4731-0.

Manuscript B: "Surveillance and genomics of toxigenic *Vibrio cholerae* O1 from fish, phytoplankton and water in Lake Victoria, Tanzania". **Hounmanou YMG**, Leekitcharoenphon P, Hendriksen RS, Dougnon TV, Mdegela RH, Olsen JE and Dalsgaard A (**2019**). Article published in ***Frontiers in Microbiology*** 10:901; doi: 10.3389/fmicb.2019.00901.

Manuscript C: "Genomic insights into *Vibrio cholerae* O1 responsible for cholera epidemics in Tanzania between 1993 and 2017". **Yaovi Mahuton Gildas Hounmanou**, Pimlapas Leekitcharoenphon, Egle Kudirkiene, Robinson H. Mdegela, Rene S. Hendriksen, John Elmerdahl Olsen and Anders Dalsgaard (**2019**). Article published in ***PLoS Neglected Tropical Diseases*** 13(12): e0007934 doi.org/10.1371/journal.pntd.0007934.

Manuscript D: "Tilapia (*Oreochromis niloticus*) as a putative reservoir host for survival and transmission of *Vibrio cholerae* O1 biotype El Tor in the aquatic environment". **Hounmanou YMG**, Mdegela RH, Dougnon TV, Madsen H, Withey JH, Olsen JE and Dalsgaard A (**2019**). Article published in ***Frontiers in Microbiology*** 10:1215*;* doi: 10.3389/fmicb.2019.01215

Articles published during the PhD, which are not part of the Thesis

a) Tilapia lake virus threatens tilapiines farming and food security: Socio-economic challenges and preventive measures in Sub-Saharan Africa. **Hounmanou, YMG**, Mdegela, RH, Dougnon, TV, Achoh, ME, Mhongole, OJ, Agadjihouèdé, H, Gangbè, L & Dalsgaard, A (**2018**). ***Aquaculture***, 493, s. 123-129. https://doi.org/10.1016/j.aquaculture.2018.05.001

b) The impact of inactivation of the purine biosynthesis genes, *purN* and *purT*, on growth and virulence in uropathogenic *E. coli*. Andersen-Civil, Audrey Inge Schytz; Ahmed, Shahana; Guerra, Priscila Regina; Andersen, Thomas Emil; **Hounmanou, Yaovi Mahuton Gildas**; Olsen, John Elmerdahl; Herrero-Fresno, Ana. ***Molecular Biology Reports***, 45, Nr. 6, **2018**, s. 2707-2716. https://doi.org/10.1007/s11033-018-4441-z

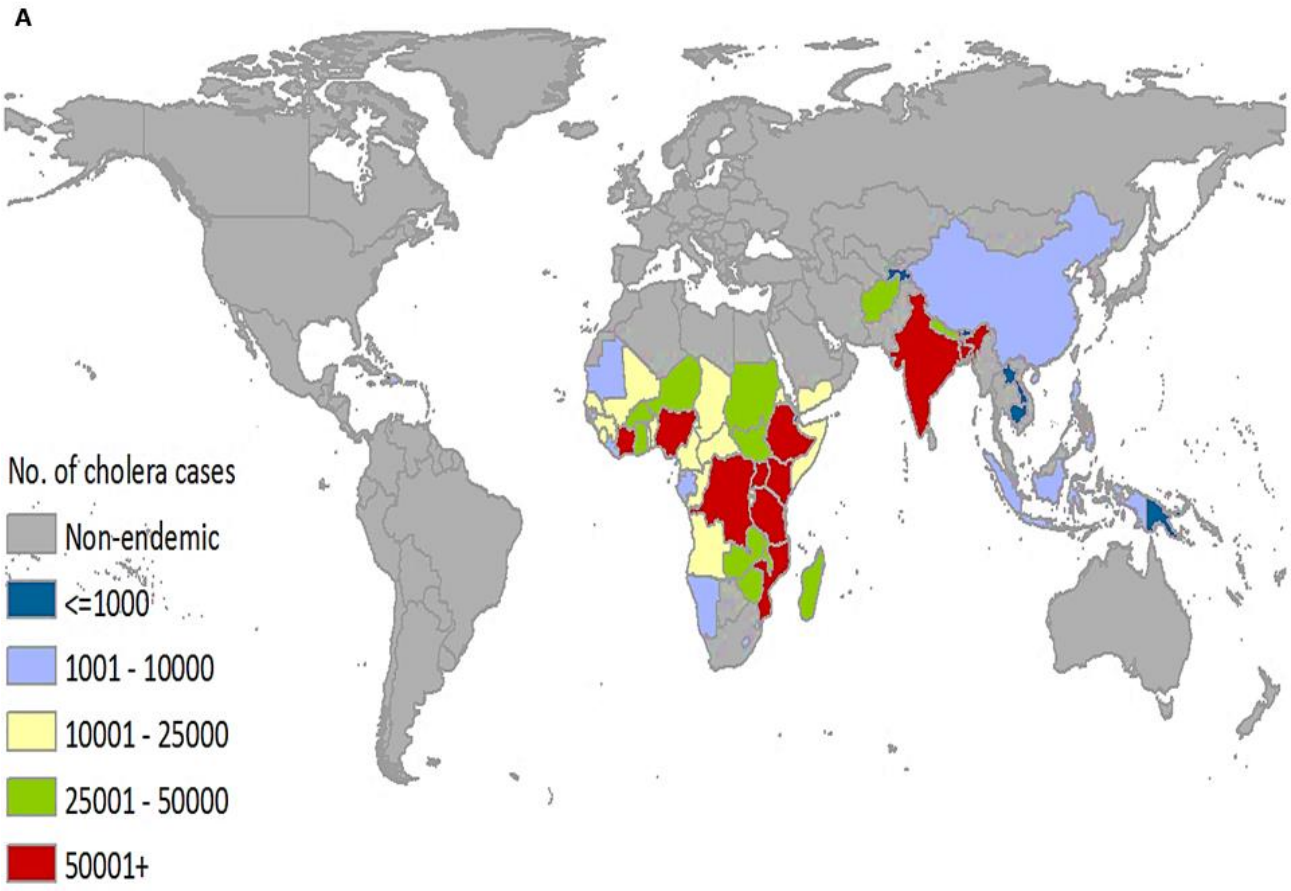
c) Molecular typing and antimicrobial susceptibility of Methicillin-Resistant Staphylococcus aureus isolated from bovine milk in Tanzania. Mohammed, Jibril; Ziwa, Michael Henry; **Hounmanou, Yaovi Mahuton Gildas**; Kisanga, Adela; Tuntufye, Huruma Nelwike. ***International Journal of Microbiology***, **2018**, 4287431, 2018. https://doi.org/10.1155/2018/4287431

d) Antimicrobial resistance among clinically relevant bacterial isolates in Accra: a retrospective study. Mohammed, Jibril; **Hounmanou, Yaovi Mahuton Gildas**; Thomsen, Line Elnif. ***BMC Research Notes***, 11, Nr. 1, 254, 2018. https://doi.org/10.1186/s13104-018-3377-7

Chapter 1. Introduction

1.1. Global dynamics of cholera in an African context

Cholera is an acute life-threatening diarrheal disease and remains a major health threat, particularly in developing countries of Africa, Asia, and Latin America, claiming about 4.3 million reported cases and 142,000 deaths worldwide every year (Ali et al., 2015; Griffith et al., 2006; Weil and Ryan, 2018). In the past 40 years, the burden of cholera has shifted from Asia to Africa (Fig. 1), which is currently the source of over 94% of reported global cases (Ali et al., 2015; Mengel et al., 2014; Reyburn et al., 2011). Sub-Saharan Africa bears, to date, the major impact of the global cholera burden (Lessler et al., 2018; Rebaudet et al., 2013; Weil and Ryan, 2018).



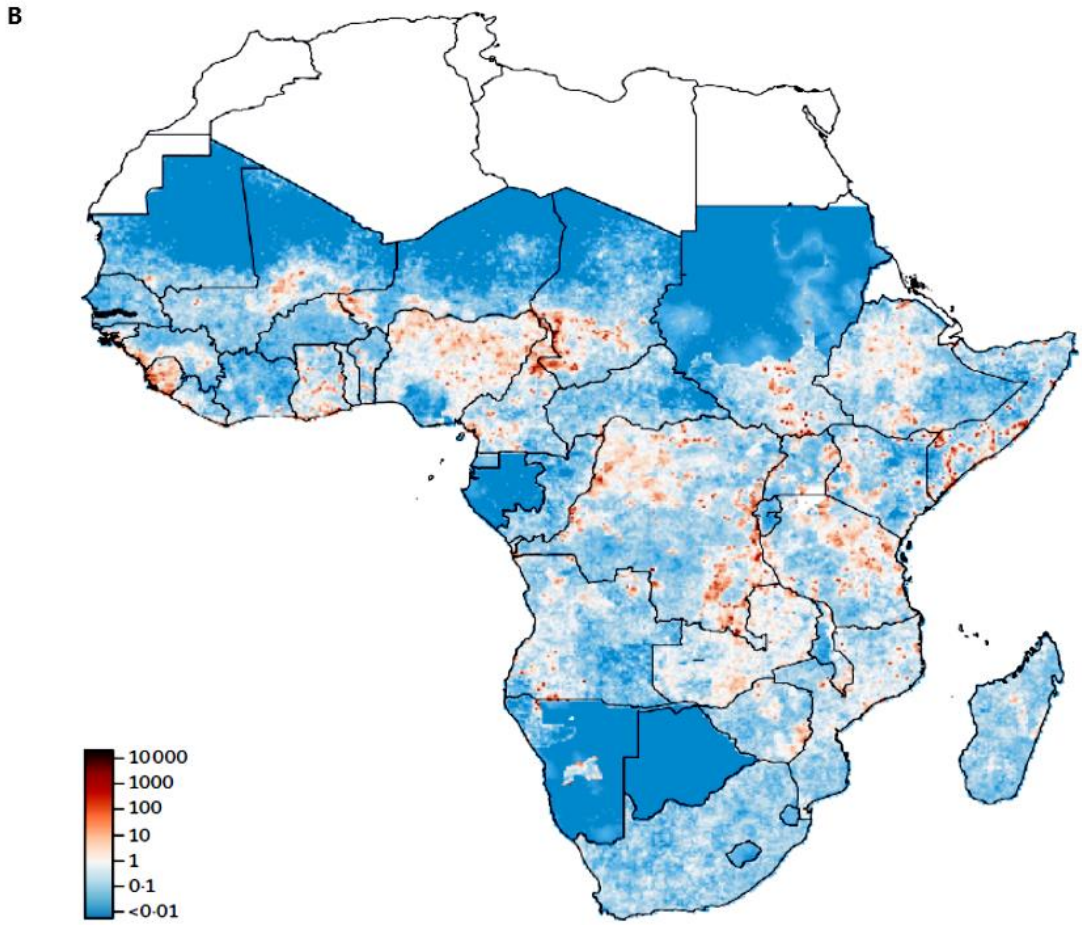


Figure 1. Annual cholera incidences. (A) Global annual number of cholera cases in endemic countries; (B) Average annual cholera incidence per 100,000 people in Sub-Saharan Africa between 2010 and 2016. Data adopted from previous studies (Ali et al., 2015; Lessler et al., 2018).

An estimated 2.8 million cholera cases occur each year in endemic countries, and the average global annual incidence rate is 2 cases per 1000 people at risk, with countries of highest incidence rates being in Africa and Southern Asia (Ali et al., 2015). In Sub-Saharan Africa, an annual average of 141,918 cholera cases are reported from only 4% of districts, which is home to about 88 million people (Lessler et al., 2018). Since 1817, cholera has spread from the Indian sub-continent and seven pandemics have been observed (Siriphap et al., 2017). The seventh pandemic is still ongoing and genomic studies have shown that this pandemic has spread from the Bay of Bengal in at least three independent but overlapping

waves (I, II and III) from a common ancestor in the 1950s, with several transcontinental transmission events (Mutreja et al., 2011). The classification into waves is based on genetic characteristics of the *V. cholerae* outbreak strains and is described in **manuscripts B** and **C** for the studied Tanzanian strains.

Figure 2, below, shows that outbreaks of wave I were more geographically dispersed globally but were replaced in the early 90s by wave II outbreaks, which seemed to be more geographically restricted to mostly Asia. Meanwhile, wave III outbreaks continue to occur and are reported in Asia, Africa and the Americas. In Africa, the seventh pandemic has been investigated and until 2015 researchers determined that the different epidemics could all be traced back to a single lineage, which has been introduced at least 11 times (T1 to T11) since the first epidemic in the 1970s (Weill et al., 2017). Additional analyses following the 2016-2017 Yemen outbreaks revealed two more transmission events, T12 and T13, involving strains currently circulating in East Africa (Weill et al., 2019). The "T" identity of the Tanzanian strains analyzed in this study is described in **Manuscript C**.

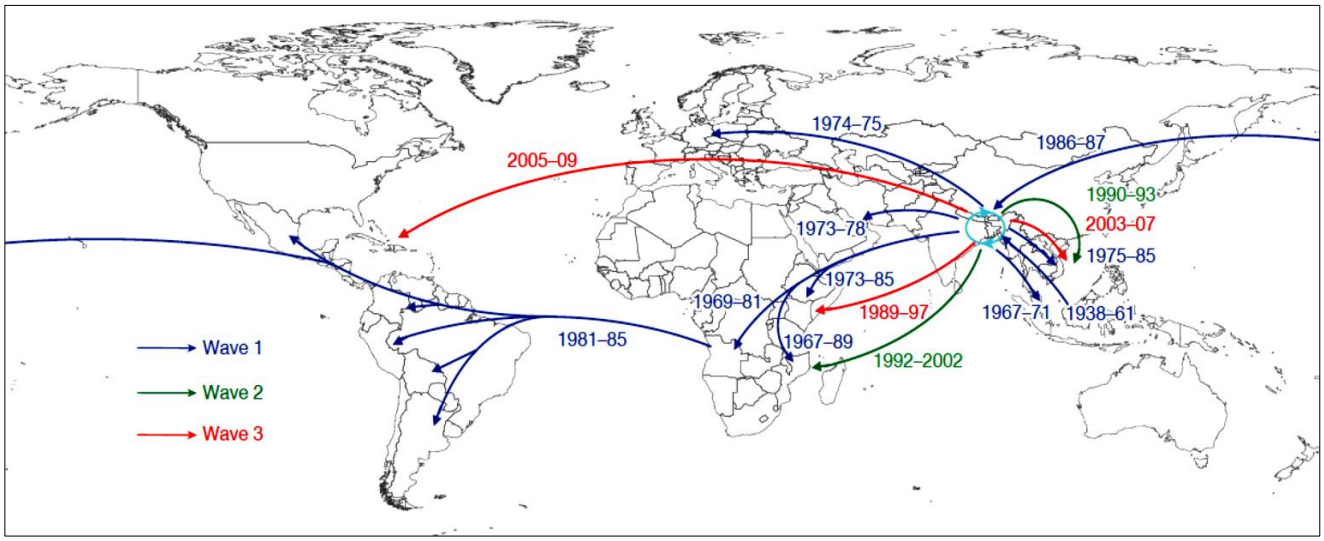


Figure 2. Global spread of the seventh cholera pandemic into three waves. Adopted from a previous report (Mutreja et al., 2011).

1.2. History and burden of cholera in Tanzania

More than 40 years after its resurgence in Africa in 1970, cholera remains a serious public health problem, characterized by large disease burden, frequent outbreaks, persistent endemicity, and high case fatality rates, particularly in the region of the African Great Lakes, which might act as reservoirs for cholera on the continent (Mengel et al., 2014; Rebaudet et al., 2013). After cholera emerged from the original reservoir in the Ganges delta of India in 1817, Tanzania was the first African country affected during the second pandemic in 1836 with *V. cholerae* O1 strains introduced through outbreaks on the Indian Ocean coast by ships transporting slaves, causing about 20,000 deaths in Zanzibar (Olago et al., 2007).

The contemporary ongoing seventh pandemic reached Tanzania in 1974 when cholera occurred in Kyela district on the shores of Lake Nyasa bordering Malawi (Mbwette, 1987). The outbreak lasted only one month and killed seven out of the ten notified cases. The next episode was reported in 1977 when a Tanzanian resident of Rufiji district in the coastal region died of cholera after hosting a merchant from Asia where outbreaks were occurring at the time (WHO, 2018b). Between October 1977 and April 1980, the outbreak spread to about 18 regions of the country including Kilimanjaro, Dar-es-Salaam, Dodoma, Kigoma, and Mwanza (Mbwette, 1987). It was during this outbreak that Zanzibar was affected by the seventh pandemic. The first cases were identified in 1978 around fishing villages on Unguja island and claimed 411 cases, including 51 deaths (WHO, 2018b). In 1983, further outbreaks were reported on both islands of Zanzibar (Pemba and Unguja) before the largest outbreaks of Zanzibar were recorded in 1997, 1998 and 2004 with 520 to 650 cases per outbreak (WHO, 2018b). In 2006, Zanzibar experienced another large outbreak, which started in South Pemba in March, with the first case originating from Mkoani district. The outbreak spread a week later to Unguja island and by July 2006, a total of 315 cases including eight deaths were reported (WHO, 2018b). In Tanzania as a whole, and over the decade following the

outbreak of 1977, seasonal relapses of cholera were reported across the country both in mainland and Zanzibar until the first major outbreak occurred in 1992 when 18,526 cases including 2,173 deaths were recorded countrywide (WHO, 2018b).

The second major countrywide cholera outbreak occurred in 1997. The epidemic started at the end of January in Dar es Salaam and accounted for 40,249 cases and 2,231 deaths with seven regions affected. Between 2002 and 2006, most Tanzanian regions have reported cholera cases and nine of them, including Zanzibar, Dar es Salaam and Dodoma reported more than 2000 cases each during this five years period (WHO, 2018b). In 2006, Tanzania reported up to 14,297 cholera cases with 254 deaths (WHO, 2018b).

Since the seventh cholera pandemic reached the country in 1974, Tanzania reports outbreaks almost every year with over 250,000 cases and 13,078 deaths reported up until 2018 (Lessler et al., 2018; WHO, 2018a). Cholera has become rampant in Tanzania and its islands where very frequent cases with high case fatality rates occur every year. An analysis of 11 years (2007-2017) of cholera data from mainland Tanzania in **Manuscript A** reveals 39,444 cholera cases with 600 deaths, giving a case fatality rate of 1.5% and an average annual incidence rate of 8.39 per 100,000 population with a strong correlation between living near lakes and the increase in cholera incidence (Fig. 3). In Zanzibar, from 1997 and 2017, approximately 11,921 cholera cases were reported, representing an average annual incidence rate of 4.4 per 10 000 (Bi et al., 2018).

The historical dynamics of cholera outbreaks show a significant impact of imported cases mainly from Asia and other African countries due to political unrest and population displacements like the 2015 outbreak that occurred in Kigoma and Kagunga refugee camps in Tanzania among Burundian refugees and accounted for 3,000 cases including 31 deaths (Kachwamba et al., 2017; WHO, 2018a, 2018b). There is also a notable endemicity around the lake zones and Zanzibar as well as city slums like in Dar es

Salaam and central regions due to anthropogenic activities (**Manuscript A**). Also, seasonal outbreaks have often been seen during the seventh cholera pandemic in Tanzania (Bi et al., 2018; Reyburn et al., 2011).

It is however perceived that the number of cholera cases in Africa in general and in Tanzania in particular could possibly be much higher than what is reported to the WHO (Azman et al., 2019). Reasons include underreporting and inadequate surveillance systems as described in **Manuscript A**, and the lack of healthcare facilities in remote areas, which underlines a number of home-based cases that feed the epidemic cycle significantly (Griffith et al., 2006; Mengel et al., 2014). In addition, many countries are reluctant to conduct cholera surveillance or to officially report cholera cases or deaths since embargos from trading partners and tourism losses usually follow the confirmation of cholera outbreaks (Azman et al., 2019). This is because international organizations put restrictions on produce and other imports from countries reporting cholera, which significantly harm their economy through exports, private consumption, consumer prices, employment, and the overall GDP (Lonappan et al., 2019). Such stigma and associated fear of loss of economic opportunities lead to a lack of adequate disease reporting and constitute some of the challenges hampering effective cholera control programs.

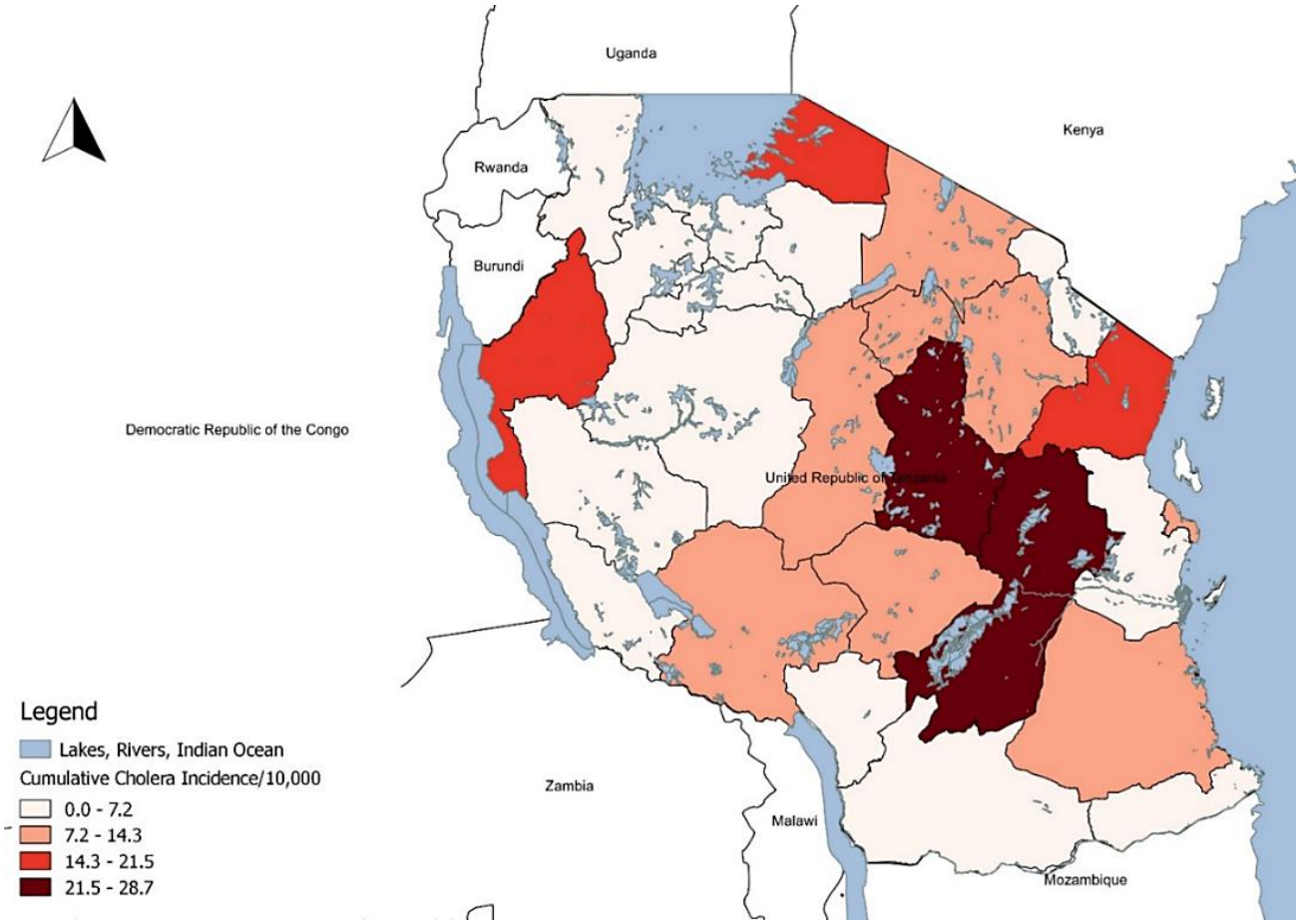


Figure 3. Spatial distribution the average number of cholera incidences per 10,000 people over a period of 11 years (2007-2017) in mainland Tanzania. Figure produced in QGIS using data described in **Manuscript A**.

1.3. Molecular biology of *Vibrio cholerae* with emphasis on *V. cholerae* O1

Vibrio cholerae, the ethiological agents of cholera, are Gram-negative comma shaped Gamma-proteobacteria of the Family *Vibrionaceae* and described as autochthonous of aquatic environments (Islam et al., 2019). They are classified into over 200 serogroups based on their O-antigen (Shimada et al., 1994). However, only the serogroups O1 and O139 are recognized to cause cholera and are involved in pandemic cholera (Colwell et al., 2003; Faruque et al., 1998). The O1 serogroup is categorized into three serotypes: Ogawa, Inaba and the Hikojima (a variant of the Ogawa serotype). This classification is

however outdated and rarely used in current literature (Shimada et al., 1994). While *V. cholerae* O139 seemed to have been reported only in Asian epidemics (Chun et al., 2009; Faruque et al., 2003), the non-O1/non-O139 strains are mostly nonpathogenic and commonly encountered in the aquatic environments during non-epidemic periods, although some have been associated with sporadic cholera-like infections (Dalsgaard et al., 2001; Fang et al., 2019; Haley et al., 2014). Regardless of their serogroup, the ability of *V. cholerae* to cause cholera depends upon the main virulence factor, the cholera toxin encoded by the *ctxAB* operon, which resides in the genome of a filamentous bacteriophage (CTX) specific to toxigenic *V. cholerae* strains (Aliabad et al., 2012; Bakhshi et al., 2008; Sanchez and Holmgren, 2011).

The CTX prophage of *V. cholerae* is a key element in the genetic makeup of toxigenic *V. cholerae* due to its role in virulence and diversity of the pathogen (Faruque et al., 1998; Rashid et al., 2016). The genome of the CTX prophage (Fig. 4) contains a 4.5 kb central core region composed of the *ctxAB*, *zot*, *ace*, *orfU* and *cep* genes, flanked by repetitive sequences of 2.4 kb (RS2) and 2.7kb (RS1) (Bakhshi et al., 2008). The A subunit of the cholera toxin gene (*ctxA*) is the enzymatic component of cholera toxin, having ADP-ribosylating activity, while the B subunit (*ctxB*) forms a pentamer that binds to GM_1 ganglioside receptors on the epithelial cell surface and delivers the A subunit into the cell. The B subunit is also used in genotyping of the bacteria (Li et al., 2019; Mutreja et al., 2011; Naha et al., 2012; Sanchez and Holmgren, 2011). The *ctxB* genotypes of *V. cholerae* strains that have caused various cholera outbreaks in Tanzania are described in **Manuscript C**.

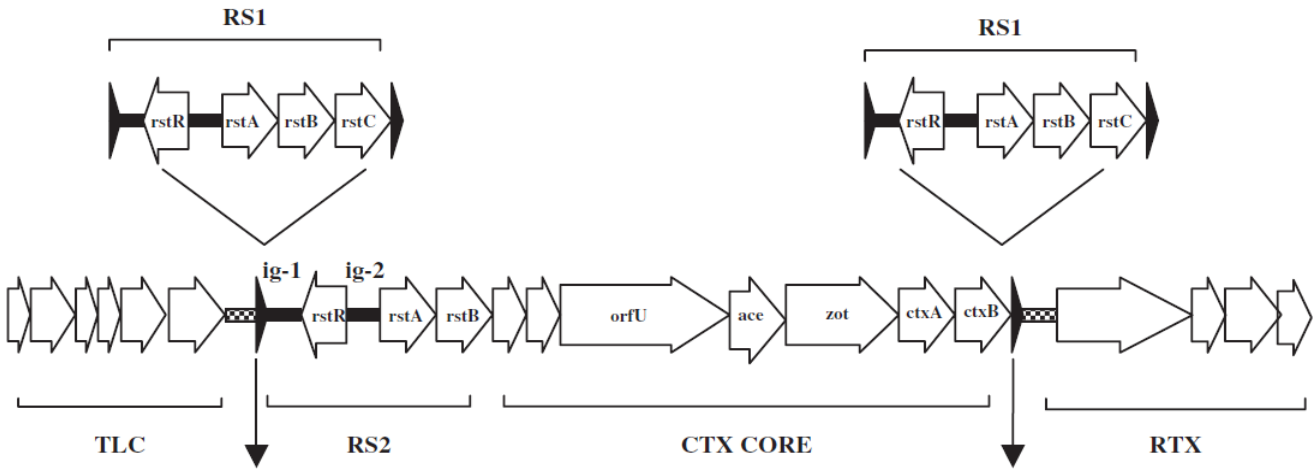


Figure 4. Genomic organization of the CTX prophage in *Vibrio cholerae*. Data adopted from previous description on *V. cholerae* strain N16961 (Bakhshi et al., 2008).

Compared to most bacteria that have a single chromosome, *V. cholerae* contain two chromosomes of approximately 4 Mb total (Fig. 5.) They consist of a large chromosome of about 3 Mb (chromosome 1) and a small chromosome of 1 Mb (chromosome 2) containing the core genes and genetic islands (Heidelberg et al., 2000; Mutreja and Dougan, 2019). These genetic islands or mobile genetic elements play pivotal epidemiological roles in cholera. The main ones are the Vibrio Seventh Pandemic Islands I and II (VSPI and VSPII), TCP, Vibrio Pathogenicity Islands (VPI-II) and the CTX prophage, implicated in pathogenesis and evolution of the bacteria (Chun et al., 2009; Dziejman et al., 2005; Faruque and Mekalanos, 2012). Variations in the mobile elements such as in the prophages, transposable elements, and potentially in plasmids, underline the diversity of the *V. cholerae* genome (Dutilh et al., 2014). In this Thesis, the genetic islands of *V. cholerae* from clinical and environmental origins have been characterized and described in **Manuscripts B** and **C.** Furthermore, one of the key mobile genetic elements in *V. cholerae* is the SXT, a self-transmissible integrating conjugative element that encodes antimicrobial resistance, including resistance to sulfamethoxazole and trimethoprim (Spagnoletti et al., 2014; Wang et al., 2016). Antimicrobial resistance in *V. cholerae* due to the SXT/ICE is described in the

studied Tanzanian clinical and environmental strains in **Manuscripts B** and **C** along with plasmid-mediated acquired resistance genes and resistance to quinolones due to chromosomal mutations.

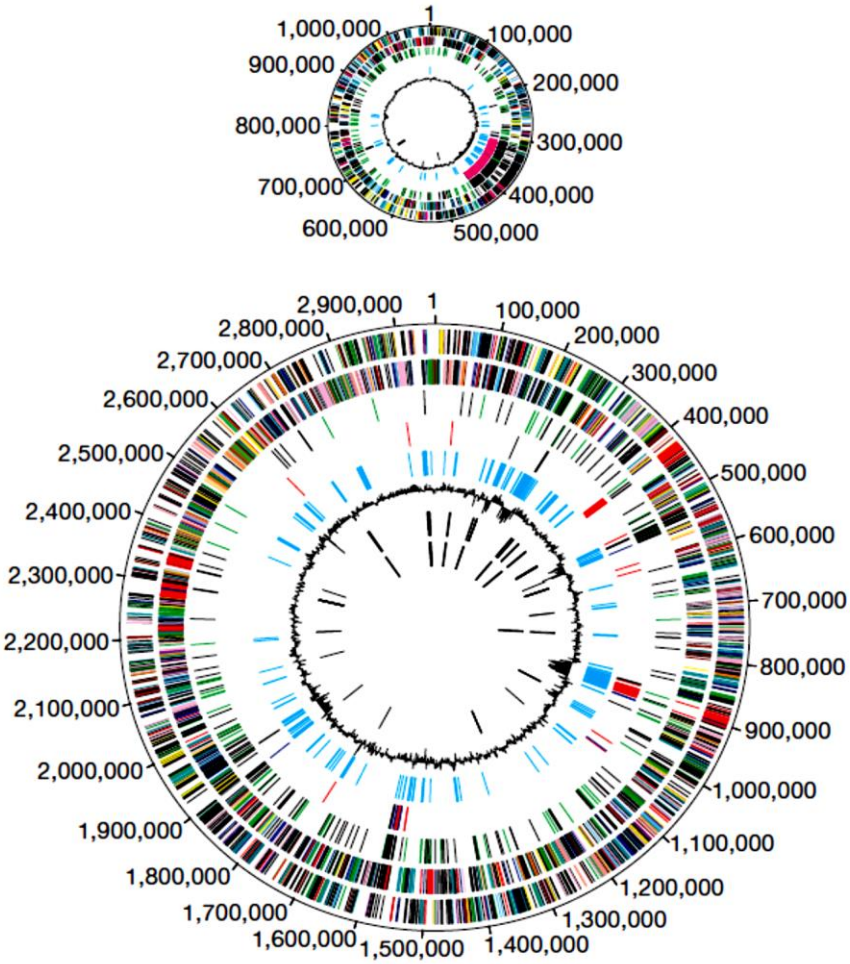


Figure 5. Representation of the two chromosomes of *Vibrio cholerae*. Data adopted from a previous study (Heidelberg et al., 2000).

V. cholerae O1 is sub-divided into two biotypes, the Classical and the El Tor biotypes (Beyhan et al., 2006; Mohammadi barzelighi et al., 2016) based on differences in their phenotypic and genotypic characteristics, pathogenic potential, infection modes and human survival abilities (Brumfield et al., 2018; Li et al., 2019; Mohammadi barzelighi et al., 2016). After their first description in 1817, *V.*

cholerae have caused to date seven different pandemics worldwide and the Classical biotype of *V. cholerae* O1 is believed to be responsible for the first six (Devault et al., 2014; Echenberg, 2011). The seventh and ongoing pandemic started in 1961 and is attributed mainly to *V. cholerae* O1 biotype El Tor with the *ctx*B3 genotype (Mutreja et al., 2011; Naha et al., 2013; Safa et al., 2010). This biotype has however evolved, leading to the emergence of El Tor variants known as "atypical El Tor" or "hybrids", through the acquisition of the *ctxB* genes from the classical biotype CTX prophage by the El Tor strains (Kim et al., 2015; Naha et al., 2012; Rashed et al., 2013). These atypical variants are the most common strains causing contemporary cholera outbreaks worldwide and are believed to produce more cholera toxin, resulting in a more severe disease than the prototype El Tor (Ghosh-Banerjee et al., 2010; Rashid et al., 2016). Two genotypes, *ctx*B1 and *ctx*B7, are found within the atypical El Tor variants (Table 1) and both are leading all contemporary outbreaks around the globe (Kim et al., 2014, 2015; Pal et al., 2017). As summarized in Table 1, all *V. cholerae* reported to date have been classified into three waves, Wave I, II and III based on their genotypes with the wave III sub-divided into early and current wave III (Mutreja et al., 2011).

Table 1. Genetic diversity of pandemic *V. cholerae* populations. Data adopted from previous studies (Kim et al., 2015; Safa et al., 2010).

***V. cholerae* strains**	**Years of occurrence**	**Pandemic**	***ctx*B genotype**
[a]Vc-O1, Classical biotype	1817 – 1923	First six pandemics	*ctx*B1
Vc-O1, pre-seventh pandemic El Tor	1923 – 1961	Pandemic free period	*ctx*B1, *ctx*B2
Vc-O1, El Tor Wave I (El Tor)	1961 – 2000s	Seventh pandemic	*ctx*B3
Vc-O139	90s – Present	Seventh pandemic	ctxB3, ctxB1
Vc-O1, El Tor Wave II/early Wave III (atypical El Tor)	90s – Present	Seventh pandemic	*ctx*B1
Vc-O1, El Tor current Wave III (atypical El Tor/ Haitian variant)	2000s – Present	Seventh pandemic	*ctx*B7

[a] *V. cholerae* serogroup O1

1.4. Environmental reservoirs of *V. cholerae* and associated ecological factors in the AGLR

Vibrio cholerae are acknowledged as autochthonous organisms of the aquatic environments where they are reported to be associated with a variety of living organisms of the aquatic fauna and flora (Colwell et al., 1977; Vezzulli et al., 2010). The environmental reservoirs of *V. cholerae* are well-established in marine and estuarine waters, but not much is known about freshwater reservoirs such as Lake Victoria and Lake Tanganyika, the main great lakes on the African continent (Bwire et al., 2018; Hounmanou et al., 2019a; Nkoko et al., 2011; Rebaudet et al., 2013). The AGLR represents the leading cholera hotspot in Africa, substantiating that the lakes may play an important role as reservoirs for *V. cholerae*, leading to the increase in cholera incidence in the region (Lessler et al., 2018; Moore et al., 2015; Rebaudet et al., 2013). However, the role of the aquatic environment in the emergence of cholera also depends on ecological factors including temperature and pH, which are optimal for *V. cholerae* in tropical waters (Plisnier et al., 2015; Stoltzfus et al., 2014). It has been demonstrated that salinity levels

of about 15%, water temperatures between 25° C and 30° C, and alkaline pH around 8.5 are favorable for the growth and survival of *V. cholerae* in the aquatic environment (Huq et al., 1984). In **Manuscript B**, these parameters were measured and discussed for Lake Victoria in relation to occurrence of *V. cholerae*. The link between cholera, the aquatic environment and climate, known as the "cholera paradigm" is well-established in the AGLR because most epidemics occur in lakeside areas, where the weekly incidence of cholera varies by season, rainfall and phytoplankton blooms (Nkoko et al., 2011).

Various hypotheses have been studied to understand the mechanism of persistence of *V. cholerae* during inter-epidemic periods, including continuous human transmissions and animals as reservoirs. The aquatic environmental reservoirs have been the most conclusive, due to the natural aquatic habitat of the pathogen (Islam et al., 2019). During inter-epidemic periods, *V. cholerae* live in the aquatic environment in association with cyanobacteria, phytoplankton, water hyacinths, free-living amoebae, copepods, blue crabs, and marine bivalves, as well as fish (Fig. 6) and disease outbreaks are triggered by seasonal changes in the environment (Hounmanou et al., 2019b; Islam et al., 2019; Vezzulli et al., 2010). Phytoplankton and zooplanktons are the main aquatic reservoirs of *V. cholerae* and have been the subject of many investigations. They release considerable amounts of organic carbon into the environment during photosynthesis, along with high nutrient concentrations, which favor the growth of *V. cholerae* in optimum temperatures and salinity (Islam et al., 2015; Tamplin et al., 1990). *V. cholerae* also feed on chitin, which is the most abundant polysaccharide and the principal component of the zooplankton exoskeleton that supports the growth of large populations of *V. cholerae* (Constantin de Magny and Colwell, 2009).

During stressful conditions in the aquatic environment, *V. cholerae* strains transition into a dormant stage known as the viable but non-culturable (VBNC) state to survive and then revert into a culturable and infectious form when conditions become more optimal, including salinity, pH, temperature and

proper nutrients (Fig. 6) (Lutz et al., 2013; Vezzulli et al., 2010, 199). Classic bacteriological culture methods usually fail to detect environmental samples containing such VBNC strains and in **Manuscript B,** we have described a more sensitive methodological approach for the detection of these strains, which can play an important role in studying the epidemiology of cholera. The environmental non-toxigenic non-O1/nonO139 strains can in some cases acquire the CTX phage in the aquatic environment and convert to the toxigenic stage (Faruque and Mekalanos, 2012), underlining that environmental strains could be progenitors of new outbreak strains. In **Manuscript B** and **C**, we reported close phylogenetic relationships between clinical outbreak strains and *V. cholerae* from the aquatic environment, mainly from fish and phytoplankton obtained from Lake Victoria. The ability of *V. cholerae* to form biofilms through quorum sensing is also a well-documented mechanism by which *V. cholerae* live and persist in the aquatic environment (Kamruzzaman et al., 2010; Lutz et al., 2013).

Chironomids, also called non-biting midges, endemic in most sub-Saharan freshwater ecosystems including Lake Tanganyika and Lake Victoria (Armitage et al., 1995; Broza et al., 2005; Halpern et al., 2004) are another potential reservoir for *V. cholerae*. The egg masses of chironomids are laid in nutrient-rich edges of the lakes, and embedded in a gelatinous layer, which *V. cholerae* utilizes as source of carbon for survival (Broza et al., 2005; Halpern et al., 2004). The ability of *V. cholerae* to persist in the environment and cause epidemics on a seasonal basis has made them one of the few pathogens that have survived the major three niche dimensions such as space, time and habitat (Dutilh et al., 2014).

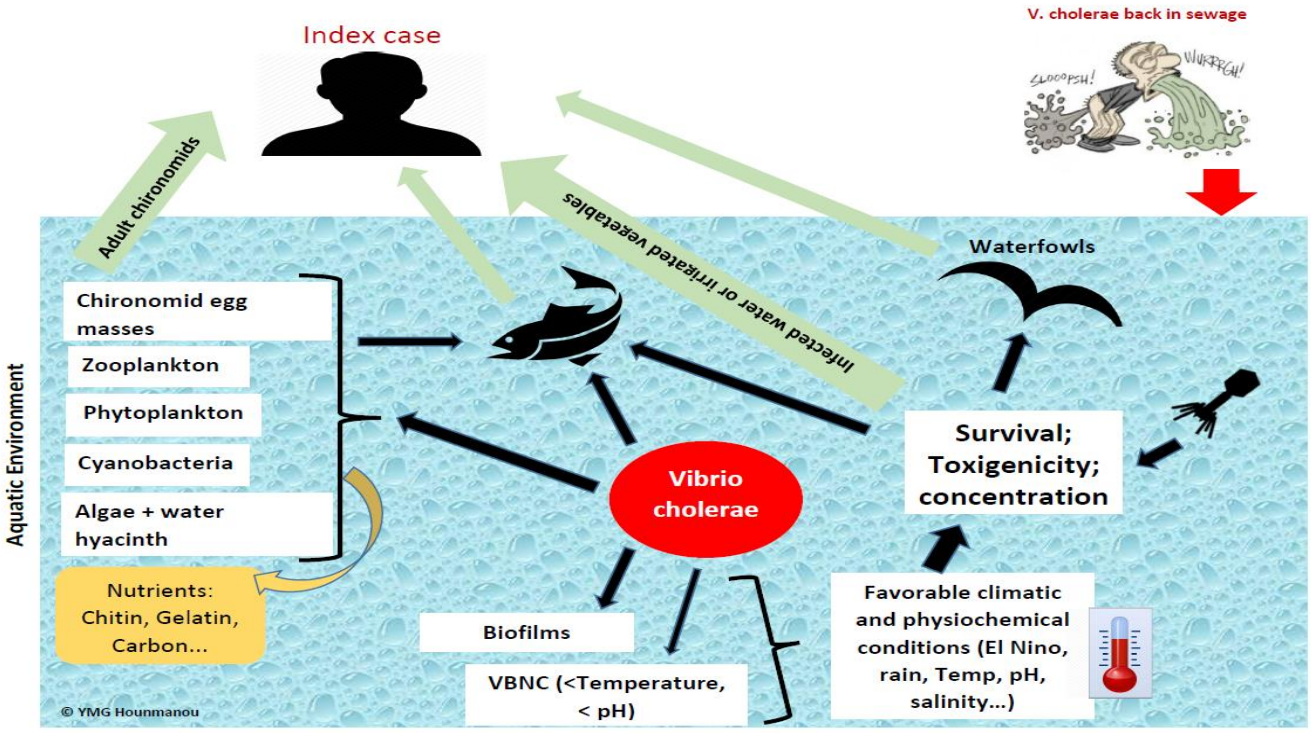


Figure 6. Schematic representation of interactions between *V. cholerae* and aquatic reservoirs in the environment and infected cholera patients. Diagram created in Microsoft Power Point by the author.

Among the aquatic reservoirs of *V. cholerae*, fish have however been less studied compared to phytoplankton and other aquatic organisms, especially in the AGLR where fish could be an underestimated reservoir for *V. cholerae* (Halpern and Izhaki, 2017; Hounmanou et al., 2016, 2019b; Senderovich et al., 2010). In Bangladesh, Hilsa fish (*Tenualosa ilisha*) have been identified as a potential source of toxigenic *V. cholerae* (Hossain et al., 2018), like in the case of *Rastrineobola argentea* (carps) in **Manuscript B**. In **Manuscript D**, we provide evidence that tilapia (*Oreochromis niloticus*), a popularly consumed fish in the AGLR, could serve as a reservoir for persistence and transmission of *V. cholerae* within the aquatic environment and subsequently increase the likelihood of human exposure (Hounmanou et al., 2019b). Similar studies using zebrafish models indicated that infected zebrafish can transmit *V. cholerae* to naïve zebrafish via excretion (Mitchell et al., 2017; Runft et al., 2014). The

implication of fish in cholera transmission in the region requires more research both on genomic characteristics of *V. cholerae* from fish and also on the host-pathogen interactions between *V. cholerae* and fish. Studies from Lake Tanganyika indicated a strong correlation between planktonic blooms, fish abundance and cholera (Echenberg, 2011; Plisnier et al., 2015). Similar implications of fish on the occurrence of cholera epidemics across Tanzania have also been reported (Dalusi et al., 2015; Rabia et al., 2017). Fish-eating birds such as great cormorants have also been identified as important drivers for dissemination of environmental *V. cholerae*, although the birds themselves may not be considered as reservoirs (Laviad -Shitrit et al., 2017). The *ctx*A gene encoding cholera toxin and *rfv*B-O1 have been identified in great Cormorants cloacal swabs, demonstrating that these birds that feed on fish can be important vehicles of global dissemination of toxigenic *V. cholerae* from one water-body to another and subsequently one place to the next (Laviad -Shitrit et al., 2017; Laviad-Shitrit et al., 2019). In experimental studies it has been shown that these migratory birds can shed *V. cholerae* that they get from infected fish for more than 72 h, a duration that is sufficient for them to cross oceans (Laviad -Shitrit et al., 2017; Laviad-Shitrit et al., 2019). There are different species of migratory water birds in the AGLR, which require further investigation regarding their role in the epidemiology of cholera in the region.

1.5. The transmission paradigm of *V. cholerae*

V. cholerae is mainly transmitted through the fecal-oral route. As an aquatic microorganism, the transmission of *V. cholerae* to humans has long been associated with exposure to contaminated water sources (Nelson et al., 2009; Ratchford and Wang, 2019). This implies that toxigenic clones of *V. cholerae* can infect an index case though ingestion of contaminated water or food (Fig. 7). The index case/cases can thereafter develop cholera and represent a source of further spread of the bacteria among the community, which will lead to an outbreak through human-to-human transmission (Lutz et al., 2013; Reidl and Klose, 2002). During outbreaks, infected individuals will re-infect water bodies with toxigenic

V. cholerae through discharge of sewage. The epidemic can be exacerbated in immunocompromised communities with poor hygienic conditions and existence of conflicts and population displacements (Hounmanou et al., 2016; Ingelbeen et al., 2019; Rajasingham et al., 2019). It is documented that in Sub-Saharan Africa, the highest number of cholera cases occur in populations around the great lakes, showing the importance of the local environment in the transmission cycle of cholera (Bi et al., 2018; Lessler et al., 2018). A recent study on the cholera dynamic in West Africa concluded that Accra, Ghana represented the main cholera hotspot in this region, with the findings incriminating mainly the poor water network systems and unprotected water sources (Moore et al., 2018). We describe the implication of the lakes and closeness to water bodies on the incidence of cholera in Tanzania in **Manuscript A.**

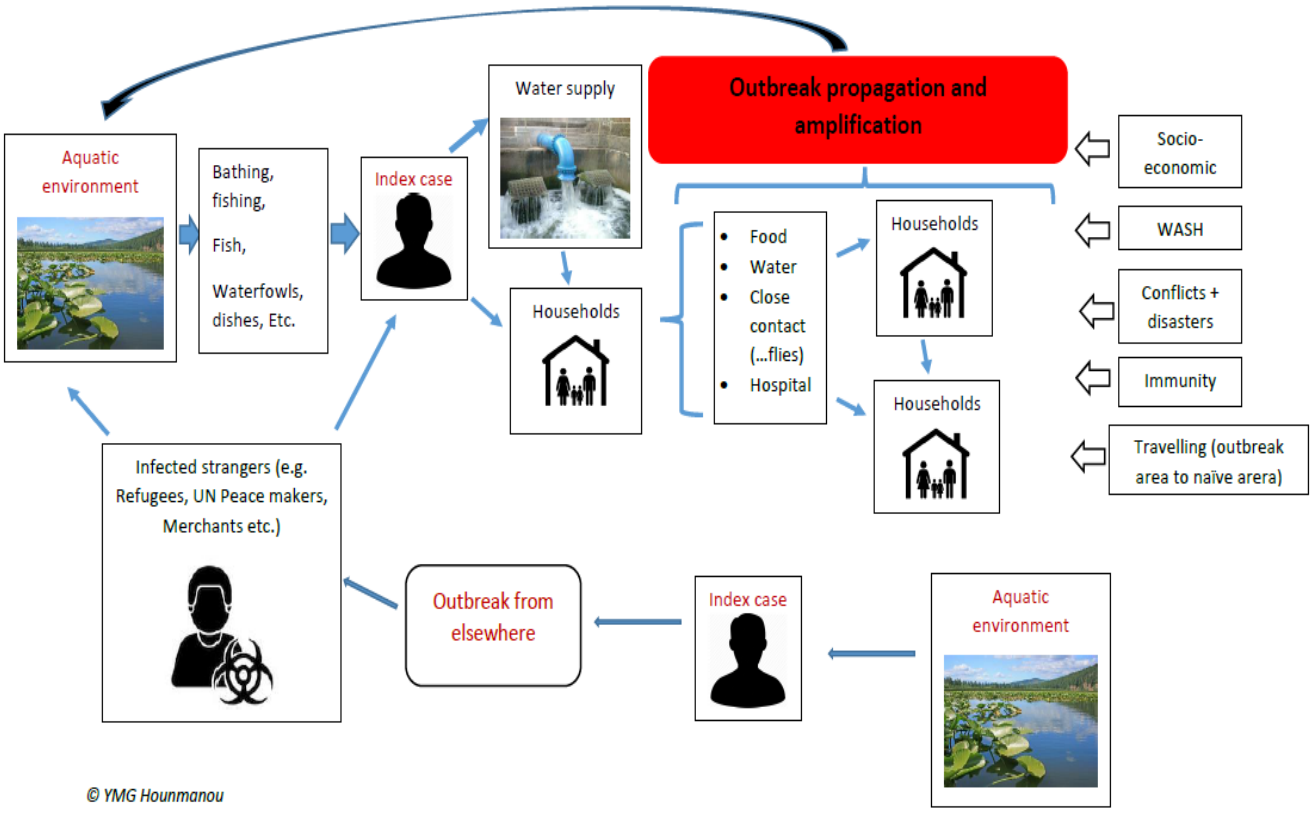


Figure 7. Schematic representation of the cholera transmission and propagation cycle. Diagram created in Microsoft Power Point by the author.

A cholera index case can also emanate from a different location, which is often the case, by acquiring *V. cholerae* O1 from an active outbreak in a different place then moving to a new geographical area. This was the case of the devastating Haitian outbreak in 2010, where *V. cholerae* O1 was introduced by UN peacemakers arriving from Nepal (Hendriksen et al., 2011; Katz et al., 2013). Similarly, *V. cholerae* O1 strains that caused the deadly 2016-2017 cholera outbreaks in Yemen were previously reported in East African outbreaks since 2014, (Weill et al., 2019), highlighting the complexity of the global dissemination of cholera. Infected individuals from imported outbreaks can then excrete *V. cholerae* O1 into the sewer system, thus maintaining an endemicity of new populations of *V. cholerae* in the aquatic environment. However, this theory of human-mediated transmission rather than environmentally driven cholera has become the topic of investigation in many recent studies as part of the ongoing debate on whether cholera epidemics are caused by local indigenous environmental toxigenic strains or are originating from imported cases (Ratchford and Wang, 2019; Weill et al., 2017). Genomic analyses from African and American cholera outbreaks suggest that human factors are more important in cholera dynamics than climate or environmental factors (Domman et al., 2017; Weill et al., 2017). Human activity like travel and poor sanitary conditions have also been reported as important sources for spread of pathogenic *V. cholerae* globally over the last century and particularly over the three decades during which cholera has established itself in Africa, Latin America, and the Caribbean (Robins and Mekalanos, 2014). Retrospective analyses of the ancient Danish cholera outbreaks in 1853 proposed the exclusion of the environmental sources in the cholera transmission paradigm because their data supported the human-mediated transmission called “short-cycle” (Phelps et al., 2017, 2018). Thus, it is suggested that humans are potential reservoirs of *V. cholerae,* at least for the African cases, even though this aquatic pathogen is not recognized as a fecal bacteria; e.g. to date there seems only to have been one report of a long term

carrier of *V. cholerae* in the world (Azurin et al., 1967). Moreover, *V. cholerae* cannot survive long-term in the acidic conditions in the human gut and humans can therefore not be considered as reservoirs for *V. cholerae*. Without further prospective microbiome studies revealing that humans can be reservoirs and long-term carriers of toxigenic *V. cholerae*, the transmission cycle of *V. cholerae* will still be built around the aquatic environments because the human-mediated transmission cannot stand alone without a prior introduction of the bacteria from the environment at one point, even for imported cases. Like the case of "eggs and chicken" rhetoric, the discussion about whether the environment or humans contaminate the other first remains a prospective area of study as the direction is still unknown. This issue has been addressed in **Manuscripts B** and **C** by characterizing clinical and environmental *V. cholerae* O1 strains from Tanzania using genomic tools. Furthermore, with accumulating evidence that waterbirds can carry and disseminate toxigenic *V. cholerae* from one water body to another between and within continents (Laviad-Shitrit et al., 2019), the aquatic environment will remain the focus of cholera transmission even in the occurrence of clonal strains in different locations.

1.6. Effect of climatic events on *V. cholerae* and cholera around the African Great Lakes

The seasonal pattern of cholera is one of the most incontestable facts on cholera, notably in cholera endemic countries (Lemaitre et al., 2019). Rainfall, for instance, plays a major role in water surface contamination, through the washout of open-air defecation sites and raw sewage circulation in the environment. In addition, rainfall will lead to increased washout/introduction of nutrients into coastal areas, with subsequent algae blooms due to eutrophication, with proliferation of *V. cholerae*. Outbreaks in Tanzania, Bangladesh and other endemic countries have usually been reported after flooding (Lemaitre et al., 2019; Reyburn et al., 2011; Rinaldo et al., 2012). A large ecological and microbiological study of Lake Tanganyika indicated seasonal fecal contamination of the Tanganyika surface water along all

sampling sites in Tanzania, Zambia, Burundi and DRC, where runoff water serves as vehicle for contamination of the lake with *V. cholerae* from sewage and open pit latrines (Plisnier et al., 2015).

El- Niño events, characterized by the rise in sea surface temperatures, can affect reservoirs of *V. cholerae* such as phytoplankton blooms in the lakes and thereby influence subsequent exposure and incidence of cholera. These parameters include increased water surface temperature, changed wind direction or upwelling of nutrients from deeper water (Reyburn et al., 2011), leading to higher incidence of cholera in the following years (Fig. 8). El- Niño events are reported to favor excess rainfall in the AGLR and have thus influenced cholera epidemics with low incidence during the dry season in some regions but a trend in increased outbreaks by the end of the dry season (Nkoko et al., 2011; Plisnier et al., 2015). Around Lake Victoria, significant fluctuations in cholera incidence were associated with increased rainfall (Stoltzfus et al., 2014). The correlation between rainfall and cholera incidence in mainland Tanzania from 2007 to 2017 is described in **Manuscript A** for all regions in mainland Tanzania. In Zanzibar, analysis of cholera and rainfall data between 1993 and 2017 revealed that outbreaks were highly seasonal, with high-risk periods corresponding to the annual rainy seasons. This could represent an opportunity for cholera preparedness and control (Bi et al., 2018).

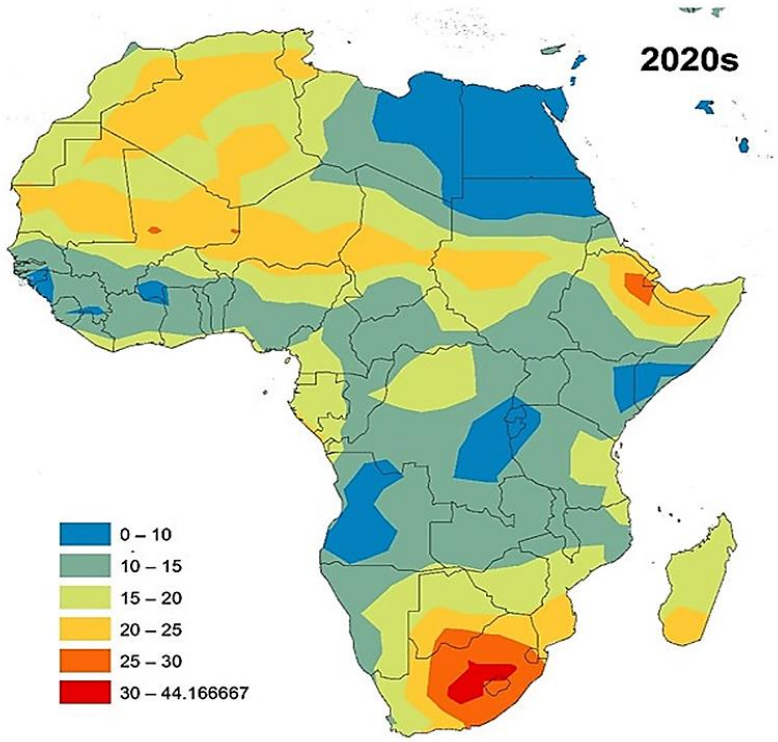


Figure 8. Map showing the strength (%) of temperature and precipitation that may amplify the incidence of cholera in Africa according to the climate projected for 2020. Warm colors represent new areas that may see increases in cholera outbreaks. Adopted from a previous report (Wendel, 2015).

Apart from rainfall, other climatic factors such as temperature and pH have a significant impact on the dynamics of *V. cholerae* in the environment and thereby on the epidemiology of cholera (Huq et al., 1984; Lipp et al., 2002). With the current climate change, global average temperature is expected to increase, which will result in the rise of sea levels. The increased global temperature may influence the temporal fluctuations of cholera, potentially increasing the frequency and duration of cholera outbreaks (Emch et al., 2008). Climatic factors like El- Niño events can alter the local aquatic environment in the lakes, such as water surface temperature, precipitation, salinity and nutrient concentration, which may favor *V. cholerae* either directly or through its environmental reservoirs (Emch et al., 2008). In the AGLR, encompassing both Lakes Tanganyika and Victoria and other nearby lakes with similar ecological

conditions, positive correlation between increase in cholera incidence and periods of warm El- Niño events have been reported (Nkoko et al., 2011; Plisnier et al., 2015). Focusing on five African Great Lakes, a positive correlation was also reported in Uganda between changes in climatic parameters and the recovery of toxigenic *V. cholerae* in the lakes (Bwire et al., 2018). Similar situations have been described in **Manuscript B**, where we described the occurrence of toxigenic *V. cholerae* O1 of outbreak potential obtained from Lake Victoria during a non-outbreak period under optimum water temperature, pH and salinity. The influence of increased temperature as a major risk factor affecting the survival and growth of *V. cholerae* in southeastern Africa has been suggested as a predicting variable for future cholera outbreaks in the region (Paz, 2009). On the global scale, cholera outbreaks seem recurrent and more constant in tropical countries close to the equator with high and generally constant temperatures, a region where Lake Tanganyika and Lake Victoria, the two main African Great Lakes are located (Emch et al., 2008; Ingelbeen et al., 2019). There is, however, an interconnection between temperature and other environmental factors like algal blooms, grazing of other organisms, or CTX phage acquisition for the growth of toxigenic *V. cholerae* that can lead to subsequent cholera (Faruque and Mekalanos, 2012; Vital et al., 2007). Moreover, wind direction through increased or decreased upwelling in temperature is reported to significantly influence the amount of nutrients in the water and therefore interact with the effect of temperature on *V. cholerae* and cholera outbreaks (Broza et al., 2005; Paz, 2009; Vital et al., 2007). The role and cost of climatic factors and climate change should not only be observed as risk factors; rather, they should be considered as assets for prediction and control of cholera in endemic settings like around the AGLR (Lipp et al., 2002; Trærup et al., 2011). Socioeconomic activities like fishing, leading to settlements along the lakeside, together with poor hygiene practices and unhygienic conditions, and social conflicts and unrest can interact with climatic factors to favor the emergence of

cholera outbreaks due either to indigenous strains of *V. cholerae* or newly introduced lineages (Lessler et al., 2018; Nkoko et al., 2011; Rebaudet et al., 2013).

1.7. Application of WGS in studying the evolution and resistance of *V. cholerae*

Before the advent of WGS, identification and characterization of *V. cholerae* was performed using a number of methods, ranging from phenotypic techniques to DNA based molecular typing. Most DNA based approaches included PFGE, *CTX*-genotyping, ribotyping and MLST (Rahaman et al., 2015). Various other methods have been developed and applied in characterization of *V. cholerae* after the establishment of WGS. These include VNTRs, variable number of tandem repeats; MLVA, multi-locus variable tandem repeat; and MVLST, multi-virulence locus sequencing typing (Table 2) (Rahaman et al., 2015). WGS however, offers so far the best resolution into the molecular epidemiology of pathogens (Bayliss et al., 2017). The first whole genome project for *V. cholerae* was published in the year 2000 when *V. cholerae* O1 El Tor N16961 was fully sequenced, launching the current genomic era of this pathogen (Heidelberg et al., 2000). Since then, WGS data coupled with bioinformatics tools have been adopted for evolutionary studies of *V. cholerae* and have been very effective in many regions of the globe. One example is the Haitian outbreak in 2010 that was tracked back to Nepalese peacemakers based on genomic analyses (Grad and Waldor, 2013; Hendriksen et al., 2011). Similarly, WGS data of 136 *V. cholerae* seventh pandemic El Tor isolates collected over 40 years were analyzed and three independent overlapping waves of *V. cholerae*, all coming from the Bay of Bengal, were identified with transcontinental outbreaks of cholera caused by genetically similar strains (Mutreja et al., 2011; Rahaman et al., 2015). Furthermore, the source and transmission routes of the devastating cholera outbreaks in Yemen in 2016-2017 were screened using genomic tools, revealing that the outbreak strains originated from south Asia, and have been involved in previous epidemics in East Africa before spreading to Yemen (Weill et al., 2019). In the African context, a large collection of 1,070 whole-genome sequences of *V.*

cholerae collected in 45 countries over almost 50 years were studied and revealed that all epidemics on the continent since the 1970s were caused by a single lineage of *V. cholerae* from Asia (7PET), which had been introduced on at least 11 occasions (Weill et al., 2017). The lineages involved in Tanzanian epidemics were described in **Manuscript C**. Genomic tools were also applied to decipher the evolution of cholera in the Americas, where the whole genomes of *V. cholerae* responsible for outbreaks in Central and South America were investigated and revealed that they originated from Asia (Domman et al., 2017). Although imported strains from Asia are found in most of these outbreaks to be the lead fuel of epidemics (Robins and Mekalanos, 2014), locally circulating strains both in people and the aquatic environment also cause many outbreaks (Domman et al., 2017; Hounmanou et al., 2016).

Table 2. Evolution of methods developed and applied in the characterization of *V. cholerae*. Data adopted from previous reports (Rahaman et al., 2015)

Periods	**Year of establishment**	**Methods**
Pre-sixth pandemic	1884	*V. cholerae* grown as pure culture
Sixth pandemic	1916	Serological classification
Pandemic free period	1923	Phage typing
	1959	Biotyping
Seventh pandemic	1986	Ribotyping
	1990	PFGE, RAPD
	1995	AFLP
	1998	MLST
	2000	WGS
	2002	VNTR
	2009	MLVA
	2011	MVLST

Complete genome reads from WGS data can be used to directly detect virulence genes, antimicrobial resistance determinants and the sequence types of *V. cholerae* strains in real time (Rahaman et al., 2015). Although antimicrobial resistance is not as alarming in cholera compared to other infectious diseases, the use of antimicrobials in the treatment regimen has contributed to the spread of resistance genes across the world following the spread of cholera. Genomic tools are employed to determine the occurrence of conjugative plasmids and integrative conjugative elements (ICE) in *V. cholerae* encoding resistance to many antibiotics, including sulfamethoxazole and trimethoprim, as well as various other acquired AMR genes and resistance due to chromosomal mutations (Hendriksen et al., 2011; Wang et al., 2016). In addition to evolutionary analyses, WGS data can be used for new diagnostics, molecular typing and vaccination strategies (Hendriksen et al., 2019; Ramamurthy et al., 2019). Furthermore, genomic analyses have provided understanding in the different molecular mechanisms of antibiotic resistance in *V. cholerae*, leading to suggestions of alternative strategies that can be used to treat the disease such as the recently proposed use of anti-virulence compounds (Narendrakumar et al., 2019).

Despite the advances provided by WGS in cholera epidemiology, there is still a need for standardization of pipelines and databases. For instance in the 2010 Haitian outbreak, Nepalese peacemakers were incriminated (Hendriksen et al., 2011), but the same strains were found as genetically related to Indian and Cameroonian strains (Reimer et al., 2011) and further concluded to have originated from South Asia, not Nepal (Chin et al., 2011). These confusions substantiate the limitations of genomic analyses, which always need epidemiological data for more comprehensive interpretations. In this Thesis, epidemiological data presented in **Manuscript A** have been used to complement the genomic information described in **Manuscripts B** and **C**.

1.8. Study objectives

1.8.1. Overall objective

The overall objective of this Thesis is to provide insights into the dynamics of cholera epidemics, the genomic evolution, the pathogenesis and aquatic reservoirs of the seventh pandemic *Vibrio cholerae* O1 in Tanzania. Its goal is to contribute to the pool of knowledge that will enable Tanzania to get rid of the cholera epidemics by 2030.

1.8.2. Specific objectives

To achieve the general objective, four studies were designed and conducted with the following specific objectives:

i. Study 1 (**Manuscript A**): To describe the epidemiology of cholera in Tanzania and address weaknesses in the current surveillance system in achieving the objectives of the global roadmap to 2030 for cholera control

ii. Study 2 (**Manuscript B**): To assess the occurrence of toxigenic *V. cholerae* O1 during non-cholera outbreak periods in Lake Victoria and the genetic characteristics that support environmental persistence and relatedness to pandemic strains

iii. Study 3 (**Manuscript C**): To investigate the evolution of *Vibrio cholerae* O1 isolated in Tanzania during the past three decades, including evolution in determinants of pathogenicity and antimicrobial resistance

iv. Study 4 (**Manuscript D**): To determine the role of fish (*Oreochromis niloticus*) as a reservoir host for the survival and transmission of toxigenic *V. cholerae* O1 in the aquatic environment

Chapter 2: Summary of Materials and Methods

This chapter describes the study designs and the study area where all four studies included in this Thesis were conducted and provides a brief overview of the analyses carried out in each study with further details provided in the **manuscripts A, B, C** and **D**.

2.1. Description of study area and study designs

All four studies of this Thesis were carried out using data and *V. cholerae* strains collected in the United Republic of Tanzania, an East African country, which is part of the AGLR (Nkoko et al., 2011). Tanzania shares borders with Kenya and Uganda to the north; Rwanda, Burundi, and the Democratic Republic of the Congo to the west; Zambia, Malawi, and Mozambique to the south; and the Indian Ocean to the East (Fig. 9). Tanzania is administratively divided into two parts, namely mainland Tanzania (also known as Tanganyika) and Zanzibar islands (NBS, 2012).

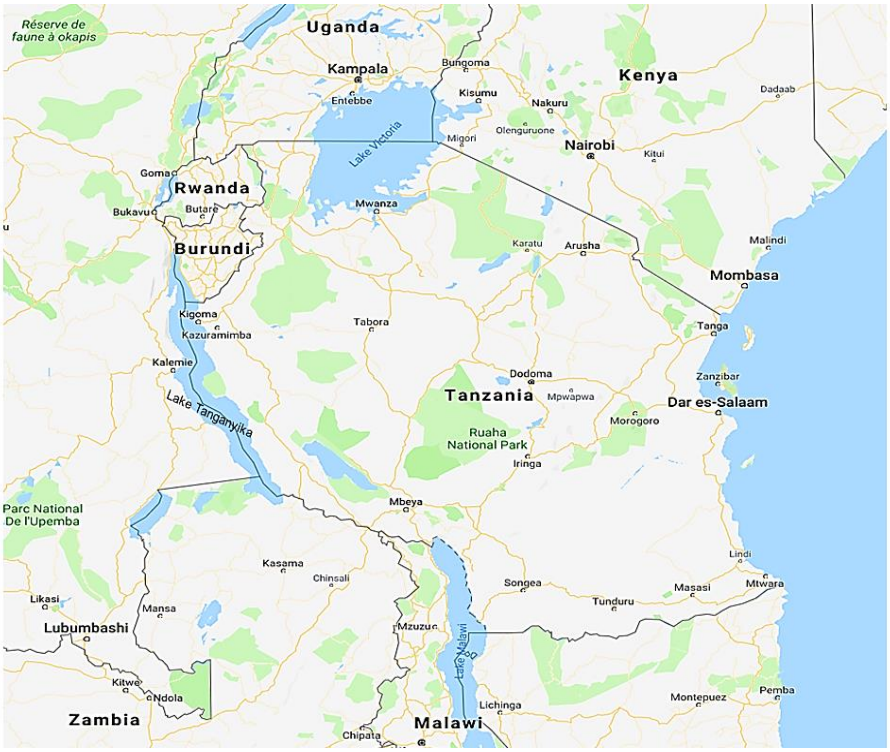


Figure 9. Localization of Tanzania, the study area within East Africa. Adopted from Google Maps.

In study 1, we performed a retrospective analysis on 11 years of cholera surveillance data collected from 2007 to 2017 by the Ministry of Health for all 25 regions in mainland Tanzania. Data from Zanzibar was not included because such analysis has already been conducted and published for the period 1993 to 2017 (Bi et al., 2018). This study was conducted using spatio-temporal and regression analyses to identify risk factors and hotspots in line with recommendations of the global task force on cholera control, which stipulates that prioritizing high-risk areas in endemic countries can substantially increase the efficiency of cholera control programs (**Manuscript A**). This is because only detailed analysis of local data in endemic countries can provide an in-depth understanding of local dynamics that can be used as data for action to support effective control measures.

Study 2 was a cross-sectional study conducted during the dry and rainy seasons along Lake Victoria where fish, phytoplankton and water samples were collected and analyzed for isolation of *V. cholerae* in off-shore and on-shore waters along the Mwanza basin of the lake as detailed in **Manuscript B**. In this study, we investigated the effect of seasonal pattern, the physico-chemical measurements of the Lake water and the absence of ongoing cholera outbreak on the occurrence of toxigenic *V. cholerae* O1 in Lake Victoria. The environmental dynamic of toxigenic strains during the non-epidemic periods is crucial in the epidemiology of cholera, but in such periods *V. cholerae* in the environment can be in a dormant, non-culturable state. We therefore employed culture-based methods and developed a PCR detection method from enriched samples aiming to increase the sensitivity of existing methods in the recovery of *V. cholerae* positive samples from the aquatic environment. This study furthermore employed genomic tools on the isolated *V. cholerae* strains to decipher their intrinsic genetic characteristics that support environmental persistence and their potential to cause epidemics through pathogenicity determinants and relatedness to pandemic strains.

In study 3, we analyzed clinical *V. cholerae* O1 samples collected during cholera epidemics across various regions in the country including Zanzibar, Tanga, Songwe, Singida, Ruvuma, Mwanza, Morogoro, Mbeya, Mara, Kigoma and Dar es Salaam (**Manuscript C**). Here, we aimed to provide a holistic picture of the evolutionary genomics of *V. cholerae* O1 involved in cholera outbreaks in Tanzania during the past three decades, including evolution in determinants of pathogenicity and antimicrobial resistance. The Tanzanian outbreaks strains were then fully sequenced and genome-wide analyses were performed to compare them to environmental isolates obtained in Lake Victoria in order to provide evidence for environmental survival of pandemic clones, which could be progenitors of future epidemics. Both clinical and environmental strains were thereafter analyzed in a global context of the seventh pandemic cholera, which helped to characterize introductions and spread of various sub-lineages into and outside the AGLR.

Study 4 was an experimental animal study that took place in laboratory conditions, where fish were kept in aquaria as described in **Manuscript D**. The genomic findings in study 2 and 3, along with the literature, provided accumulating evidence of persistence of pandemic clones of *V. cholerae* in the aquatic environment. Since fish are one of the suspected aquatic reservoirs of *V. cholerae* that are directly part of human food chain but could also play a role in global dissemination of the pathogen during migration, the study 4 presented in **Manuscript D** was conducted. In this study, we assessed the role of a popular African fish species, tilapia (*Oreochromis niloticus*) as a reservoir host for the survival and transmission of toxigenic *V. cholerae* in the aquatic environment. The fish were infected in aquaria using 5×10^7 cfu/mL of *V. cholerae* O1 Classical, El Tor wild type, El Tor Δ*toxT* and *V. cholerae* non-O1 to study their colonization, duration of shedding and transmission in tilapia guts.

2.2. Overview of collected data and analyses perfomed

Manuscripts A to **D** provide detailed descriptions of the data collection procedures and methodologies used in the studies. Table 3 displays a schematic overview of the methods developed and applied.

Table 3. Summary of collected data and subsequent analyses.

St**udies**	**Data collected**	**Data processing and analyses**
Study 1 (**Manuscript A**)	Cholera cases and deaths (2007-2017)	Descriptive statistics
	Total water area in a region (km^2)	Calculation of cholera incidence
	Total water perimeter/100 km in a region	Calculation of case fatality rates
	Demographic and socio-economic data	Hotspots analysis using SatScan
	Proportion of regions having sea border (%)	Poisson regression for risk factors analysis
	Proportion of regions with international border (%)	Estimation of incidence rate ratios (IRR)
	Rainfall data during year of epidemic	
	Lag rainfall in year before epidemic	
	GIS data and country shape files	
Study 2 (**Manuscript B**)	Water samples from Lake Victoria	Enrichment of samples in alkaline peptone water
	Phytoplankton samples	Bacteriological culture for identification of *V. cholerae*
	Fish samples (*Rastrineobola argentea*)	DNA extraction from both cultured isolates and enriched samples
	pH, temperature, salinity, dissolved oxygen measurements in Lake Victoria	PCR on DNA from cultured isolates and enriched samples
		Chi-square and regression analysis
		Antimicrobial susceptibility testing
		Whole genome sequencing (WGS) and bioinformatics analyses
Study 3 (**Manuscript C**)	Clinical *V. cholerae* O1 (2015-2017)	Culture, confirmatory PCR and antimicrobial susceptibility testing
	Public *V. cholerae* genomes originating from Tanzania in Genbank and ENA (1993-2015)	DNA extraction from cultured isolates
	Environmental *V. cholerae* genomes from Study 2	WGS and bioinformatics analysis
Study 4 (**Manuscript D**)	Farmed Tilapia (*Oreocrhomis niloticus*)	Aquaria experiment with exposure of fish to *V. cholerae* strains
	V. cholerae strains O395, E7946, JW612 and *V. cholerae* non-O1	Colonization experiments for 14 days with repetition for each strain
	Intestinal *V. cholerae* counts and counts from aquaria water	Cohabitation experiments with repetitions
	Optical density measurements of aquarium water	Generalized linear regression analysis for counts

2.3. Ethical Consideration

Studies reported in **Manuscripts A, B** and **C** contained no human data or animal manipulation protocols and did not require any ethical approval. However, **Manuscript D** with the aquarium experiments and infection of fish with *V. cholerae* required an ethical approval, which was obtained from the ethical review board through the ethical clearance certificate for conducting animal-related research in Tanzania with issue number SUA/CVMBS/018/07. Euthanasia of fish and disposal of waste materials were performed according to instructions in the ethical approval.

Chapter 3. Summary of results

In this chapter, we provide a concise description of results obtained in the four manuscripts.

The epidemiological study carried out in **Manuscript A** aimed to provide understanding in the mechanisms of emergence and transmission of cholera in Tanzania as stipulated by the global roadmap for cholera control by 2030. Descriptive statistics in this short-communication article revealed a cumulated 39,444 cholera cases and 600 deaths across all regions of mainland Tanzania between 2007 and 2017, giving a case fatality rate of 1.5% and an average annual incidence rate of 8.39 per 100,000 people. It is shown in **Manuscript A** that cholera high-risk populations in Tanzania were mostly those living in central regions due to urbanization and also populations living near the Great Lakes such as Lake Victoria, Tanganyika and Nyaza. The risk of experiencing cholera in these regions was up to 2.9 times higher than elsewhere in the country (**Manuscript A**). Besides identifying the lake zones as cholera hotspots in Tanzania, this study also confirmed that living near water bodies increases the risks for cholera, i.e. for every 100 Km of water perimeter in a region, the cholera incidence increased by 1.5% (**Manuscript A**).

In study 2, we reported the occurrence of pandemic sub-lineages of *V. cholerae* O1 in the aquatic environment including in fish, phytoplankton and water in Lake Victoria during non-epidemic periods (**Manuscript B**), confirming the implication of waterbodies on cholera incidence as identified in **study 1**. Genome analyses of the sequences of these environmental isolates recovered in the Lake more than a year after cholera outbreaks have ceased in the lake zone revealed that they were pathogenic strains belonging to the seventh pandemic lineage. The majority of them, i.e. F2, F4, W1, W3, P2 and P3 (**Manuscript B**) belonged to the third wave of the seventh cholera pandemic and are phylogenetically closely related to strains that caused cholera epidemics in Tanzania, Kenya, and Uganda in 2015, with as low as three SNPs difference. Some of these environmental strains, notably F1, F3 and W2 of the T10

sub-lineage (described in **Manuscript C**), have persisted in the aquatic environment for relatively longer periods due to their relatedness to strains from outbreaks that have occurred going back to 1998 (**Manuscript B** and **C**). The environmental persistence of these strains recovered from the Lake was attributed to the circulation of clonal strains in rampant outbreaks in the Eastern African region but also to their intrinsic genetic features enhancing survival in the aquatic environment and their interaction with fish and phytoplankton supported by the optimum physico-chemical parameters of the Lake (**Manuscript B**). Amongst others, the genetic elements found in the genomes of *V. cholerae* O1 isolated from Lake Victoria that could support their environmental fitness include the putative CRISPR/Cas system for phage tolerance, autoinducers (AI-2 *LuxP* and *LuxQ*) involved in quorum sensing and biofilm formation for environmental survival. Various environmental stress response regulator proteins were conserved in the strains, mainly the response regulators of the *Vie*SAB transduction system of *Vibrio*, the two-component response regulator proteins, histidine kinase, and Vibrio Polysaccharides (VPS) biosynthesis proteins (**Manuscript B**). In study 2, we also designed and applied a method with high sensitivity for the recovery of toxigenic *V. cholerae* from environmental samples, which consisted of a multiplex PCR reaction (targeting *ompW* and *ctxA*) on DNA extracted directly from samples enriched for 6 h in alkaline peptone water. This method as shown in **Manuscript B** is more suitable for surveillance of *V. cholerae* in the aquatic environment.

In **Manuscript C**, we studied clinical *V. cholerae* O1 isolated from cholera patients in Tanzania during the past three decades between 1993 and 2017 in a national and global context along with strains recovered from Lake Victoria in study 2, aiming to investigate their genomic evolution and propose guidance for control based on the genetic patterns of past epidemics. This study revealed that all cholera outbreaks that occurred in Tanzania were caused by the 7PET lineage predominated by three time-separated sub-lineages, T5, T10 and T13. The T5 sub-lineage of genotype *ctx*B3 of *V. cholerae* O1

El Tor, known to be part of Wave I of the seventh cholera pandemic, caused all epidemics in Tanzania until 1997 (**Manuscript C**). Between 1998 and 2012, all cholera epidemics occurring in Tanzania were caused by the T10 sub-lineage of atypical El Tor genotype *ctx*B1. The environmental strains F1, F3 and W2 misclassified in study 2 were here confirmed as T10 (**Manuscript C**). Interestingly, *V. cholerae* O1 causing outbreaks in the Tanzanian refugee camps between January and May 2015 were also identified as genotype *ctx*B1 and T10 sub-lineage known as early Wave III strains of the 7PET lineage. The origin of these refugee camp strains was however traced back to DRC and Zambia based on their sequence type ST515, where identical clones have caused outbreaks in 2012 and 2013. It is worth noting that the locally circulating T10 strains in Tanzania from 1998 to 2012 and those found in Lake Victoria were all ST69, highlighting a potential environmental persistence of the outbreak strains (**Manuscript C**). Except for the refugee camp outbreaks, all analyzed strains in Tanzania after 2013 were T13 atypical El Tor of the current wave III genotype *ctx*B7 causing most cholera outbreaks until 2017. These T13 strains were phylogenetically related to strains from other East African countries and Yemen occurring in the same periods suggesting cross-border transmission of the same clones (**Manuscript C**). There was also a significant involvement of the African Great Lakes substantiated by clonality at core and accessory genomes level between environmental strains from Lake Victoria and *V. cholerae* O1 responsible for older, as well as recent epidemics in Tanzania (**Manuscript C**). This study also revealed that the T13 strains are less drug resistant and present approximately 10-kb nucleotide deletions in the SXT element. Moreover, nucleotide deletions were observed in the CTX prophage of some strains within T13 and T10 sub-lineages, which suggest recombination but also requires further virulence studies for their clinical relevance (**Manuscript C**).

The environmental reservoirs of toxigenic *V. cholerae* are of crucial epidemiological importance for cholera in the sense that knowledge about the inter-epidemic reservoirs of *V. cholerae*

may provide useful insights that can guide control of cholera. Besides cyanobacteria and other phyto- and zooplanktons that are well-studied and documented aquatic reservoirs for *V. cholerae*, Study 4 identified a common fish species, tilapia (*Oreochromis niloticus*), as a potential reservoir host for survival and transmission of toxigenic *V. cholerae* in the aquatic environment (**Manuscript D**). The study revealed that the seventh pandemic El Tor *V. cholerae* O1 colonized tilapia intestines and persisted at stable concentrations around 10^3 cfu/intestine for over two weeks. When water was renewed in the aquaria every day by fresh sterile water (see **Manuscript D**), *V. cholerae* counts in water decreased from 10^7 to 10^3 cfu/ml and intestinal counts went from 10^6 to 10^2 cfu/intestine with and without feeding. The tested *V. cholerae* strains were transmitted from infected to naïve tilapia after 24 h of cohabitation. This experimental study confirms for the first time that tilapia and possibly other fish species provide means for colonization and multiplication for the pandemic strains of *V. cholerae* in the aquatic environment while also favoring its horizontal transmission, thus supporting persistence within the fish population and the aquatic environment. It will therefore increase the likelihood of human exposure to the pathogen via direct contact with fish, fish infected water or further spread from fish-eating migratory birds.

Manuscript A: "Cholera hotspots and surveillance constraints contributing to recurrent epidemics in Tanzania"

Short-communication article published in ***BMC Research Notes*** *(2019) 12:664*

https://doi.org/10.1186/s13104-019-4731-0

Yaovi M.G. Hounmanou[1*], Kåre Mølbak[1,2], Jonas Kähler[2], Robinson H. Mdegela[3], John E. Olsen[1], Anders Dalsgaard[1,4]

*Corresponding address: gil@sund.ku.dk

Keywords: Cholera; Tanzania; spatial-temporal analysis; Great Lakes; cholera dynamics

Hounmanou et al. BMC Res Notes (2019) 12:664
https://doi.org/10.1186/s13104-019-4731-0

BMC Research Notes

RESEARCH NOTE

Open Access

Cholera hotspots and surveillance constraints contributing to recurrent epidemics in Tanzania

Yaovi M. G. Hounmanou[1*], Kåre Mølbak[1,2], Jonas Kähler[2], Robinson H. Mdegela[3], John E. Olsen[1] and Anders Dalsgaard[1,4]

Abstract

Objective: We described the dynamics of cholera in Tanzania between 2007 and 2017 and assessed the weaknesses of the current surveillance system in providing necessary data in achieving the global roadmap to 2030 for cholera control.

Results: The Poisson-based spatial scan identified cholera hotspots in mainland Tanzania. A zero-inflated Poisson regression investigated the relationship between the incidence of cholera and available demographic, socio-economic and climatic exposure variables. Four cholera hotspots were detected covering 17 regions, home to 28 million people, including the central regions and those surrounding the Lakes Victoria, Tanganyika and Nyaza. The risk of experiencing cholera in these regions was up to 2.9 times higher than elsewhere in the country. Regression analyses revealed that every 100 km of water perimeter in a region increased the cholera incidence by 1.5%. Due to the compilation of surveillance data at regional level rather than at district, we were unable to reliably identify any other significant risk factors and specific hotspots. Cholera high-risk populations in Tanzania include those living near lakes and central regions. Successful surveillance require disaggregated data available weekly and at district levels in order to serve as data for action to support the roadmap for cholera control.

Keywords: Cholera, Tanzania, Spatial–temporal analysis, Great Lakes, Cholera dynamics

Introduction

Half of all cholera reported cases from Africa between 1970 and 2011 were notified by seven countries, including Tanzania, which has remained one of the top cholera reporting countries until 2018 [1, 2]. Since the seventh cholera pandemic reached the country in 1974, Tanzania reports outbreaks almost every year and has notified over 250,000 cases and 13,078 deaths by 2018 [1, 3].

In 2017, the global task force on cholera control established a roadmap to 2030 for elimination of cholera with Tanzania being one of the 48 targeted endemic countries [4]. The strategies recommended by the task force include rigorous surveillance for early detection and response with emphasis on cholera high-risk populations at local levels for optimal interventions [4, 5]. Therefore, prioritizing high-risk areas in endemic countries can increase the efficiency of cholera control programs because only detailed analysis of local data in each country can provide better understanding of local cholera dynamics for effective control [1]. Successful identification of high-risk areas, however depends on robust surveillance, which is difficult to achieve in many countries including Tanzania due to the stigma related to cholera reporting and associated economic losses in the tourism sector leading to inaccurate reporting in many countries [5]. Nevertheless, a number of spatio-temporal studies have been conducted with surveillance data in Uganda, in the Democratic Republic of Congo (DRC), in India and on Zanzibar islands to contribute with scientific knowledge

*Correspondence: gil@sund.ku.dk
[1] Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, 1870 Frederiksberg C, Copenhagen, Denmark
Full list of author information is available at the end of the article

enabling these countries to get rid of epidemic cholera by 2030 [6–10].

In the present study, we analyzed 11 years of available cholera surveillance data between 2007 and 2017 from all regions of mainland Tanzania. Together with geographical, climatic and socio-demographic data, hotspots identification and risk factors analyses were performed to describe the epidemiology of cholera in Tanzania and address weaknesses in the current surveillance system in achieving the objectives of the global roadmap to 2030 for elimination of cholera in the country.

Main text

Methods

Data collection

Retrospective data of cholera cases and deaths compiled at regional levels from 2007 to 2017 were obtained from the Epidemiology and Disease Control Section of the Tanzanian Ministry of Health. Data included in this study are those reported based on the WHO standards for cholera case definition and are described as follows:

(i) Patient aged 5 years or more with severe dehydration and acute watery diarrhea or individual who died from the same symptoms in an area without a confirmed outbreak;
(ii) A patient aged 2 years or above having acute watery diarrhea, with or without vomiting in an area where there is an ongoing cholera epidemic;
(iii) Confirmation of *Vibrio cholerae* O1 isolated in the stool of suspected patients.

Population data for the country's 25 regions (Additional file 1) were extracted from the latest 2012 Population and Housing Census report from the National Bureau of Statistics [11]. Using the inter-censual growth rate of each region, annual population in each region was calculated to estimate the population at risk per year in each region. Proportion of households with access to improved drinking water and households with access to improved toilets were retrieved and included in the analysis. For indicators of socioeconomic status, proportion of households possessing mobile phone and television during the census was used (Additional file 1).

Rainfall data were obtained from the Tanzanian Meteorological Agency, where annual rainfall data in millimeter was obtained per region for each of the 11 years.

Country shape files were obtained from the National Bureau of Statistics [11]. GIS data of the water areas were obtained online from diva-GIS website [12]. All shape files were analyzed in quantum GIS version 2.18, Las Palmaras (https://qgis.org/en/site/) for validity. In QGIS, regional and waterbodies polygons were joined using the Union vector to determine two new variables: total water area and water perimeter found in each region and included in the analysis (Additional file 1).

Data analyses

Identification of high risk clusters (hotspots) were performed using SaTScan v 9.6 (https://www.satscan.org/) with the Poisson-based spatial scan. The centroid coordinates of each region was detected from the regional shape-file using QGIS 2.18. These coordinates were used as the geographic references of the regions. In the Poisson model, the expected cholera cases in each part of the regions are assumed proportional to the population size of the region. The model detected clusters in a multidimensional point process and allowed variable window sizes to scan for cholera cases within the region. Variable window size was used, because a prior estimation of the size of the area covered by a cluster was not known. A circular scan window was selected, which moved over the entire region with a radius that varied from zero to 25% of the population at risk. The clusters covered areas with lower rates outside a circular scan window compared with higher rates inside the circle. The likelihood ratio for a specific window was determined as previously described [6, 8]. The output files were displayed in Google Map.

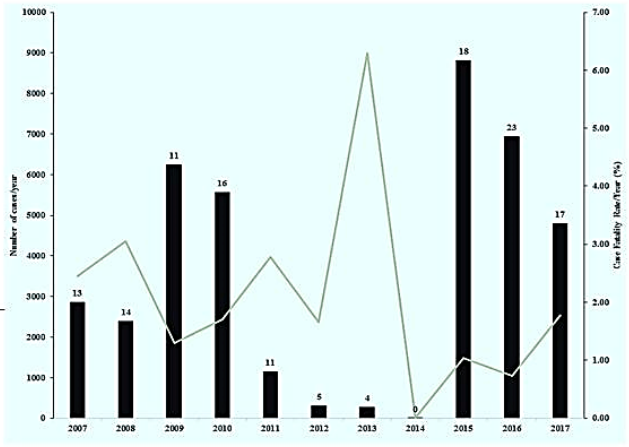
Poisson regression was used for the analyses of potential risk factors for cholera. In this model, the total number of cases (2007 to 2017) reported at the regional level was the dependent variable, exposure variables were obtained from the 2012 census as well as parameters from the geographical analysis (Fig. 1a). The logarithm to the population was included as an offset variable and thereby the analysis represents a log-linear model of the incidence. We used a zero-inflated Poisson regression for the analysis of the relation between rainfall in the year before reporting and the number of reported cases. In both models, we applied robust standard errors. In the analysis of the relation between rainfall and number of reported cases, we adjusted the standard error for the 25 regions because it was assumed that there is less variance within a region than between regions. The log to the estimated yearly population was included as offset variable. Stata version 14 was used for the regression models.

Results and discussion

From 2007 to 2017, mainland Tanzania reported 39,444 cholera cases with 600 deaths, giving a case fatality rate of 1.5% and an average annual incidence rate of 8.39 per 100,000 people. A similar analysis of 10 years data from the DRC revealed a higher number of cases and case fatality rates of 1.9% [9]. DRC is the country reporting the most cholera cases in Africa and for many years have been a devastated country because of wars and

Variables	Incidence Rate Ratio (95% Confidence Interval)	Significance (p)
Proportion of urban population (%)	1.24 (0.47-3.30)	0.663
Water Area in a region (Km^2)	1.000 (0.99-1.000)	0.927
Water Perimeter/100km	1.015 (1.000-1.030)	0.042
Proportion of households with TV (%)	0.99 (0.98-1.009)	0.878
Proportion of households with Mobile Phone (%)	0.98 (0.96-1.011)	0.289
Literacy rate (%)	1.008 (0.946-1.076)	0.788
Population Density (%)	0.999 (0.999-1.000)	0.721
Proportion of households with improved sanitation (%)	1.003 (0.988-1.018)	0.668
Proportion of households with improved water (%)	1.013 (0.989-1.038)	0.265
Growth rate (%)	0.913 (0.754-1.106)	0.356
Proportion of regions having sea boarder (%)	1.167 (0.649-2.096)	0.605
Proportion of regions having international boarder (%)	0.924 (0.464-1.836)	0.82
Rainfall during year of epidemic (dm)	0.897 (0.797-1.009)	0.071
Lag rainfall in year before epidemic (dm)	0.867 (0.718-1.049)	0.144

a



b

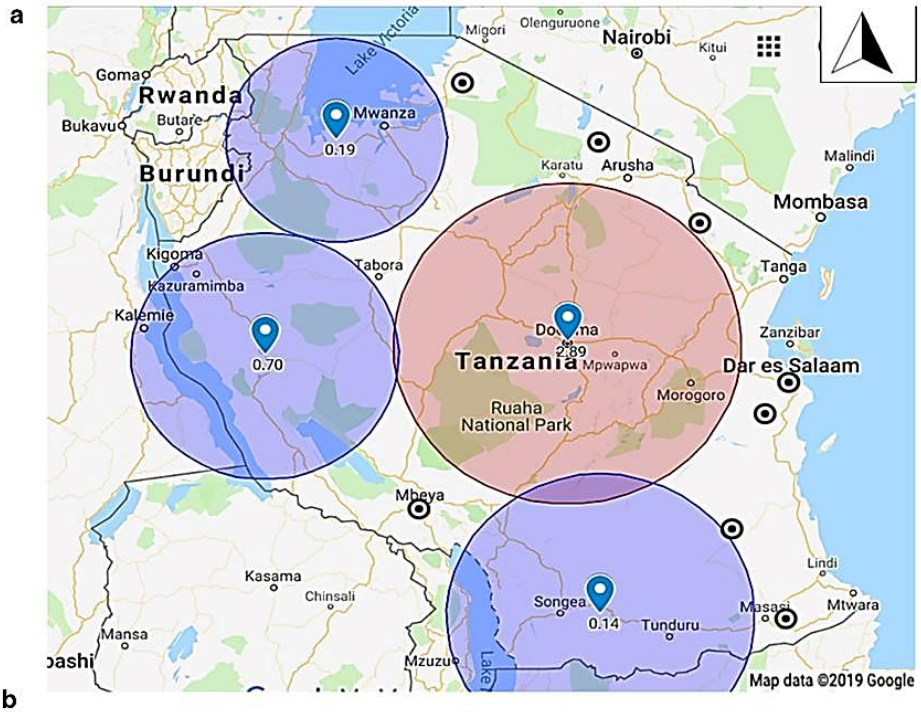
Fig. 1 **a** Poisson regression of determinants for cholera reporting in Tanzania based on regional data. **b** Number of cholera cases (bars) and case fatality rates (line) per year in mainland Tanzania. Numbers of affected regions are shown on top of each bar

population displacement associated with higher risks for cholera [9]. However, the relatively lower incidence in Tanzania could be attributed to the fact that reporting cholera has negative impacts on tourism and exports of affected countries, leading to underreporting [5]. In 2016, 23 of the 25 regions in Tanzania reported cholera but most cases (8821) were reported in 2015 (Fig. 1b). The highest case fatality rate was recorded in 2013 (6.3%) when only 270 cases were reported countrywide. It is likely that initial cases in an outbreak experience elevated case-fatality, whereas the official recognition of an outbreak leads to improved management and increased case-finding thereby identifying milder cases as well. At regional level, Shinyanga had the highest case fatality rate in the study period (7%) while Dodoma reported the highest number of cases (5988), although 94.1% of these cases were recorded exclusively between 2015 and 2017 where the countrywide incidence reached 14.2 per 100,000 people (Fig. 1b). Every region except Kagera, reported cholera at least once during the 11 years. Fifteen regions reported cholera in at least six of the 11 years and can be considered cholera endemic regions [6, 8].

Spatial analyses revealed four high-risk areas of different sizes (Fig. 2a). Seventeen regions had their centroids within the identified hotspots. The risk of having cholera in these regions was up to 2.89 times higher compared to elsewhere in the country ($p < 0.0001$, Fig. 2b). The hotspot with highest risk of cholera was Dodoma region, including its neighboring regions where high magnitude outbreaks were recorded after 2015. This corroborates the role of urbanization and population displacements in cholera dynamics [13–15]. The increasing number of outbreaks in 2015–2017 in Dodoma coincides with the time where the Tanzanian government moved offices to Dodoma as the capital city of the country. This was associated with significant movement of government employees and affiliated business from Dar es Salaam, but also a number of people working with the construction of new government office buildings [16].

Approximately 28 million people live in the regions found in the hotspots based on the 2012 census. Three of the four hotspots were around major lakes in the country mainly Lake Victoria, Tanganyika and Nyasa. Living near a lake was also reported in Uganda, the DRC and elsewhere as a factor associated with increased cholera incidence [8, 17, 18].

According to existing literature, Dar es Salaam is one of the endemic cities experiencing cholera outbreaks in Tanzania [19]. This city was however not identified in our analysis as part of the hotspot areas. Such a discrepancy reveal one of the limitations of the data and weaknesses of the existing surveillance system in which cholera reports were aggregated at regional level rather than at district level. The observed clusters should therefore be treated with caution. Reliable disease hotspot identification cannot be effective when surveillance and risk factor



b

Relative Risk of clusters	Regions covered	Population at risk
2.89	Dodoma, Iringa, Singida, Manyara, Morogoro, Tanga	10105096
0.70	Katavi, Rukwa, Kigoma, Tabora	6006575
0.19	Geita, Mwanza, Kagera, Shinyanga, Simiyu	10116439
0.14	Ruvuma, Njombe	2081435

Fig. 2 Cholera hotspots in mainland Tanzanian and relative risks as identified by SatScan and displayed using Google map. **a** Mapping of the relative risks of cholera within the identified hotspots. **b** Regions at risk and population size in the identified cholera hotspots

data are not available from the smallest geographical structures in affected countries [6, 8]. Reliability of data also includes consistency in case definitions because only a small proportion of suspected cases are laboratory confirmed and may bias identification of priority areas where interventions are needed [5].

Compared to previous studies in the DRC, Uganda, India and Zanzibar [6–9] which had cholera data and exposure variables from districts, the available regional data in Tanzania were not able to detect many risk factors investigated against the incidence of cholera. The rainfall pattern, both before and during the year of epidemics did not significantly affect the cholera incidence in mainland Tanzania based on currently available data ($p=0.07$; 0.14, respectively, Additional file 1). This could mean that in reality there is no positive correlation between rainfall and risk for cholera in mainland Tanzania. Nevertheless, cholera is normally expected to have a seasonal pattern as is the case in Zanzibar and elsewhere [7, 20, 21]. Only water perimeter in a region was significantly associated with incidence of cholera with a 1.5% increase in incidence for 100 km increase in water perimeter (IRR 1.015; 95% CI 1.001 to 1.030; $p=0.042$, Fig. 3). This correlates findings of the hotspots analyses where three significant clusters covered regions around the Great Lakes. Living near water bodies mainly great Lakes seems therefore a significant risk factor for cholera which has also been documented globally [22]. Recent findings in Tanzania where *V. cholerae* O1 isolated from Lake Victoria were phylogenetically identical to those causing cholera outbreaks in the African Great Lakes region further illustrates the association between Lakes and cholera [23].

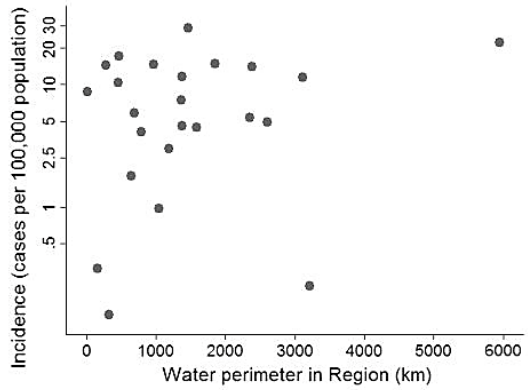


Fig. 3 Relationship between cholera incidence (log incidence per 100,000 population) and total water perimeter (in kilometer) in a region. One region has been omitted, i.e. Kagera because of zero reporting

Limitations

The principal limitation of this study is the weakness of the cholera surveillance system in Tanzania to provide disaggregated data available weekly and at district levels in order to serve as data for action to support the roadmap for cholera control. Moreover, there is a low incidence of cholera in Tanzania based on reported data and this could be attributed to the fact that reporting cholera has negative impacts on tourism and exports of affected countries, leading to underreporting. Furthermore, Dar es Salaam is one on of the main cities experiencing recurrent cholera outbreaks in Tanzania but was not identified as part of the hotspot areas because data were aggregated at regional level and could bias the spatial analysis. Data aggregated at regional levels were not able to detect rainfall as a significant risk factor on cholera incidence. During data curation we observed that only a small proportion of suspected cases are laboratory confirmed and this may affect the consistency in case definitions and identification of priority areas where interventions are needed.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s13104-019-4731-0.

Additional file 1. Number of cholera reported cases and deaths by region and cholera risk factors.

Abbreviation

DRC: Democratic Republic of Congo.

Acknowledgements

We are grateful to Dr. George C.Kauki and Janneth Mghamba from the Epidemiology and Disease Control Section of the Ministry of Health Community Development, Gender Elderly and Children of Tanzania for providing the data.

Authors' contributions

YH collected the data, carried out initial descriptive statistics and GIS data analyses, constructed graphs, made data interpretation and drafted the original Manuscript. KM carried out the regression analyses, result interpretation, and revised the final manuscript. JK provided guidance in GIS data analysis and revised the final manuscript. RM, JO and AD provided guidance in data interpretation and critical revision of the final manuscript. KM and AD supervised this study and all authors have approved the final version of the manuscript for submission.

Funding

Authors received funding from the Danish International Development Assistance (Danida) through the project "Innovations and Markets for Lake Victoria Fisheries" (IMLAF); Grant Number: DFC File No. 14-P01-TAN and the International Foundation for Sciences; Grant Number: 1-2-A-6100-1.
The funding bodies had no roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The dataset supporting this study has been submitted in Additional file 1.

Ethics approval and consent to participate

Not applicable. The study involved retrospective surveillance data.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1] Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, 1870 Frederiksberg C, Copenhagen, Denmark. [2] Division of Infectious Disease Preparedness, Statens Serum Institut, Artillerivej 5, 2300 Copenhagen, Denmark. [3] Department of Veterinary Medicine and Public Health, College of Veterinary and Biomedical Sciences, Sokoine University of Agriculture, PO Box: 3021, Morogoro, Tanzania. [4] School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore, Singapore.

Received: 3 September 2019 Accepted: 14 October 2019
Published online: 21 October 2019

References

1. Lessler J, Moore SM, Luquero FJ, McKay HS, Grais R, Henkens M, et al. Mapping the burden of cholera in sub-Saharan Africa and implications for control: an analysis of data across geographical scales. Lancet. 2018;391:1908–15. https://doi.org/10.1016/S0140-6736(17)33050-7.
2. Mengel MA, Delrieu I, Heyerdahl L, Gessner BD. Cholera outbreaks in Africa. In: Nair GB, Takeda Y, editors. Cholera outbreaks. Berlin: Springer; 2014. p. 117–44. https://doi.org/10.1007/82_2014_369.
3. WHO. Cholera—United Republic of Tanzania. WHO. 2018. http://www.who.int/csr/don/12-january-2018-cholera-tanzania/en/. Accessed 15 Oct 2018.
4. WHO. Ending cholera. A global roadmap to 2030. 2017. http://www.who.int/cholera/publications/global-roadmap/en/. Accessed 15 Oct 2018.
5. Azman AS, Moore SM, Lessler J. Surveillance and the global fight against cholera: setting priorities and tracking progress. Vaccine. 2019. https://doi.org/10.1016/j.vaccine.2019.06.037.
6. Ali M, Sen Gupta S, Arora N, Khasnobis P, Venkatesh S, Sur D, et al. Identification of burden hotspots and risk factors for cholera in India: an observational study. PLoS ONE. 2017;12:e0183100. https://doi.org/10.1371/journal.pone.0183100.

7. Bi Q, Abdalla FM, Masauni S, Reyburn R, Msambazi M, Deglise C, et al. The epidemiology of cholera in Zanzibar: implications for the zanzibar comprehensive cholera elimination plan. J Infect Dis. 2018. https://doi.org/10.1093/infdis/jiy500.
8. Bwire G, Ali M, Sack DA, Nakinsige A, Naigaga M, Debes AK, et al. Identifying cholera "hotspots" in Uganda: an analysis of cholera surveillance data from 2011 to 2016. PLOS Negl Trop Dis. 2017;11:e0006118. https://doi.org/10.1371/journal.pntd.0006118.
9. Ingelbeen B, Hendrickx D, Miwanda B, van der Sande MAB, Mossoko M, Vochten H, et al. Recurrent cholera outbreaks, Democratic Republic of the Congo, 2008–2017. Emerg Infect Dis. 2019;25:856–64. https://doi.org/10.3201/eid2505.181141.
10. Bwire G, Mwesawina M, Baluku Y, Kanyanda SSE, Orach CG. Cross-border cholera outbreaks in sub-Saharan Africa, the mystery behind the silent illness: what needs to be done? PLoS ONE. 2016;11:e0156674. https://doi.org/10.1371/journal.pone.0156674.
11. NBS. Tanzania Census 2012—National Bureau of Statistics. 2012. http://dataforall.org/dashboard/tanzania/. Accessed 29 Oct 2018.
12. DIVA-GIS. Download data by country|DIVA-GIS. 2018. http://www.diva-gis.org/gdata. Accessed 23 Sept 2018.
13. Phelps M, Perner ML, Pitzer VE, Andreasen V, Jensen PKM, Simonsen L. Cholera epidemics of the past offer new insights into an old enemy. J Infect Dis. 2018;217:641–9. https://doi.org/10.1093/infdis/jix602.
14. Phelps MD, Azman AS, Lewnard JA, Antillón M, Simonsen L, Andreasen V, et al. The importance of thinking beyond the water-supply in cholera epidemics: a historical urban case-study. PLOS Negl Trop Dis. 2017;11:e0006103. https://doi.org/10.1371/journal.pntd.0006103.
15. Sasaki S, Suzuki H, Igarashi K, Tambatamba B, Mulenga P. Spatial analysis of risk factor of cholera outbreak for 2003-2004 in a peri-urban area of Lusaka. Zambia. Am J Trop Med Hyg. 2008;79:414–21.
16. Lugongo B. Tanzania: Government Move to Dodoma Now At 86 Per Cent. Tanzania Daily News (Dar es Salaam). 2019. https://allafrica.com/stories/201902060401.html. Accessed 8 Jul 2019.
17. Bompangue D, Giraudoux P, Handschumacher P, Piarroux M, Sudre B, Ekwanzala M, et al. Lakes as source of cholera outbreaks, Democratic Republic of Congo. Emerg Infect Dis. 2008;14:798–800. https://doi.org/10.3201/eid1405.071260.
18. Nkoko D, Giraudoux P, Plisnier P-D, Tinda A, Piarroux M, Sudre B, et al. Dynamics of cholera outbreaks in Great Lakes Region of Africa, 1978–2008. Emerg Infect Dis. 2011. https://doi.org/10.3201/eid1711.110170.
19. Rajasingham A, Hardy C, Kamwaga S, Sebunya K, Massa K, Mulungu J, et al. Evaluation of an emergency bulk chlorination project targeting drinking water vendors in cholera-affected wards of Dar es Salaam and Morogoro, Tanzania. Am J Trop Med Hyg. 2019. https://doi.org/10.4269/ajtmh.18-0734.
20. Emch M, Feldacker C, Islam MS, Ali M. Seasonality of cholera from 1974 to 2005: a review of global patterns. Int J Health Geogr. 2008;7:31. https://doi.org/10.1186/1476-072X-7-31.
21. Lemaitre J, Pasetto D, Perez-Saez J, Sciarra C, Wamala JF, Rinaldo A. Rainfall as a driver of epidemic cholera: comparative model assessments of the effect of intra-seasonal precipitation events. Acta Trop. 2019;190:235–43. https://doi.org/10.1016/j.actatropica.2018.11.013.
22. Islam MS, Zaman MH, Islam MS, Ahmed N, Clemens JD. Environmental reservoirs of Vibrio cholerae. Vaccine. 2019. https://doi.org/10.1016/j.vaccine.2019.06.033.
23. Hounmanou YMG, Leekitcharoenphon P, Hendriksen RS, Dougnon TV, Mdegela RH, Olsen JE, et al. Surveillance and genomics of toxigenic *Vibrio cholerae* O1 from fish, phytoplankton and water in Lake Victoria, Tanzania. Front Microbiol. 2019. https://doi.org/10.3389/fmicb.2019.00901.

Publisher's Note

Manuscript B: "Surveillance and Genomics of Toxigenic *Vibrio cholerae* O1 from Fish, Phytoplankton and Water in Lake Victoria, Tanzania"

Published in *Frontiers in Microbiology*

Hounmanou YMG, Leekitcharoenphon P, Hendriksen RS, Dougnon TV, Mdegela RH, Olsen JE and Dalsgaard A (2019). Front. Microbiol. 10:901. doi: 10.3389/fmicb.2019.00901

frontiers in Microbiology

ORIGINAL RESEARCH
published: 30 April 2019
doi: 10.3389/fmicb.2019.00901

Surveillance and Genomics of Toxigenic *Vibrio cholerae* O1 From Fish, Phytoplankton and Water in Lake Victoria, Tanzania

Yaovi M. Gildas Hounmanou[1], Pimlapas Leekitcharoenphon[2], Rene S. Hendriksen[2], Tamegnon V. Dougnon[3], Robinson H. Mdegela[4], John E. Olsen[1] and Anders Dalsgaard[1]*

[1] Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [2] National Food Institute, WHO Collaborating Center for Antimicrobial Resistance in Food Borne Pathogens and Genomics and European Union Reference Laboratory for Antimicrobial Resistance, Technical University of Denmark, Kongens Lyngby, Denmark, [3] Research Unit in Applied Microbiology and Pharmacology of Natural Substances, Laboratory of Research in Applied Biology, Polytechnic School of Abomey-Calavi, University of Abomey-Calavi, Cotonou, Benin, [4] Department of Veterinary Medicine and Public Health, College of Veterinary Medicine and Biomedical Sciences, Sokoine University of Agriculture, Morogoro, Tanzania

OPEN ACCESS

Edited by:
Swapan Banerjee, Health Canada, Canada

Reviewed by:
Yan Boucher, University of Alberta, Canada
Hongxia Wang, University of Alabama at Birmingham, United States

****Correspondence:***
Yaovi M. Gildas Hounmanou
gil@sund.ku.dk

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *07 January 2019*
Accepted: *09 April 2019*
Published: *30 April 2019*

Citation:
Hounmanou YMG, Leekitcharoenphon P, Hendriksen RS, Dougnon TV, Mdegela RH, Olsen JE and Dalsgaard A (2019) Surveillance and Genomics of Toxigenic Vibrio cholerae O1 From Fish, Phytoplankton and Water in Lake Victoria, Tanzania. Front. Microbiol. 10:901. doi: 10.3389/fmicb.2019.00901

The occurrence of toxigenic *Vibrio cholerae* O1 during a non- outbreak period in Lake Victoria was studied and genetic characteristics for environmental persistence and relatedness to pandemic strains were assessed. We analyzed 360 samples of carps, phytoplankton and water collected in 2017 during dry and rainy seasons in the Tanzanian basin of Lake Victoria. Samples were tested using PCR (*ompW* and *ctxA*) with DNA extracted from bacterial isolates and samples enriched in alkaline peptone water. Isolates were screened with polyvalent antiserum O1 followed by antimicrobial susceptibility testing. Whole genome sequencing and bioinformatics tools were employed to investigate the genomic characteristics of the isolates. More *V. cholerae* positive samples were recovered by PCR when DNA was obtained from enriched samples than from isolates (69.0% vs. 21.3%, $p < 0.05$), irrespectively of season. We identified ten *V. cholerae* O1 among 22 *ctx*A-positive isolates. Further studies are needed to serotype the remaining *ctx*A-positive non-O1 strains. Sequenced strains belonged to El Tor atypical biotype of *V. cholerae* O1 of MLST ST69 harboring the seventh pandemic gene. Major virulence genes, *ctxA*, *ctxB*, *zot*, *ace*, *tcpA*, *hlyA*, *rtxA*, *ompU*, *toxR*, *T6SS*, *alsD*, *makA* and pathogenicity islands VPI-1, VPI-2, VSP-1, and VSP-2 were found in all strains. The strains contained *Vibrio* polysaccharide biosynthesis enzymes, the *mshA* gene and two-component response regulator proteins involved in stress response and autoinducers for quorum sensing and biofilm formation. They carried the SXT integrative conjugative element with phenotypic and genotypic resistance to aminoglycoside, sulfamethoxazole, trimethoprim, phenicol, and quinolones. Strains contained a multidrug efflux pump component and were resistant to toxic compounds with copper homeostasis and cobalt-zinc-cadmium resistance proteins. The environmental strains belonged to the third wave of the seventh pandemic and most are genetically closely related to recent outbreak strains from

Tanzania, Kenya, and Uganda with as low as three SNPs difference. Some strains have persisted longer in the environment and were more related to older outbreak strains in the region. *V. cholerae* O1 of outbreak potential seem to persist in Lake Victoria through interactions with fish and phytoplankton supported by the optimum water parameters and intrinsic genetic features enhancing survival in the aquatic environment.

Keywords: *Vibrio cholerae*, genomics, aquatic reservoirs, African Great Lakes, microbial ecology

INTRODUCTION

Vibrio cholerae, the causative agent of cholera, is one of the oldest pathogens known to cause disease in humans, yet there is still much to be learned regarding its transmission and evolution. Although the majority of the worldś cholera cases occur in Africa, there has been limited research on the occurrence of toxigenic *V. cholerae* O1 in aquatic environments in countries like Tanzania located around the African Great Lakes (Dalusi et al., 2015a; Hounmanou et al., 2016; Bwire et al., 2018a). Knowledge on occurrence and transmission in and around the lakes are important when establishing preventive and control measures of epidemic cholera, which often affects countries situated around the lakes. Moreover, *V. cholerae* is ubiquitous in aquatic environments and phytoplankton and fish have been reported as potential reservoirs, e.g., in Tanzania (Haque et al., 2012; Rabia et al., 2017; Hossain et al., 2018a; Nyambuli et al., 2018). The tropical waters of Lake Victoria and Lake Tanganyika are likely natural habitats for the El Tor variant of *V. cholerae* O1 (Echenberg, 2011).

The incidence of cholera in countries around the Great Lakes varies by season, rainfall, plankton blooms, and level of fishing activities (Nkoko et al., 2011). *V. cholerae* O1 resistant to common antimicrobials have been isolated in fish and water from the Kenyan side of Lake Victoria (Onyuka et al., 2011). A study in Lake Tanganyika indicated a strong correlation between planktonic blooms, fish abundance and cholera (Plisnier et al., 2015). A recent study in Uganda has also identified *V. cholerae* O1 in Lake Victoria although the strains were non-toxigenic (Bwire et al., 2018a).

Tanzania has consistently been affected by cholera since 1974 and outbreaks continue to re-occur in various towns, especially in the coastal and Great Lakes regions (Urassa et al., 2009; Reyburn et al., 2011). Since the 7th pandemic reached the country in 1974, cholera has affected more than 230,596 people causing 17,714 deaths, most of which lived in lake zones (Mengel et al., 2014). Despite suggestions of genetic similarity and clonality of environmental and clinical *V. cholerae* strains (Dalusi et al., 2015b), recent studies have revealed different clonal complexes based on MLVA typing among strains implicated in cholera outbreaks in Tanzania (Kachwamba et al., 2017). Analyses of cholera outbreak strains from Uganda and Mozambique however showed a high level of similarity among strains suggesting a clonal transmission in East African countries (Garrine et al., 2017; Bwire et al., 2018b) probably facilitated by the lakes they share. Thus, there is a need to determine the relatedness of environmental and clinical isolates of *V. cholerae* and the role of the aquatic lake environments as reservoirs for the pathogen.

Furthermore, *V. cholerae* has evolved through the emergence of multidrug resistant strains with the acquisition of the SXT integrative conjugative elements (Hendriksen et al., 2011; Spagnoletti et al., 2014; Kaas et al., 2016) and transfer of the cholera toxin gene through phages (Waldor and Mekalanos, 1996). The emergence of the El Tor variant biotype carrying the *ctx*B1 gene of the Classical biotype and *V. cholerae* O1 carrying the *ctx*B7 gene of the Haitian strain is another example of evolution of *V. cholerae* and stress the importance of a continuous monitoring of the genetic characteristics of environmental *V. cholerae* (Ghosh-Banerjee et al., 2010; Kim et al., 2014).

The aim of this study was to investigate the occurrence of toxigenic *V. cholerae* O1 during a non-cholera outbreak period in Lake Victoria and determine the genetic characteristics that support environmental persistence and genetic relatedness to pandemic strains. Whole genome sequencing (WGS) and bioinformatics analysis coupled with environmental surveillance data were applied to determine the occurrence and outbreak potential of *V. cholerae* O1 in Lake Victoria. The data generated add to our understanding on aquatic reservoirs of toxigenic *V. cholerae* O1 in Lake Victoria.

MATERIALS AND METHODS

Sampling Area and Sample Collection

During the dry season of June to September 2017 and the rainy season of October to December 2017, 360 samples of lake water, carps (*Rastrineobola argentea*) and phytoplankton were collected from seven landing sites in the Mwanza Gulf of the Tanzanian basin of Lake Victoria (**Figure 1**). These included 120 water samples (60 offshore and 60 collected near the lakeshore), 120 carp samples and 120 phytoplankton samples (vertical and horizontal sampling as described below). Samples were collected during both seasons (180 samples in the dry season and 180 samples in the rainy season). Sampling sites were Shadi, Kijiweni and Mkuyuni in Mwanza South and Igombe, Mihama, Bwiru and Kayenze in Mwanza North. Mwanza is located in the Lake Victoria basin between the latitude −2° 31′ 0.01″ S and the longitude 32° 53′ 60.00″ E (**Figure 1**). Samples of fresh carp (about 200 g per sample, consisting of about 200 individual fish) were purchased from local fishermen immediately after landing at the beach and placed in sterile labeled plastic bags which were transported to the National Fish Quality Control Laboratory (NFQCL) in Mwanza in an insulated box with cooling elements and processed within 4 h of collection.

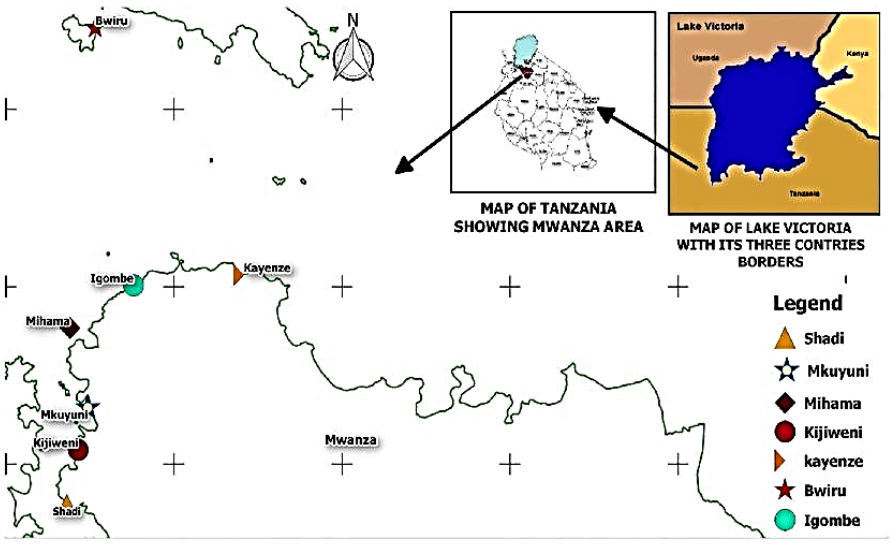


FIGURE 1 | Sampling points in Lake Victoria, Mwanza, Tanzania.

Surface water samples were collected both onshore and offshore (at least 200 m away from the nearest shore) and were collected in 250 ml sterile glass bottles and transported as shown above for carp samples.

Approximately 100 ml of phytoplankton samples were collected by boat with a phytoplankton towing net of 13-μm mesh size. To maximize chances of recovering *V. cholerae* from the Lake, phytoplankton samples were obtained from the water surface (horizontal sampling) and from beneath to top (vertical sampling) in order not to miss species/taxa that do not float on the surface water which could harbor *V. cholerae* of interest. For horizontal samples, the net was set right below the surface of the water after which the boat moved about 50 m. The net was then raised and the phytoplankton sample collected. Vertical samples were collected at a minimum depth of 5 m (up to 10 m depending on the depth offshore) where the net was submerged with a lead to collect the phytoplankton sample by a movement from bottom to top.

In the field, water temperature, pH, electrical conductivity, total dissolved solids and dissolved oxygen were measured using a portable multi-parameter meter (Hengkaituo, Guangdong, China).

Sample Preparation and Laboratory Analysis

Fresh carp samples (25 g consisting of about 20–30 individual carps) were homogenized in a stomacher bag containing 225 ml of alkaline peptone water, pH: 8.5 (APW) (Oxoid Ltd., Hampshire, United Kingdom). The homogenized samples contained carps with all body parts (skin, gills, intestines, flesh etc.). Hundred mL of water samples were filtered through a 0.45 μm pore diameter membrane (Millipore, Bedford, MA, United States) which was directly transferred into 225 mL APW; then the bottle was shaken vigorously before incubation. Similarly, phytoplankton was concentrated on a 0.45 μm pore diameter membrane filter paper (Millipore, Bedford, MA, United States) which was transferred into 225 mL APW. Samples in APW were enriched at 37°C for 6 h. Each sample was then analyzed as described below by a culture-based procedure and by PCR of DNA extracted from enriched APW samples.

After enrichment in APW, samples were streaked onto thiosulfate-citrate-bile salts-sucrose agar (TCBS) plates (Oxoid Ltd) and incubated at 37°C for 18–24 h. Characteristic dark yellow colonies with about 2 mm diameter were selected, purified and subsequently characterized by biochemical tests for identification of *V. cholerae* (Hounmanou et al., 2016). Pure cultures of presumptive *V. cholerae* colonies were then subjected to DNA extraction. DNA was extracted from boiled lysates of isolates and subjected to PCR for detection of the outer membrane protein gene (*omp*W) generating a 588 bp amplicon confirming them as *V. cholerae* (Dalusi et al., 2015a). Isolates confirmed as *V. cholerae* were subjected to agglutination with polyvalent *V. cholerae* O1 antiserum (Bio-Rad, France) and were tested for toxigenicity by PCR targeting the cholera enterotoxin subunit A gene (*ctx*A) along with the *omp*W primers in a multiplex reaction.

Antimicrobial Susceptibility Testing

Strains confirmed as toxigenic *V. cholerae* O1 were subjected to antimicrobial susceptibility testing by the Kirby-Bauer disk diffusion method with *Escherichia coli* ATCC 22925 included for quality control. Twelve different antimicrobial disks (Oxoid

Ltd) were used and results interpreted according to the Clinical and Laboratory Standards Institute (CLSI) guidelines (CLSI, 2017). Antimicrobials tested included: streptomycin (10 μg), gentamicin (30 μg), ampicillin (10 μg), nalidixic acid (30 μg), ciprofloxacin (5 μg), trimethoprim (5 μg), ceftazidime (30 μg), tetracycline (30 μg), chloramphenicol (30 μg), amoxicillin-clavulanicc acid (30 μg), cefotaxime (30 μg), and sulfamethoxazole (25 μg).

Detection of Toxigenic *V. cholerae* by PCR From Enriched Samples

About 10 ml aliquots of APW enriched samples were stored at −20°C until analysis. Total DNA was extracted using a Zymo DNA extraction Kit (The Epigenetics Company, CA, United States) following the manufacturer's protocol. DNA was extracted from 200 randomly selected samples of the 360 samples enriched in APW. A multiplex PCR was performed to target the 588-bp region of the *omp*W gene (*V. cholerae* species specific) and a 301-bp region of the subunit A of the cholera enterotoxin gene (*ctx*A). Samples that generated only one band (588 bp) were *V. cholerae* whereas samples showing both bands (588 and 301 bp) were concluded toxigenic *V. cholerae ctx*A-positive.

Primers used and PCR conditions were as previously described (Nandi et al., 2000; Dalusi et al., 2015a) with a modification of the annealing temperature which was set at 55.6°C during optimization.

Plankton Diversity

Immediately after sampling in dark bottles, 250 mL phytoplankton samples were fixed with five mL Lugol's solution and 10% formalin and then transported to the Water Quality Laboratory of the Lake Zone in Mwanza. Using an inverted microscope at resolutions of 40 × 10 and 10 × 10 and following standards keys (Verlecar and Desai, 2004), two ml of fixed sample was used by a trained taxonomist to identify taxons and species groups of phytoplankton present in the water samples.

Statistical Analysis

Proportions of positive *V. cholerae* samples from the culture-based method and PCR of DNA extracted from APW enriched samples were compared using chi-square in Epi-Info software version 7.2[1] at a probability of 0.05. Seasonal differences were also tested. Using the statistical software R v3.5.1[2] logistic regression was performed to estimate the effect of water temperature, pH, conductivity, dissolved oxygen, and total dissolved solids on the occurrence of *ctx*A-positive *V. cholerae* between sampling points.

DNA Extraction and Whole Genome Sequencing (WGS)

Eight of the ten *ctx*A-positive *V. cholerae* O1 recovered from fish, phytoplankton and water and two *ctx*A-positive non-O1 *V. cholerae* from phytoplankton and fish (Plankton1, Water1) were selected for WGS analysis. DNA was extracted using the automated Maxwell DNA extraction system (Promega Maxwell RSC, Madison, WI, United States) and the Maxwell DNA extraction and purification kit. Prior to extraction, samples were treated with Proteinase K (Sigma-Aldrich, St. Louis, MO, United States) at 56°C for 1 h followed by 10 min RNase (Sigma-Aldrich) treatment at room temperature. Concentrations of the extracted DNA were determined using a Qubit dsDNA HS assay kit (Invitrogen, United States). The DNA was run in 1% agarose gel to check for quality followed by the preparation for Illumina paired-end WGS using Illumina Miseq (Illumina, Inc., San Diego, CA, United States) according to the procedures previously described (Kaas et al., 2016). Raw sequences are submitted to the European Nucleotide Archive under the project number PRJEB30604 with the accession numbers of each sample indicated in **Supplementary Table S1**.

In silico Serogroup Typing, Multi-Locus Sequence Typing, Determination of Major Virulence Genes, Pathogenicity Islands and Phage Susceptibility

The raw paired-end reads were assembled using SPAdes assembler (Bankevich et al., 2012) available online[3]. Assembled sequences were analyzed in the batch upload pipeline of the CGE platform[4]as previously described (Thomsen et al., 2016), where KmerFinder 2.1 identified the species and possible contaminations. The species *V. cholerae* was further confirmed based on the species-specific *omp*W gene using MyDbFinder 1.2 tool (Siriphap et al., 2017) with a threshold set at 98% identity. This revealed the closest genome to the analyzed sequences as *V. cholerae* 2010EL-1786 (Reimer et al., 2011). Moreover, the sequences were further analyzed using MyDbFinder 1.2 tool with default options to identify known virulence genes and pathogenicity islands (**Supplementary Table S1**). This included the identification of *V. cholerae* serogroup-specific genes (*rfb*V-O1, *wbf*Z-O139), biotypes-specific genes (*ctx*B, *rst*R, *tcp*A), putative virulence genes (including *ctx*A), and *VC2346* specific for the 7th pandemic *V. cholerae* using a threshold of 98% identity (Kaas et al., 2016; Siriphap et al., 2017). MyDbFinder coupled with nucleotides BLAST were used to genotype the strains based on the *ctx*B of the CTX prophage that they carried. This helps to identify the wave of the seventh pandemic that the strains belonged to using *ctx*B1, *ctx*B3, and *ctx*B7 sequences (Naha et al., 2012; Kaas et al., 2016; Rashid et al., 2016). Detection of genomic islands of *V. cholerae* mainly VPI-1, VPI-2, VSP-1, VSP-2, was also carried out *in silico*. Moreover, *mak*A, *als*A and Type VI secretion system (T6SS) genes were searched. The CTX prophage genomic region containing the core region and RS1 was also assessed in the samples. Detection of the PICI like elements, i.e., PLE1 and PLE2, responsible for phage susceptibility in *V. cholerae* was performed using MyDbFinder 1.2. The MLST 2.0 tool reached through the batch upload pipeline analyzed the sequence types (ST) of the *V. cholerae* strains in order to establish genetic relatedness with known STs that have

[1]https://www.cdc.gov/epiinfo/index.html
[2]https://cran.r-project.org/
[3]https://cge.cbs.dtu.dk/services/SPAdes/
[4]https://cge.cbs.dtu.dk/services/cge/

been implicated in cholera outbreak. The analysis was based on the seven housekeeping genes: *adk*, *gyrB*, *metE*, *mdh*, *pntA*, *purM*, and *pyrC* (Kaas et al., 2016). Moreover, a ribosomal multilocus sequence typing which is a more discriminative MLST typing method based on variations of the 53 genes encoding the bacterial ribosome protein subunits (*rps* genes) was used to investigate differences among the environmental strains as previously described (Jolley et al., 2012).

Identification of Antimicrobial Resistance Genes, SXT Element, Class 1 Integron and Plasmids Typing

In the batch upload pipeline (Thomsen et al., 2016), ResFinder 3.0 assessed acquired antimicrobial resistance (AMR) genes in the assembled sequences. In ResFinder with default options, we also searched for beta-lactam resistance genes including *blaVCC-1*, a carbapenamase gene that is emerging in environmental *V. cholerae* as recently reported in strains from Canada and Germany (Mangat et al., 2016; Hammerl et al., 2017).

Detection of AMR genotype was further strengthened with the search for mobile genetic elements using MyDbFinder 1.2 (Larsen et al., 2012) for the detection of the SXT integrative conjugative element, the different classes of integrons, and the presence of mutations in the DNA gyrase (*gyr*A gene) and in the DNA topoisomerase IV (*par*C gene) (Siriphap et al., 2017). MyDbFinder 1.2 was used to compare sample sequences with the integrating conjugative elements (ICE) of *V. cholerae* O1 ICE*Vch*Hai1(JN648379) and of SXTMO10 (AY034138) at a threshold of 98% identity. Due to high levels of mutations in the SXT element (Wang et al., 2016) the threshold for detection of *int*SXT was set at 95% for % ID of 40% minimum length.

The search for plasmids was done in three steps. The PlasmidFinder 1.3 tool of the batch upload pipeline in CGE (Thomsen et al., 2016) was used to search for plasmid replicons. Due to the limitation of PlasmidFinder, which detects only replicons that are available in the CGE database, we conducted another specific search. This second step used MyDbFinder 1.2, where a local search was performed between our samples' genomes and sequences of representative IncA/C plasmids (known for carrying multidrug resistance genes in *V. cholerae*) available from the Genbank (Accession numbers: KY399978, KM083064, KF551948, CP007636, CP033514, and KJ817377) (Carraro et al., 2014; Folster et al., 2014; Wang et al., 2018). This step also involved two cryptic plasmids recently isolated in *V. cholerae* strains (Acc. KY486774 and KY486775) (Ceccarelli et al., 2017). In the third step, we used Blast atlas in GView[5] where genomes were analyzed against reference plasmids.

Moreover, the genome sequences were annotated in RAST v.2.0 (Brettin et al., 2015) and each annotated file was analyzed through the SEED viewer (Overbeek et al., 2014) for subsystem categorization of the genetic elements involved in survival and persistence in the aquatic environment.

Single Nucleotide Polymorphism-Based Phylogenetic Analyses

To further identify evolution and genetic similarities within and between the environmental *V. cholerae* isolates as compared to cholera outbreak strains, assembled genomes were analyzed using CSIPhylogeny version 1.4 with default options (Kaas et al., 2014), where high quality SNPs of the environmental genomes were identified. Sequences of *V. cholerae* from Uganda and Kenya that have borders to Lake Victoria as well as previous outbreak strains from Tanzania were obtained from the GenBank and compared with our environmental strains to determine genome wide SNPs. Additionally, *V. cholerae* O1 outbreak strains from other countries of the African Great Lakes region including the Democratic Republic of Congo (DRC), Burundi and Rwanda but also from Zambia were included in the tree to access the spatial-temporal phylogenetic evolution of the environmental strains. Accession numbers of strains used in the SNP tree are reported in **Supplementary Table S2**. This generated a regional and time-scale phylogenetic tree showing the genetic relatedness and evolution between the environmental non-outbreak *V. cholerae* O1 strains and pandemic strains from the African Great Lakes region. The reference genome of *V. cholerae* O1 strain N16961 (Biosample SAMN02603969) was used to root the tree (Kaas et al., 2016). The Newick files obtained in CSIPhilogeny 1.4 were downloaded and the final tree was amended in iTOL[6].

RESULTS

Detection of *V. cholerae* in Samples From Lake Victoria

Using standard bacteriological procedures followed by PCR for the species-specific gene *omp*W, 108 isolates were identified as *V. cholerae*. These strains were isolated from 77 out of 360 samples (21.4%) analyzed with up to three characteristic colonies being selected from each TCBS agar plate for confirmation. However, APW-enriched samples subjected directly to multiplex PCR yielded 138 DNA samples that were positive for *V. cholerae* (*omp*W) out of the 200 samples selected for analysis (69.0%). A Chi-square comparison revealed that PCR of DNA extracted directly after enrichment in APW recovered significantly higher proportions of *V. cholerae*-positive samples than identification of *V. cholerae* from isolates obtained on TCBS agar plates (69% vs. 21.4%, $p < 0.05$). A total of 22 *V. cholerae* isolates out of the 108 isolates originally recovered were positive for the subunit A of the cholera enterotoxin gene (*ctx*A). Ten of these strains belonged to serogroup O1 and originated from phytoplankton (five isolates), carps (three isolates), and water (two isolates) samples. Phytoplankton collected offshore by the vertical sampling method yielded all five *V. cholerae* O1 strains from this source. Two water samples collected offshore contained *V. cholerae* O1 (**Table 1**).

In DNA extracted from APW enriched samples, 23 samples contained the *ctx*A gene in phytoplankton (14), water (6)

[5]https://server.gview.ca/

[6]https://itol.embl.de/

TABLE 1 | Detection of toxigenic *V. cholerae* O1 by culture and PCR.

Sample type	Sampling	Number of samples	Total *ctx*A-positive samples[a]	Total *ctx*A-positive O1 samples[a]	Total *ctx*A positive samples (APW-PCR)[b]
Phytoplankton	Horizontal method	60	4	0	5
	Vertical method	60	6	5	9
Water	Onshore	60	4	0	4
	Offshore	60	3	2	2
Carps	Landing sites	120	5	3	3

[a]*Strains isolated on TCBS agar and confirmed by PCR;* [b] *Toxigenic V. cholerae identified by PCR of samples pre-enriched in alkaline peptone water (APW).*

and carps (3). Fifteen of the PCR-positive samples did not yield any colonies by the culture-based technique. The vertical phytoplankton sampling method yielded slightly more toxigenic *V. cholerae ctx*A-positive samples than the horizontal sampling method ($p > 0.05$), which was also the case for samples collected onshore compared with offshore water samples (**Table 1**). The total DNA extracted from samples enriched in APW were however, not used for further characterization because the quality of the DNA does now allow for sequencing analysis.

Seasonal Variation of *V. cholerae* in Lake Victoria

There was no statistical seasonal difference ($p > 0.05$) in recovery of *V. cholerae* when DNA was obtained from pure cultures (32.8 vs. 27.2%, dry and rainy seasons, respectively) compared to when DNA was obtained from APW-enriched samples (71.0 vs. 67.0%). Nevertheless, a statistically higher proportion of *ctx*A-positive *V. cholerae* was obtained in the dry season as compared to the rainy season for samples analyzed by culture procedures ($p = 0.04$).

Physico-Chemical Water Parameters and Phytoplankton Diversity

The water temperature at the different sampling sites ranged from 24.9 to 25.9°C and water pH ranged between 7.5 and 8.5. The electrical conductivity fluctuated between 90 and 150 μS/cm (**Supplementary Table S3**). None of these parameters varied significantly between sampling points when they were regressed on the total toxigenic *V. cholerae ctx*A-positive samples ($p > 0.05$). The analyzed phytoplankton samples contained 45 different taxa belonging to three major phytoplankton groups including green algae (62.2%), cyanobacteria (31.1%), and diatoms (6.7%).

Genomic Characterization of the *V. cholerae* Strains

We sequenced ten strains including eight of the isolated *ctx*A-positive *V. cholerae* O1 and two *ctx*A-positive *V. cholerae* non-O1, which were characterized by analysis of WGS data. Based on WGS analysis, all the ten sequenced strains including the two phenotypically non-O1 contained the *rfbV-O1* gene and should therefore be regarded as of serogroup O1. **Table 2** shows genomic characteristics of the strains when variations was seen (see further details in the **Supplementary Table S1**), whereas similar characteristics are described in the text below. All strains belong to the third wave of the seventh pandemic as they are all atypical El Tor biotype variants of *V. cholerae* O1, carrying the *ctx*B7 genotype of the *ctx*B gene while possessing the *rst*R and *tcp*A genes of El Tor biotype. The *in silico* MLST revealed that all strains belonged to the same sequence type ST69 and harbored the seventh pandemic-specific gene (VC2346) suggesting that they belong to the same clonal linage.

The occurrence of virulence-associated genes and pathogenicity islands among the environmental *V. cholerae* was similar in all sequenced strains. This included the major virulence-associated genes such as *ctxA*, *ctxB*, *zot*, *ace*, *tcpA*,

TABLE 2 | Genomic sequence data, virulence profile and occurrence of antimicrobial resistance genes in the *V. cholerae* strains.

Strain ID	No. of contigs	Genome size (bp)	rMLST	intSXT	SXT/R391
Fish1	63	40505002	rST14417	+	ICEVchHai1
Fish2	71	4029796	rST78290	+	ICEVchHai1_del[a](floR,strA/B,sul2)
Fish3	63	40505002	rST14417	+	ICEVchHai1
Fish4	71	4029706	rST78290	+	ICEVchHai1_del(floR,strA/B,sul2)
Plankton1	127	4009065	rST78277	–	Not Found
Plankton2	81	4022914	rST78290	+	ICEVchHai1_del(floR,strA/B,sul2)
Plankton3	89	4008427	rST78290	+	ICEVchHai1_del(floR,strA/B,sul2)
Water1	84	4022791	rST78290	+	ICEVchHai1_del(floR,strA/B,sul2)
Water2	66	4048306	rST14417	+	ICEVchHai1
Water3	98	4030711	rST78290	+	ICEVchHai1_del(floR,strA/B,sul2)

[a]*del, deletions in specified genes.*

hlyA, *mshA*, *rtxA*, *ompU*, and *toxR*, whereas *stn* and *chxA* genes were absent. Moreover, the strains all contained genes of the Type VI secretion system, the gene involved in glucose metabolism, *als* and the flagella-mediated cytotoxin gene *mak*A. All strains contained the pathogenicity islands VPI-1, VPI-2, VSP-1, and VSP-2.

Phenotypic and Genotypic Antimicrobial Resistance, Mobile Genetic Elements and Phage Resistance

All ten sequenced *V. cholerae* strains showed phenotypic resistance to streptomycin, while eight of them showed resistance to amoxicillin and clavulanic acid. Phenotypic resistance to ampicillin and sulfamethoxazole was observed in 6/10 strains whereas resistance to nalidixic acid and trimethoprim was observed in three strains. All isolates were susceptible to gentamicin, ciprofloxacin, ceftazidime, tetracycline, chloramphenicol and cefotaxime (**Figure 2**). Sequence analysis revealed, however, that some genes encoding resistance to certain antimicrobial classes were present in strains that are phenotypically susceptible. For instance, all strains harbored the *cat*B9 gene conferring resistance to chloramphenicol, but were phenotypically susceptible. Similarly, all strains except the strain Plankton1 harbored the *dfr*A1 gene conferring resistance to trimethoprim, while only three of them showed phenotypic resistance.

Moreover, all strains except Plankton1 contained the SXT integrative conjugative element (*int*SXT gene), but they all lacked specific integrase genes of the class 1 integron (*int*I gene). Nevertheless, not all the antimicrobial resistance genes normally found on the SXT element such as *sul*2, *dfr*A1, *dfr*A18, *flo*R, *str*A, and *str*B which are associated with sulfamethoxazole and trimethoprim, chloramphenicol and streptomycin resistance, respectively were present in the strains according to MyDbFinder 1.2. Therefore, a Blast Atlas analysis was performed where the genomes where compared to the reference *V. cholerae* strain 2010EL-1786. This revealed that except for strains Fish1, Fish3, and Water2 that were positive for aminoglycoside, phenicol and sulfonamide resistance genes, the remaining strains showed nucleotide deletions in their genomes (**Figure 3**). A more detailed analysis of the strains that have deletions showed that they contained fragments of the concerned genes (*str*A, *str*B, *Sul*2, *flo*R) on the ICE fragment, with about 1100 bp gaps between bp position 98500–102450 (**Figure 3**).

Strains Fish1, Fish3, and Water2 showed resistance to nalidixic acid and contained expected amino acid substitutions in gyrA (Ser83-Ile) and parC (Ser85Leu), but none of the ten strains contained the fluoroquinolone resistance gene *qnr*VC1. Likewise, none of the ten strains contained beta-lactam resistance genes including *blaVCC-1*, a carbapenamase gene. Nine environmental strains harboring the SXT element had a genomic organization of the integrating conjugative element (ICE) similar to that of the *V. cholerae* ICEVchHai1 (Reimer et al., 2011) reference strain with some common deletions seen in loci VC1786ICE78 for all strains and in VC1786ICE14 for strains Fish1, Fish3 and Water2. The other six strains had deletions in loci VC1786ICE6-13, VC1786ICE81, and VC1786ICE83-84 (**Supplementary Table S4**).

Our strains have no plasmid replicons according to PlasmidFinder. Moreover, in MyDbFinder and in Blast atlas, the strains had no copies of the IncA/C plasmid, as well as the cryptic plasmids pSDH1-2. The 10 strains irrespective of their biotype lacked the phage susceptibility region of the PICI like elements (PLE1, PLE2).

Genetic Elements Supporting Environmental Persistence in Studied Genomes

All ten analyzed genomes present molecular machinery for attachment, survival and defense for environmental persistence. The magnesium and cobalt efflux protein (*Cor*C), the cobalt-zinc-cadmium resistance protein (*Czc*D) and the multidrug efflux pump component (*Mtr*F) were conversed in the defense system of all the sequenced strains (**Figure 4**). Moreover, the strains all possess the *msh*A gene involved in attachment to chitin, as well as autoinducers (AI-2 LuxP and LuxQ) involved in quorum sensing and biofilm formation for environmental survival. A number of environmental stress response regulator proteins were conserved in the strains mainly the response regulators of the *Vie*SAB transduction system of *Vibrio*, the two-component response regulator proteins, histidine kinase, and Vibrio polysaccharides (VPS) biosynthesis proteins (**Figure 4**).

Phylogenetic Comparison Analysis of the Environmental *V. cholerae* O1 and Cholera Outbreak Strains in Countries Surrounding Lake Victoria

Overall, the SNP tree revealed a wide diversity among the environmental strains with up to 174 SNP differences (found between strains Plankton1 and Plankton3). Strains Fish1 and Fish3, as well as Fish2 and Fish4 were strictly clonal, respectively. Very close relatedness was also observed between strains Water2, Fish1, and Fish3 (only four SNPs difference), between Water2, Fish2, and Fish4 (only five SNPs difference) and between strains Plankton2, Fish2, and Fish4 (only nine SNPs difference). We further performed a ribosomal MLST typing (Jolley et al., 2012) to assess the diversity of the environmental strains and found little diversity as six of the ten strains belong to the same rMLST type ST78290. The strain Plankton1, which shows above 100 SNP differences compared to its pairs, belongs to a separate rMLST type 78277 whereas strains Fish1, Fish3, and Water2 belong to rMLST type 14417.

When compared to published genomes of the Genbank, the environmental *V. cholerae* O1 strains were found in three different clusters among the pandemic strains. The strain Plankton1 was distant from all genomes as a singleton but closer to strains associated with cholera outbreaks in Tanzania from 1993 and 1997 with 88 to 95 SNPs (**Figure 5A** and **Supplementary Table S5**). Strains Fish1, Fish3, and Water2, however, formed a separate cluster (29–33 SNPs, **Supplementary Table S6**) with a 2013 pandemic strain from DRC (**Figure 5B**). This clade also contains older pandemic

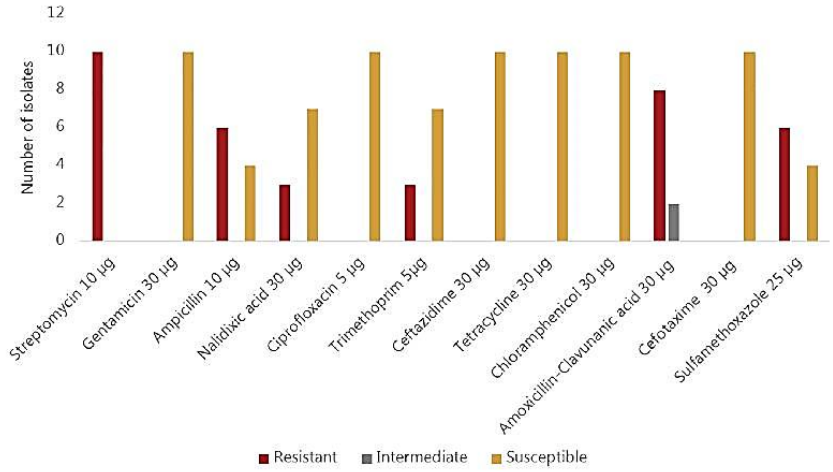


FIGURE 2 | Phenotypic antimicrobial susceptibility of environmental *Vibrio cholerae* O1 strains.

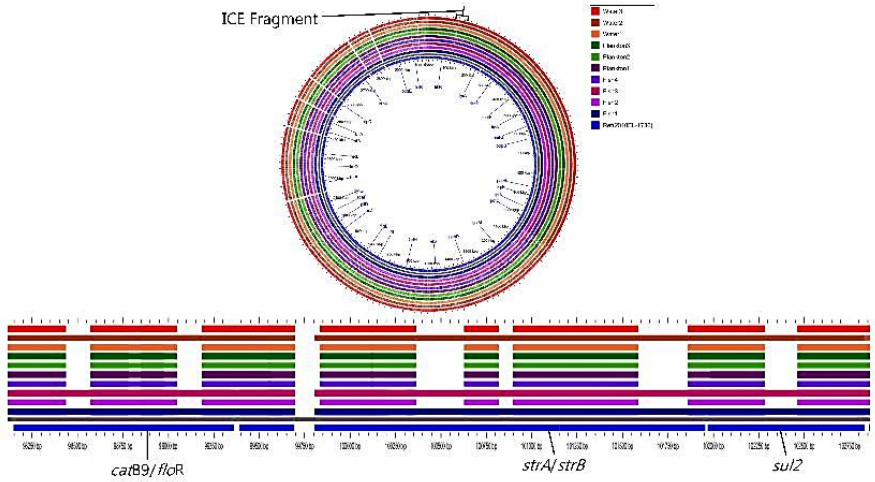


FIGURE 3 | Nucleotide deletions in resistance genes on the ICE fragment.

strains from Burundi (2001) and Tanzania (2009). The other environmental strains were highly related to recent epidemic strains with less than 22 SNPs difference between our environmental O1 strains Fish2, Fish4, Plankton2, Plankton3, Water1, Water 3 and the *V. cholerae* O1 strains associated with cholera outbreaks in Tanzania and Kenya in 2015 and in Uganda in 2016 (**Figure 5C**). The strains Fish2 and Fish4 for instance, were highly clonal when compared to the 2015 Tanzanian and Kenyan outbreak strains (only three SNPs apart; **Figure 5C**) and the Uganda outbreak strain (nine SNP differences) (Biosample Accession SAMN08744331). The largest difference in SNPs between the environmental and pandemic *V. cholerae* O1 strains was 183, found between strain Plankton1 and one strain associated with the Ugandan cholera outbreak (Biosample Accession: SAMN08744333). The pairwise SNP differences among all strains are reported in **Supplementary Table S5**.

DISCUSSION

We report the occurrence of toxigenic *V. cholerae* O1 and non-O1 in the aquatic environment of the Tanzanian basin of Lake Victoria during a non-cholera outbreak period. The presence

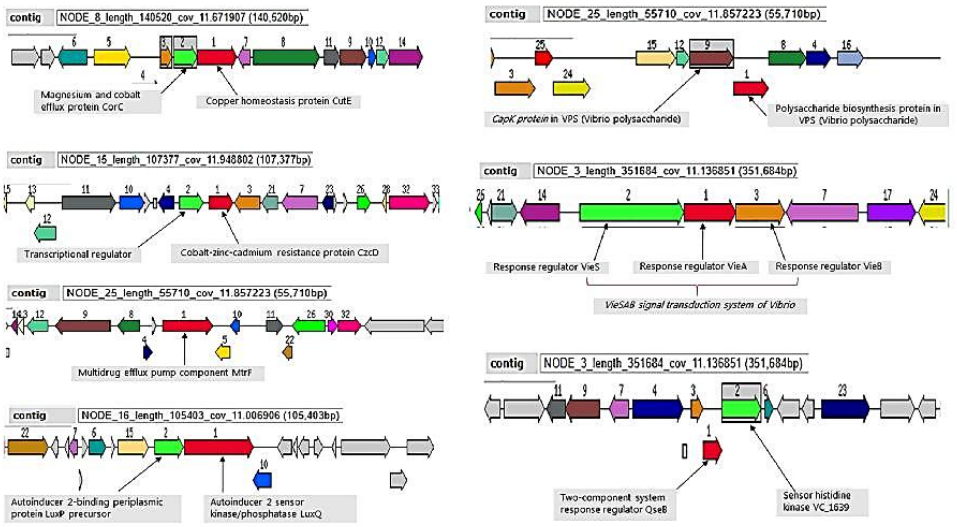


FIGURE 4 | Selected genetic elements encoding for environmental fitness. Strain Plankton2 was used for this illustration. The contig in which the genetic complex are found is indicated.

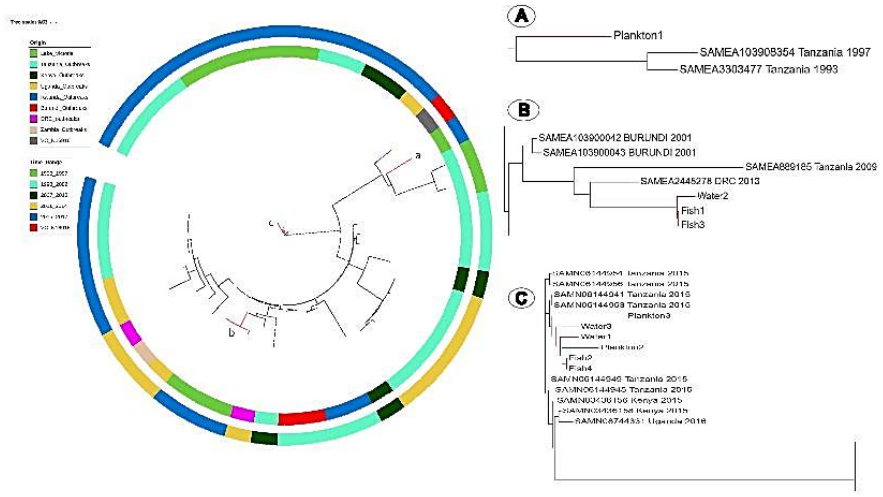


FIGURE 5 | SNP-Tree showing spatial-temporal phylogeny of environmental strains and pandemic *V. cholerae* from seven countries in the African Great Lakes Region. Colors on the inner lane indicate the origin of strains and the outer lane colors indicate the timeline evolution of the strains as shown in the legend. The tree was rooted with the reference *V. cholerae* N16916 (blue branch). Red branches identify the position of environmental genomes on the tree. **(A)** Localization of the strain Plankton1 in a clade of older strains involved in the initial waves of the seventh pandemic; **(B)** position of Fish1, 3 and Water2, and **(C)** a clade of recent outbreak strains in Tanzania, Kenya, and Uganda with most environmental strains.

of cholera toxin-positive *V. cholerae* non-O1 in Lake Victoria suggests that the *ctx*A gene is not limited to only the O1/O139 serogroups. It should be noted that we did not agglutinate the non-O1 strains in O139 antiserum. However, *V. cholerae* O139 do not seem to occur in Africa and a recent report on non-toxigenic O139 needs to be confirmed (Bwire et al., 2018a). Studies have identified *V. cholerae* O141 and O75 harboring the CTX prophage including the *ctx*A gene which have been responsible for cholera-like outbreaks, e.g., in the United States (Dalsgaard et al., 2001; Crump et al., 2003; Haley et al., 2014). Moreover, the occurrence

of non-O1 toxigenic strains in Lake Victoria underlines that serotyping beyond O1 and O139 antisera should be implemented as the CTX prophage seem to be present in more serogroups than initially documented (Waldor and Mekalanos, 1996).

The two toxigenic non-O1 *V. cholerae* sequenced in this study showed that they were in fact O1 serogroup strains based on the presence of the *rfb*V-O1 gene. This discrepancy between the phenotype and the genotype could be due to changes in the O antigen expression leading to delayed and non-conclusive agglutination in these strains during testing. It is unlikely that our phenotypically non-O1 *ctx*A-positive strains might be of another serogroup than O1 since they possess the *rfb*V-O1 gene but also belong to the atypical El Tor biotype of serogroup O1.

The most recent cholera outbreak in the study area occurred in Mwanza in March 2016 (Unpublished report from the Tanzanian Ministry of Health and Social Welfare). Thus, the current study was conducted more than a year after the latest reported cholera outbreak; yet toxigenic *V. cholerae* O1 strains were recovered from phytoplankton, fish (carps) and water samples, suggesting that Lake Victoria is a reservoir of *V. cholerae*. Isolation of *V. cholerae* O1 strains between inter-epidemic periods are rare (Alam et al., 2006; Faruque and Mekalanos, 2012) partly because *V. cholerae* O1 may become dormant, i.e., viable-but-non-culturable (Kamruzzaman et al., 2010; Bari et al., 2013). Nevertheless, our results are consistent with other findings, which also detected toxigenic strains of *V. cholerae* O1 in aquatic environments during non-cholera outbreak periods (Onyuka et al., 2011; Dalusi et al., 2015a; Hounmanou et al., 2016). Furthermore, the sequenced environmental *V. cholerae* isolates belong to the seventh cholera pandemic lineage similar to strains that caused recent epidemics (2015–2016) in countries surrounding Lake Victoria (Tanzania, Kenya, and Uganda) with SNPs ranging between 3 and 22 for samples Fish2, Fish4, Plankton2, Plankton3, Water1, and Water 3. This very close genetic relatedness to outbreak strains further substantiates that such environmental isolates are of outbreak potential given that their closest progenitors are epidemic strains. The environmental *V. cholerae* O1 isolates are likely of outbreak origin and potential progenitors of cholera outbreak strains. On the other hand, the spatial-temporal analysis of our environmental strains revealed a regional spread and long-term environmental persistence of pandemic *V. cholerae* within the African Great Lakes region. The close genetic relatedness of strain Plankton1 to the 1993 and 1997 pandemic strains from Tanzania suggests that this strain did not emerge from the recent outbreaks like the others but has probably persisted in the lake for at least 20 years. Although Plankton1 is an atypical El Tor strain whereas the 1993 and 1997 pandemic strains were prototype El Tor, their closer phylogenetic relatedness compared to the other environmental and pandemic strains show that this strain most probably underwent *ctx*B-mediated mutations throughout the years to become atypical El Tor (Rashid et al., 2016). Furthermore, Fish1, Fish3, and Water2 were distant from their environmental counterparts but closely related to older pandemic strains from DRC, Tanzania and Burundi. The finding of strains Fish1, Fish3, and Water2 suggests not only a long-term persistence of *V. cholerae* in the lake but also an environmental adaptation aided by a human-mediated spread between the countries in the African Great Lakes region enhanced by connecting water bodies. This phylogenetic heterogeneity of *V. cholerae* isolated in Lake Victoria substantiate that the lake serves as reservoir for long-term persistence of pandemic clones of *V. cholerae* with a much-dispersed geographical distribution. Sequence analysis revealed that the *V. cholerae* O1 strains possessed all major virulence-associated genes found in clinical O1 strains, e.g., the cholera enterotoxin genes, toxin co-regulated pilus, hemolysis genes, as well as the flagella-mediated toxin gene *mak*A (Aliabad et al., 2012; Castillo et al., 2018; Dongre et al., 2018). Genes involved in the T6SS that enables *V. cholerae* to overcome commensals and immune cells in the human gut during infection were also present in our environmental isolates (Unterweger et al., 2014; Logan et al., 2018). Moreover, the environmental strains were *V. cholerae* O1 El Tor variants of the Haitian *ctx*B7 genotype and therefore similar to strains implicated in the ongoing third wave of the seventh pandemic known to cause more severe disease as compared to the typical El Tor biotype (Ghosh-Banerjee et al., 2010; Mutreja et al., 2011; Kim et al., 2014). The El Tor biotype variants of *V. cholerae* O1 are proposed to be common in the African Great Lakes region with potential to cause cholera outbreaks (Echenberg, 2011). Our study provides evidence that cholera outbreak isolates may persist in the aquatic environment during non-outbreak periods. This may explain why clonal isolates are found to cause different outbreaks many years apart, e.g., as seen in Mozambique (Garrine et al., 2017).

Vibrio cholerae of sequence type ST69 belonging to the third wave of the seventh cholera pandemic are the ones responsible for most cholera outbreaks worldwide (Mutreja et al., 2011; Kachwamba et al., 2017; Weill et al., 2017). The presences of the pathogenicity islands VPI-1, VPI-2, VSP-1, and VSP-2 in the strains further confirm their outbreak potential. Our strains clustered in two major ribosomal MLST types mainly rMLST 78290 and 14417, which are identical to rMLST types of cholera outbreaks strains in the region (Mutreja et al., 2011; Kachwamba et al., 2017). We observed that strains of rMLST type 78920 clustered together on the SNP tree with outbreak strains of the same rMLST with strains of rMLST type 14417 clustering in a similar manner (**Figure 5**). Most cholera outbreak strains isolated from 2015 and the majority of our environmental strains are of an identical rMLST type 78920 suggesting not only a transmission of strains between the aquatic environment and humans, but also a time-scale evolution of *V. cholerae* from one rMLST type to another. Our results suggests that Lake Victoria is a reservoir for *V. cholerae* O1 with outbreak potential which are consistent with findings in previous studies (Faruque et al., 1998, 2007; Dalusi et al., 2015b), but are not in agreement with a recent report which questioned whether environmental strains may be progenitors of outbreak strains (Weill et al., 2017). In Haiti, a close phylogenetic relationship has also been reported between clinical and environmental toxigenic *V. cholerae* O1 strains (Azarian et al., 2014). Moreover, some *ctx*-negative environmental *V. cholerae* O1 clustered with strains responsible for clinical cholera and possessed genomic characteristics of the seventh pandemic lineages (Azarian et al., 2016). This provides further arguments that the aquatic environment is a reservoir

in the survival and generation of progenitors of outbreak strains of *V. cholerae*.

The persistence of epidemic *V. cholerae* O1 strains in the lake can be attributed to a number of environmental factors. One is the presence of suitable hosts. The symbiotic relationship between phytoplankton and *V. cholerae* is well documented and known to enhance the survival and maintenance of the pathogenicity in these strains even after cholera outbreaks have ceased (Haque et al., 2012; Mitchell et al., 2017). Attachment to aquatic organisms to take advantage of chitin and other nutrients is dependent on the MSHA (mannose-sensitive hemagglutinin) pili (Meibom et al., 2004; Sinha-Ray and Ali, 2017), which were present in our *V. cholerae* strains and whose expression by the *msh*A gene increases in warm tropical water with temperatures above 15°C (Lutz et al., 2013). The conductivity measurements and the surface water temperature in the lake averaging 25°C with an alkaline pH averaging 8.5, provides optimum conditions for survival and growth of *V. cholerae* (Lugomela et al., 2014; Dalusi et al., 2015a; Plisnier et al., 2015). The WGS data analysis further indicates that the *V. cholerae* strains have survived in and adapted to the lake environment for about two decades. Among phytoplankton groups that may harbor *V. cholerae* O1 in Lake Victoria, green algae and cyanobacteria were the major species detected in our study which are both documented aquatic reservoirs for *V. cholerae* O1 and O139 in Asia and Africa (Eiler et al., 2007; Dalusi et al., 2015b; Islam et al., 2015). We noticed that most *V. cholerae* O1 strains were recovered in phytoplankton samples obtained in deep water which suggests that the vertical sampling technique for phytoplankton should be used in future environmental surveillance studies. Food safety aspects and transmission of *V. cholerae* O1 in carps to humans are questionable as carps are sun-dried before consumption, but needs to be further explored. *V. cholerae* O1 was isolated in Nile perch from the Kenyan side of the Lake Victoria (Onyuka et al., 2011), but the role of the small carps as reservoir hosts of *V. cholerae* O1 remains uncertain. Experimental studies suggest that *V. cholerae* strains can use accessory toxins and structural components to survive in the gut of zebrafish over long periods (Runft et al., 2014; Mitchell et al., 2017). Moreover, Hilsa fish from Bangladesh and many other species were also reported aquatic hosts of *V. cholerae* (Halpern and Izhaki, 2017; Hossain et al., 2018a).

Although seasonal variation was not a significant predictor of recovery of toxigenic strains in this study, *V. cholerae* O1 are generally more frequently isolated during rainy seasons (Reyburn et al., 2011). Our strains were negative for the PICI-like elements responsible for phage susceptibility. These elements are genetic islands of *V. cholerae* that inhibit lytic phages and therefore serve as a resilience factor toward vibrio phages (Seed et al., 2013). It has been shown that seasonal fluctuations of cholera cases can be associated with predation of *V. cholerae* O1 by phages in the environment leading to a decrease in the epidemic curve (Faruque and Mekalanos, 2012). A long-term persistence of epidemic *V. cholerae* O1 clones in the lake can also be attributed to their adaptive genetic machinery made of defense genes against toxic compounds and heavy metals. The strains contained magnesium and cobalt efflux protein (*Cor*C) and the cobalt-zinc-cadmium resistance protein (*Czc*D). Studies have found different levels of heavy metals and pesticide residues in Lake Victoria (Kishe and Machiwa, 2003; Ogwok et al., 2009), which could be detrimental to bacteria like *V. cholerae* and limit their growth. The capability of *V. cholerae* O1 strains to survive in such an environment is justifiable by the genetic elements enhancing their ability to resist these toxic compounds. Moreover, in a complex environment like Lake Victoria, *V. cholerae* may be subjected to a number of other environmental stresses. The presence of autoinducers, AI-2 LuxP and LuxQ (Joelsson et al., 2007) in the genetic makeup of our strains favors their persistence since they are capable of quorum sensing and forming biofilm to survive and thrive (Kamruzzaman et al., 2010; Bari et al., 2013; Sinha-Ray and Ali, 2017). The strains also contained various stress response regulator proteins like the *Vie*SAB transduction system, the two-component response regulator proteins, and Vibrio polysaccharides (VPS) biosynthesis proteins, which are essential for environmental fitness and persistence in *V. cholerae* (Fong et al., 2010; Vesth et al., 2010; Lutz et al., 2013). Thus, the molecular characteristics of our *V. cholerae* strains substantiate their spatial-temporal persistence in the lake from where they may emerge and cause outbreaks.

As the concentration of *V. cholerae* O1 is low in the aquatic environment between cholera outbreaks, it is important to apply methods with high sensitivity for environmental surveillance (Bwire et al., 2018a; Nyambuli et al., 2018). We observed that PCR of DNA extracted from APW-enriched samples detected more *V. cholerae*-positive samples than using subculture onto TCBS agar with subsequent confirmation of selected isolates by PCR. Similar observations were reported analyzing sediments and plankton from Tanzanian estuaries using culture-independent PCR techniques (Dalusi et al., 2015a). Viable but non-culturable forms of toxigenic and non-toxigenic *V. cholerae* are important in the ecological dynamics and epidemiology of cholera (Kamruzzaman et al., 2010; Faruque and Mekalanos, 2012). However, these forms of *V. cholerae* cells cannot be detected by culture directly on selective agar media. We therefore suggest the detection of toxigenic *V. cholerae* by PCR of APW-enriched samples as a sensitive and efficient method in routine monitoring of *V. cholerae* in the environment.

Increasingly, antimicrobials are used in cholera treatment to reduce the volume of diarrhea, and *V. cholerae* O1 concentrations in stools. Our study revealed that isolates were resistant to commonly used antimicrobials such as streptomycin, amoxicillin + clavulanic acid, ampicillin, trimethoprim and sulfamethoxazole; findings which are similar to earlier studies of *V. cholerae* from Tanzania (Hounmanou et al., 2016; Rabia et al., 2017). Most of our strains carried the SXT integrative conjugative element containing resistance genes to aminoglycosides, sulfamethoxazole, trimethoprim, phenicols and quinolones with genomic similarity to *V. cholerae* ICEVchHai1 (Hendriksen et al., 2011; Kaas et al., 2016). Some of our strains carrying genes encoding phenicol resistance were susceptible to chloramphenicol (Siriphap et al., 2017; Hossain et al., 2018b) which shows that the expression of a phenotype is not solely related to the possession of encoding gene. Moreover, the nucleotide deletions that occurred within the SXT element are

not rare, as studies have underlined frequent mutation rates in the ICE cassette (Spagnoletti et al., 2014; Wang et al., 2016; Hossain et al., 2018b). This could also explain why the strain Plankton1, which was negative for the SXT element, was genetically distinct from the rest (**Figure 5**). Plasmids were not a significant factor encoding antimicrobial resistance because antimicrobial resistance genes in our *V. cholerae* strains was rather related to the SXT element (Wang et al., 2016). The lack of plasmids in our strains may also explain the absence of beta-lactam genes as beta-lactamases are commonly carried by conjugative plasmids (Kudirkiene et al., 2018). Nevertheless, some cryptic plasmids have been found in non-O1 strains (Ceccarelli et al., 2017), while the IncA/C conjugative plasmids responsible for multidrug resistance were detected in some *V. cholerae* in Haiti (Wang et al., 2018). Acquisition of these resistance profiles not only is of clinical therapeutic relevance but may also be part of the defense mechanisms of *V. cholerae* in aquatic environments as protection against antimicrobial compounds and antimicrobial residues that may be present in the lake. Besides the ICE element, all sequenced strains contained a multidrug efflux pump component (*Mtr*F) which could support their ability to persist in the lake in case of exposure to antimicrobial residues favoring their environmental persistence as it may also enhance their interaction with aquatic organisms (Alvarez-Ortega et al., 2013), as well as virulence (Alcalde-Rico et al., 2016).

This study reports the occurrence of multidrug resistant *V. cholerae* O1 in Lake Victoria that are genetically closely related to recent pandemic strains in Tanzania, Kenya and Uganda. The strains identified are also closely related to older pandemic strains recovered in the Democratic Republic of Congo, Burundi and Tanzania up until 1993, suggesting a long-term persistence and wide spatial distribution of pandemic strains within the region with the lake serving as a reservoir. These environmental isolates likely emerged from previous cholera outbreaks and survived in the lake environment for decades through various relations with reservoirs such as phytoplankton and fish. The *V. cholerae* O1 strains in the lake are potential progenitors of future cholera outbreak strains. Our findings are important for surveillance of *V. cholerae* O1 and understanding the epidemiology of cholera in countries around the lake.

ETHICS STATEMENT

The present study required no ethical approval since the analyzed samples were collected from the environment, namely water, phytoplankton and carp from Lake Victoria. We, however, obtained a research permit from local authorities in Mwanza, the study area for collection of these samples in the Mwanza basin of Lake Victoria. The permit issued in Kiswahili Language is submitted with the Manuscript only for editorial use.

AUTHOR CONTRIBUTIONS

YH collected samples and carried out the study in the laboratory, analyzed the results, and drafted the manuscript. PL participated in genomic data analysis, critical reviewing and editing of original draft manuscript. RH provided guidance and participated in genomic data analysis, critical reviewing, and editing of the manuscript. TD participated in critical reviewing and editing of the manuscript. RM, JO, and AD conceived, designed, and contributed to the revision of the draft and final approval of the version to be published. AD was the principal supervisor of the project. All authors read and approved the final manuscript.

FUNDING

This study was supported by the Danish International Development Assistance (Danida)-funded project "Innovations and Markets for Lake Victoria Fisheries" (IMLAF), DFC File No. 14-P01-TAN as well as the International Foundation for Sciences (IFS), Stockholm, Sweden through grant 1-2-A-6100-1.

ACKNOWLEDGMENTS

The authors extend their gratitude to the funders and colleagues at the National Fish Quality Control Laboratory in Mwanza as well as colleagues from the Water Quality Laboratory of the Lake Zone in Mwanza, Tanzania.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.00901/full#supplementary-material

REFERENCES

Alam, M., Sultana, M., Nair, G. B., Sack, R. B., Sack, D. A., Siddique, A. K., et al. (2006). Toxigenic *Vibrio cholerae* in the Aquatic Environment of Mathbaria, Bangladesh. *Appl. Environ. Microbiol.* 72, 2849–2855. doi: 10.1128/AEM.72.4.2849-2855.2006

Alcalde-Rico, M., Hernando-Amado, S., Blanco, P., and Martínez, J. L. (2016). Multidrug efflux pumps at the crossroad between antibiotic resistance and bacterial virulence. *Front. Microbiol.* 7:1483. doi: 10.3389/fmicb.2016.01483

Aliabad, N. H., Bakhshi, B., Pourshafie, M. R., Sharifnia, A., and Ghorbani, M. (2012). Molecular diversity of CTX prophage in *Vibrio cholerae*: CTX prophage in *Vibrio cholerae*. *Lett. Appl. Microbiol.* 55, 27–32. doi: 10.1111/j.1472-765X.2012.03253.x

Alvarez-Ortega, C., Olivares, J., and Martinez, J. L. (2013). RND multidrug efflux pumps: what are they good for? *Front. Microbiol.* 4:7. doi: 10.3389/fmicb.2013.00007

Azarian, T., Ali, A., Johnson, J. A., Jubair, M., Cella, E., Ciccozzi, M., et al. (2016). Non-toxigenic environmental *Vibrio cholerae* O1 strain from Haiti provides evidence of pre-pandemic cholera in Hispaniola. *Sci. Rep.* 6:36115. doi: 10.1038/srep36115

Azarian, T., Ali, A., Johnson, J. A., Mohr, D., Prosperi, M., Veras, N. M., et al. (2014). Phylodynamic analysis of clinical and environmental *Vibrio cholerae*

isolates from haiti reveals diversification driven by positive selection. *mBio* 5:e1824-14. doi: 10.1128/mBio.01824-14

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bari, S. M. N., Roky, M. K., Mohiuddin, M., Kamruzzaman, M., Mekalanos, J. J., and Faruque, S. M. (2013). Quorum-sensing autoinducers resuscitate dormant *Vibrio cholerae* in environmental water samples. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9926–9931. doi: 10.1073/pnas.1307697110

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., et al. (2015). RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci. Rep.* 5:8365. doi: 10.1038/srep08365

Bwire, G., Debes, A. K., Orach, C. G., Kagirita, A., Ram, M., Komakech, H., et al. (2018a). Environmental surveillance of *Vibrio cholerae* O1/O139 in the five african great lakes and other major surface water sources in Uganda. *Front. Microbiol.* 9:1560. doi: 10.3389/fmicb.2018.01560

Bwire, G., Sack, D. A., Almeida, M., Li, S., Voeglein, J. B., Debes, A. K., et al. (2018b). Molecular characterization of *Vibrio cholerae* responsible for cholera epidemics in Uganda by PCR, MLVA and WGS. *PLoS Negl. Trop. Dis.* 12:e0006492. doi: 10.1371/journal.pntd.0006492

Carraro, N., Sauvé, M., Matteau, D., Lauzon, G., Rodrigue, S., and Burrus, V. (2014). Development of pVCR94ΔX from *Vibrio cholerae*, a prototype for studying multidrug resistant IncA/C conjugative plasmids. *Front. Microbiol.* 5:44. doi: 10.3389/fmicb.2014.00044

Castillo, D., Kauffman, K., Hussain, F., Kalatzis, P., Rørbo, N., Polz, M. F., et al. (2018). Widespread distribution of prophage-encoded virulence factors in marine Vibrio communities. *Sci. Rep.* 8:9973. doi: 10.1038/s41598-018-28326-9

Ceccarelli, D., Garriss, G., Choi, S. Y., Hasan, N. A., Stepanauskas, R., Pop, M., et al. (2017). Characterization of two cryptic plasmids isolated in haiti from clinical *Vibrio cholerae* Non-O1/Non-O139. *Front. Microbiol.* 8:2283. doi: 10.3389/fmicb.2017.02283

CLSI (2017). *Performance Standards for Antimicrobial Susceptibility Testing*, 27th Edn. Wayne, PA: Clinical and Laboratory Standards Institute.

Crump, J. A., Bopp, C. A., Greene, K. D., Kubota, K. A., Middendorf, R. L., Wells, J. G., et al. (2003). Toxigenic *Vibrio cholerae* Serogroup O141–associated cholera-like diarrhea and bloodstream infection in the United States. *J. Infect. Dis.* 187, 866–868. doi: 10.1086/368330

Dalsgaard, A., Serichantalergs, O., Forslund, A., Lin, W., Mekalanos, J., Mintz, E., et al. (2001). Clinical and environmental isolates of *Vibrio cholerae* serogroup O141 carry the CTX phage and the genes encoding the toxin-coregulated pili. *J. Clin. Microbiol.* 39, 4086–4092. doi: 10.1128/JCM.39.11.4086-4092.2001

Dalusi, L., Lyimo, T. J., Lugomela, C., Hosea, K. M. M., and Sjöling, S. (2015a). Toxigenic *Vibrio cholerae* identified in estuaries of Tanzania using PCR techniques. *FEMS Microbiol. Lett.* 362:fnv009. doi: 10.1093/femsle/fnv009

Dalusi, L., Saarenheimo, J., Lyimo, T. J., and Lugomela, C. (2015b). Genetic relationship between clinical and environmental *Vibrio cholerae* isolates in Tanzania: a comparison using repetitive extragenic palindromic (REP) and enterobacterial repetitive intergenic consensus (ERIC) fingerprinting approach. *Afr. J. Microbiol. Res.* 9, 455–462. doi: 10.5897/AJMR2014.7307

Dongre, M., Singh, B., Aung, K. M., Larsson, P., Miftakhova, R., Persson, K., et al. (2018). Flagella-mediated secretion of a novel *Vibrio cholerae* cytotoxin affecting both vertebrate and invertebrate hosts. *Commun. Biol.* 1:59. doi: 10.1038/s42003-018-0065-z

Echenberg, M. J. (2011). *Africa in the Time of Cholera: A History of Pandemics from 1817 to the Present*. New York, NY: Cambridge University Press.

Eiler, A., Gonzalez-Rey, C., Allen, S., and Bertilsson, S. (2007). Growth response of *Vibrio cholerae* and other *Vibrio* spp. to cyanobacterial dissolved organic matter and temperature in brackish water: Cyanobacterial DOM, temperature and Vibrio growth. *FEMS Microbiol. Ecol.* 60, 411–418. doi: 10.1111/j.1574-6941.2007.00303.x

Faruque, S. M., Albert, M. J., and Mekalanos, J. J. (1998). Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiol. Mol. Biol. Rev.* 62, 1301–1314.

Faruque, S. M., and Mekalanos, J. J. (2012). Phage-bacterial interactions in the evolution of toxigenic *Vibrio cholerae*. *Virulence* 3, 556–565. doi: 10.4161/viru.22351

Faruque, S. M., Tam, V. C., Chowdhury, N., Diraphat, P., Dziejman, M., Heidelberg, J. F., et al. (2007). Genomic analysis of the Mozambique strain of *Vibrio cholerae* O1 reveals the origin of El Tor strains carrying classical CTX prophage. *Proc. Natl. Acad. Sci. U.S.A.* 104, 5151–5156. doi: 10.1073/pnas.0700365104

Folster, J. P., Katz, L., McCullough, A., Parsons, M. B., Knipe, K., Sammons, S. A., et al. (2014). Multidrug-resistant IncA/C plasmid in *Vibrio cholerae* from Haiti. *Infectious Dis. J.* 20, 1951–1953. doi: 10.3201/eid2011.140889

Fong, J. C. N., Syed, K. A., Klose, K. E., and Yildiz, F. H. (2010). Role of Vibrio polysaccharide (vps) genes in VPS production, biofilm formation and *Vibrio cholerae* pathogenesis. *Microbiology* 156, 2757–2769. doi: 10.1099/mic.0.040196-0

Garrine, M., Mandomando, I., Vubil, D., Nhampossa, T., Acacio, S., Li, S., et al. (2017). Minimal genetic change in *Vibrio cholerae* in Mozambique over time: multilocus variable number tandem repeat analysis and whole genome sequencing. *PLoS Negl. Trop. Dis.* 11:e0005671. doi: 10.1371/journal.pntd.0005671

Ghosh-Banerjee, J., Senoh, M., Takahashi, T., Hamabata, T., Barman, S., Koley, H., et al. (2010). Cholera toxin production by the El Tor variant of *Vibrio cholerae* O1 compared to prototype El Tor and classical biotypes. *J. Clin. Microbiol.* 48, 4283–4286. doi: 10.1128/JCM.00799-10

Haley, B. J., Choi, S. Y., Grim, C. J., Onifade, T. J., Cinar, H. N., Tall, B. D., et al. (2014). Genomic and phenotypic characterization of *Vibrio cholerae* Non-O1 isolates from a US gulf coast cholera outbreak. *PLoS One* 9:e86264. doi: 10.1371/journal.pone.0086264

Halpern, M., and Izhaki, I. (2017). Fish as hosts of *Vibrio cholerae*. *Front. Microbiol.* 8:282. doi: 10.3389/fmicb.2017.00282

Hammerl, J. A., Jäckel, C., Bortolaia, V., Schwartz, K., Bier, N., Hendriksen, R. S., et al. (2017). Carbapenemase VCC-1–producing *Vibrio cholerae* in coastal waters of Germany. *Emerg. Infect. Dis.* 23, 1735–1737. doi: 10.3201/eid2310.161625

Haque, M. M., Alam, M., and Salam, A. (2012). Frequency of *Vibrio cholerae* in the water and plankton samples of south- western coastal aquatic habitats of Bangladesh. *J. Bangladesh Acad. Sci.* 36, 71–78. doi: 10.3329/jbas.v36i1.10922

Hendriksen, R. S., Price, L. B., Schupp, J. M., Gillece, J. D., Kaas, R. S., Engelthaler, D. M., et al. (2011). Population genetics of *Vibrio cholerae* from Nepal in 2010: evidence on the origin of the haitian outbreak. *mBio* 2:e00157-11. doi: 10.1128/mBio.00157-11

Hossain, Z. Z., Farhana, I., Tulsiani, S. M., Begum, A., and Jensen, P. K. M. (2018a). Transmission and toxigenic potential of *Vibrio cholerae* in hilsha fish (Tenualosa ilisha) for human consumption in Bangladesh. *Front. Microbiol.* 9:222. doi: 10.3389/fmicb.2018.00222

Hossain, Z. Z., Leekitcharoenphon, P., Dalsgaard, A., Sultana, R., Begum, A., Jensen, P. K. M., et al. (2018b). Comparative genomics of *Vibrio cholerae* O1 isolated from cholera patients in Bangladesh. *Lett. Appl. Microbiol.* 67, 329–336. doi: 10.1111/lam.13046

Hounmanou, Y. M. G., Mdegela, R. H., Dougnon, T. V., Mhongole, O. J., Mayila, E. S., Malakalinga, J., et al. (2016). Toxigenic *Vibrio cholerae* O1 in vegetables and fish raised in wastewater irrigated fields and stabilization ponds during a non-cholera outbreak period in Morogoro, Tanzania: an environmental health study. *BMC Res. Notes* 9:466. doi: 10.1186/s13104-016-2283-0

Islam, M. S., Islam, M. S., Mahmud, Z. H., Cairncross, S., Clemens, J. D., and Collins, A. E. (2015). Role of phytoplankton in maintaining endemicity and seasonality of cholera in Bangladesh. *Trans. R. Soc. Trop. Med. Hyg.* 109, 572–578. doi: 10.1093/trstmh/trv057

Joelsson, A., Kan, B., and Zhu, J. (2007). Quorum sensing enhances the stress response in *Vibrio cholerae*. *Appl. Environ. Microbiol.* 73, 3742–3746. doi: 10.1128/AEM.02804-06

Jolley, K. A., Bliss, C. M., Bennett, J. S., Bratcher, H. B., Brehony, C., Colles, F. M., et al. (2012). Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology* 158, 1005–1015. doi: 10.1099/mic.0.055459-0

Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M., and Lund, O. (2014). Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One* 9:e104984. doi: 10.1371/journal.pone.0104984

Kaas, R. S., Ngandjio, A., Nzouankeu, A., Siriphap, A., Fonkoua, M.-C., Aarestrup, F. M., et al. (2016). The lake chad basin, an isolated and persistent reservoir of *Vibrio cholerae* O1: a genomic insight into the outbreak in cameroon, 2010. *PLoS One* 11:e0155691. doi: 10.1371/journal.pone.0155691

Kachwamba, Y., Mohammed, A. A., Lukupulo, H., Urio, L., Majigo, M., Mosha, F., et al. (2017). Genetic Characterization of *Vibrio cholerae* O1 isolates from outbreaks between 2011 and 2015 in Tanzania. *BMC Infect. Dis.* 17:157. doi: 10.1186/s12879-017-2252-9

Kamruzzaman, M., Udden, S. M. N., Cameron, D. E., Calderwood, S. B., Nair, G. B., Mekalanos, J. J., et al. (2010). Quorum-regulated biofilms enhance the development of conditionally viable, environmental *Vibrio cholerae*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1588–1593. doi: 10.1073/pnas.0913404107

Kim, E. J., Lee, D., Moon, S. H., Lee, C. H., Kim, S. J., Lee, J. H., et al. (2014). Molecular insights into the evolutionary pathway of *Vibrio cholerae* O1 atypical El Tor variants. *PLoS Pathog.* 10:e1004384. doi: 10.1371/journal.ppat.1004384

Kishe, M. A., and Machiwa, J. F. (2003). Distribution of heavy metals in sediments of Mwanza Gulf of Lake Victoria, Tanzania. *Environ. Int.* 28, 619–625. doi: 10.1016/s0160-4120(02)00099-5

Kudirkiene, E., Andoh, L. A., Ahmed, S., Herrero-Fresno, A., Dalsgaard, A., Obiri-Danso, K., et al. (2018). The use of a combined bioinformatics approach to locate antibiotic resistance genes on plasmids from whole genome sequences of *Salmonella enterica* serovars from humans in Ghana. *Front. Microbiol.* 9:1010. doi: 10.3389/fmicb.2018.01010

Larsen, M. V., Cosentino, S., Rasmussen, S., Friis, C., Hasman, H., Marvig, R. L., et al. (2012). Multilocus sequence typing of total-genome-sequenced bacteria. *J. Clin. Microbiol.* 50, 1355–1361. doi: 10.1128/JCM.06094-11

Logan, S. L., Thomas, J., Yan, J., Baker, R. P., Shields, D. S., Xavier, J. B., et al. (2018). The *Vibrio cholerae* type VI secretion system can modulate host intestinal mechanics to displace gut bacterial symbionts. *Proc. Natl. Acad. Sci. U.S.A.* 115, E3779–E3787. doi: 10.1073/pnas.1720133115

Lugomela, C., Moyo, S., Lyimo, T. J., Namkinga, L. A., Goericke, R., and Sjöling, S. (2014). Co-variations of cholera with climatic and environmental parameters in coastal regions of Tanzania. *West. Indian Ocean J. Mar. Sci.* 13, 93–105.

Lutz, C., Erken, M., Noorian, P., Sun, S., and McDougald, D. (2013). Environmental reservoirs and mechanisms of persistence of *Vibrio cholerae*. *Front. Microbiol.* 4:375. doi: 10.3389/fmicb.2013.00375

Mangat, C. S., Boyd, D., Janecko, N., Martz, S.-L., Desruisseau, A., Carpenter, M., et al. (2016). Characterization of VCC-1, a novel ambler class A carbapenemase from *Vibrio cholerae* isolated from imported retail shrimp sold in Canada. *Antimicrob. Agents Chemother.* 60, 1819–1825. doi: 10.1128/aac.02812-15

Meibom, K. L., Li, X. B., Nielsen, A. T., Wu, C.-Y., Roseman, S., and Schoolnik, G. K. (2004). The *Vibrio cholerae* chitin utilization program. *PNAS* 101, 2524–2529. doi: 10.1073/pnas.0308707101

Mengel, M. A., Delrieu, I., Heyerdahl, L., and Gessner, B. D. (2014). "Cholera outbreaks in Africa," in *Cholera Outbreaks*, eds G. B. Nair and Y. Takeda (Berlin: Springer), 117–144. doi: 10.1007/82_2014_369

Mitchell, K. C., Breen, P., Britton, S., Neely, M. N., and Withey, J. H. (2017). Quantifying *Vibrio cholerae* enterotoxicity in a zebrafish infection model. *Appl. Environ. Microbiol.* 83:e783-17. doi: 10.1128/AEM.00783-17

Mutreja, A., Kim, D. W., Thomson, N., Connor, T. R., Lee, J. H., Kariuki, S., et al. (2011). Evidence for multiple waves of global transmission within the seventh cholera pandemic. *Nature* 477, 462–465. doi: 10.1038/nature10392

Naha, A., Pazhani, G. P., Ganguly, M., Ghosh, S., Ramamurthy, T., Nandy, R. K., et al. (2012). Development and evaluation of a PCR assay for tracking the emergence and dissemination of haitian variant ctxB in *Vibrio cholerae* O1 strains isolated from Kolkata. *India J. Clin. Microbiol.* 50, 1733–1736. doi: 10.1128/JCM.00387-12

Nandi, B., Nandy, R. K., Mukhopadhyay, S., Nair, G. B., Shimada, T., and Ghose, A. C. (2000). Rapid method for species-specific identification of *Vibrio cholerae* using primers targeted to the gene of outer membrane protein OmpW. *J. Clin. Microbiol.* 38, 4145–4151.

Nkoko, D., Giraudoux, P., Plisnier, P.-D., Tinda, A., Piarroux, M., Sudre, B., et al. (2011). Dynamics of cholera outbreaks in great lakes region of Africa, 1978–2008. *Emerg. Infect. Dis.* 17, 2026–2034. doi: 10.3201/eid1711.110170

Nyambuli, S., Mhongole, O. J., Katakweba, A. A., Dalsgaard, A., and Mdegela, R. H. (2018). Prevalence, pathogenic markers and antibiotic susceptibility of *Vibrio cholerae* in sardines, water and phytoplankton in lake Tanganyika, Tanzania. *Int. J. Agric. For. Fish.* 6:29.

Ogwok, P., Muyonga, J. H., and Sserunjogi, M. L. (2009). Pesticide residues and heavy metals in lake victoria nile perch, lates niloticus, belly flap oil. *Bull. Environ. Contam. Toxicol.* 82, 529–533. doi: 10.1007/s00128-009-9668-x

Onyuka, J., Kakai, R., Onyango, D., Arama, P. F., Gichuki, J., and Ofulla, A. V. O. (2011). Prevalence and antimicrobial susceptibility patterns of enteric bacteria isolated from water and fish in lake victoria basin of Western Kenya. *Int. J. Biol. Med. Sci.* 1, 6–13.

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., et al. (2014). The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). *Nucleic Acids Res.* 42, D206–D214. doi: 10.1093/nar/gkt1226

Plisnier, P.-D., Poncelet, N., Cocquyt, C., De Boeck, H., Bompangue, D., Naithani, J., et al. (2015). *Cholera Outbreaks at Lake Tanganyika Induced by Climate Change?*. Brussels: Belspo.

Rabia, A., Wambura, P., Misinzo, G., Kimera, S., Mdegela, R., Mzula, A., et al. (2017). Molecular epidemiology of *Vibrio cholerae* recovered from sewage drains, captured fish and humans in 2015/16 cholera outbreak in zanzibar, Tanzania. *J. Adv. Microbiol.* 5, 1–11. doi: 10.9734/JAMB/2017/36036

Rashid, M., Rashed, S. M., Islam, T., Johura, F.-T., Watanabe, H., Ohnishi, M., et al. (2016). CtxB1 outcompetes CtxB7 in *Vibrio cholerae* O1, Bangladesh. *J. Med. Microbiol.* 65, 101–103. doi: 10.1099/jmm.0.000190

Reimer, A., Domselaar, G., Stroika, S., Walker, M., Kent, H., Tarr, C., et al. (2011). Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. *Emerg. Infect. Dis.* 17, 2113–2121. doi: 10.3201/eid1711.110794

Reyburn, R., Kim, D. R., Emch, M., Khatib, A., von Seidlein, L., and Ali, M. (2011). Climate variability and the outbreaks of cholera in Zanzibar, East Africa: a time series analysis. *Am. J. Trop. Med. Hyg.* 84, 862–869. doi: 10.4269/ajtmh.2011.10-0277

Runft, D. L., Mitchell, K. C., Abuaita, B. H., Allen, J. P., Bajer, S., Ginsburg, K., et al. (2014). Zebrafish as a natural host model for *Vibrio cholerae* colonization and transmission. *Appl. Environ. Microbiol.* 80, 1710–1717. doi: 10.1128/AEM.03580-13

Seed, K. D., Lazinski, D. W., Calderwood, S. B., and Camilli, A. (2013). A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature* 494, 489–491. doi: 10.1038/nature11927

Sinha-Ray, S., and Ali, A. (2017). Mutation in flrA and mshA Genes of *Vibrio cholerae* inversely involved in vps-independent biofilm driving bacterium toward nutrients in lake water. *Front. Microbiol.* 8:1770. doi: 10.3389/fmicb.2017.01770

Siriphap, A., Leekitcharoenphon, P., Kaas, R. S., Theethakaew, C., Aarestrup, F. M., Sutheinkul, O., et al. (2017). Characterization and genetic variation of *Vibrio cholerae* isolated from clinical and environmental sources in Thailand. *PLoS One* 12:e0169324. doi: 10.1371/journal.pone.0169324

Spagnoletti, M., Ceccarelli, D., Rieux, A., Fondi, M., Taviani, E., Fani, R., et al. (2014). Acquisition and evolution of SXT-R391 integrative conjugative elements in the seventh-pandemic *Vibrio cholerae* lineage. *mBio* 5, e1356–e1314. doi: 10.1128/mBio.01356-14

Thomsen, M. C. F., Ahrenfeldt, J., Cisneros, J. L. B., Jurtz, V., Larsen, M. V., Hasman, H., et al. (2016). A bacterial analysis platform: an integrated system for analysing bacterial whole genome sequencing data for clinical diagnostics and surveillance. *PLoS One* 11:e0157718. doi: 10.1371/journal.pone.0157718

Unterweger, D., Miyata, S. T., Bachmann, V., Brooks, T. M., Mullins, T., Kostiuk, B., et al. (2014). The *Vibrio cholerae* type VI secretion system employs diverse effector modules for intraspecific competition. *Nat. Commun.* 5:3549. doi: 10.1038/ncomms4549

Urassa, W., Mhando, Y., Mhalu, F., and Mgonja, S. (2009). Antimicrobial susceptibility pattern of *Vibrio cholerae* 01 strains during two cholera outbreaks in Dar Es Salaam, Tanzania. *East Afr. Med. J.* 77, 350–353. doi: 10.4314/eamj.v77i7.46661

Verlecar, X. N., and Desai, S. R. (2004). *Phytoplankton Identification Manual*. Dona Paula: National Institute of Oceanography.

Vesth, T., Wassenaar, T. M., Hallin, P. F., Snipen, L., Lagesen, K., and Ussery, D. W. (2010). On the origins of a vibrio species. *Microb. Ecol.* 59, 1–13. doi: 10.1007/s00248-009-9596-7

Waldor, M. K., and Mekalanos, J. J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* 272, 1910–1914. doi: 10.1126/science.272.5270.1910

Wang, R., Liu, H., Zhao, X., Li, J., and Wan, K. (2018). IncA/C plasmids conferring high azithromycin resistance in *vibrio cholerae*. *Int. J. Antimicrob. Agents* 51, 140–144. doi: 10.1016/j.ijantimicag.2017.09.009

Wang, R., Yu, D., Yue, J., and Kan, B. (2016). Variations in SXT elements in epidemic *Vibrio cholerae* O1 El Tor strains in China. *Sci. Rep.* 6:22733. doi: 10.1038/srep22733

Weill, F.-X., Domman, D., Njamkepo, E., Tarr, C., Rauzier, J., Fawal, N., et al. (2017). Genomic history of the seventh pandemic of cholera in Africa. *Science* 358, 785–789. doi: 10.1126/science.aad5901

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Manuscript C: "Genomic insights into *Vibrio cholerae* O1 responsible for cholera epidemics in Tanzania from 1993 to 2017"

Article accepted for publication in ***PLoS Neglected Tropical Diseases*** 13(12): e0007934.

https://doi.org/10.1371/journal.pntd.0007934

Yaovi Mahuton Gildas Hounmanou[1*], Pimlapas Leekitcharoenphon[2], Egle Kudirkiene[1], Robinson H. Mdegela[3], Rene S. Hendriksen[2], John Elmerdahl Olsen[1] and Anders Dalsgaard[1]

*** Corresponding author:**

Yaovi M. Gildas Hounmanou

University of Copenhagen

gil@sund.ku.dk

Short title: Evolutionary genomics of *V. cholerae* O1 in Tanzania

Keywords: *Vibrio cholerae*, Genomics, Pangenome, Tanzania, Cholera, African Great Lakes Region

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Genomic insights into *Vibrio cholerae* O1 responsible for cholera epidemics in Tanzania between 1993 and 2017

Yaovi Mahuton Gildas Hounmanou[1]*, Pimlapas Leekitcharoenphon[2], Egle Kudirkiene[1], Robinson H. Mdegela[3], Rene S. Hendriksen[2], John Elmerdahl Olsen[1], Anders Dalsgaard[1,4]

1 Department of Veterinary and Animal Sciences, University of Copenhagen, Copenhagen, Denmark, **2** National Food Institute, Technical University of Denmark, Lyngby, Denmark, **3** Department of Veterinary Medicine and Public Health, Sokoine University of Agriculture, Morogoro, Tanzania, **4** School of Chemical and Biomedical Engineering, Nanyang Technological University, Singapore city, Singapore

* gil@sund.ku.dk

Citation: Hounmanou YMG, Leekitcharoenphon P, Kudirkiene E, Mdegela RH, Hendriksen RS, Olsen JE, et al. (2019) Genomic insights into *Vibrio cholerae* O1 responsible for cholera epidemics in Tanzania between 1993 and 2017. PLoS Negl Trop Dis 13(12): e0007934. https://doi.org/10.1371/journal.pntd.0007934

Editor: François-Xavier Weill, Institut Pasteur, FRANCE

Received: May 27, 2019

Accepted: November 18, 2019

Published: December 23, 2019

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files. However, Raw sequences reads generated in this study are deposited to the European Nucleotide Archive under the project accession number PRJEB30604.

Funding: AD and RM received funding from the Danish International Development Assistance (Danida) through the project "Innovations and Markets for Lake Victoria Fisheries" (IMLAF). Grant

Abstract

Background

Tanzania is one of seven countries with the highest disease burden caused by cholera in Africa. We studied the evolution of *Vibrio cholerae* O1 isolated in Tanzania during the past three decades.

Methodology/Principal findings

Genome-wide analysis was performed to characterize *V. cholerae* O1 responsible for the Tanzanian 2015–2017 outbreak along with strains causing outbreaks in the country for the past three decades. The genomes were further analyzed in a global context of 590 strains of the seventh cholera pandemic (7PET), as well as environmental isolates from Lake Victoria. All Tanzanian cholera outbreaks were caused by the 7PET lineage. The T5 sub-lineage (*ctx*B3) dominated outbreaks until 1997, followed by the T10 atypical El Tor (*ctx*B1) up to 2015, which were replaced by the T13 atypical El Tor of the current third wave (*ctx*B7) causing most cholera outbreaks until 2017 with T13 being phylogenetically related to strains from East African countries, Yemen and Lake Victoria. The strains were less drug resistant with approximate 10-kb deletions found in the SXT element, which encodes resistance to sulfamethoxazole and trimethoprim. Nucleotide deletions were observed in the CTX prophage of some strains, which warrants further virulence studies. Outbreak strains share 90% of core genes with *V. cholerae* O1 from Lake Victoria with as low as three SNPs difference and a significantly similar accessory genome, composed of genomic islands namely the CTX prophage, Vibrio Pathogenicity Islands; toxin co-regulated pilus biosynthesis proteins and the SXT-ICE element.

Conclusion/Significance

Characterization of *V. cholerae* O1 from Tanzania reveals genetic diversity of the 7PET lineage composed of T5, T10 and T13 sub-lineages with introductions of new sequence types

number: DFC File No. 14-P01-TAN, (https://dfcentre.com/). YMGH received funding from the International Foundation for Sciences. Grant number: 1-2-A-6100-1, (http://www.ifs.se/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript

Competing interests: The authors have declared that no competing interests exist.

from neighboring countries. The presence of these sub-lineages in environmental isolates suggests that the African Great Lakes may serve as aquatic reservoirs for survival of *V. cholerae* O1 favoring continuous human exposure.

Author summary

The seventh cholera pandemic has claimed >250,000 reported cases and 13,078 deaths until 2018 in Tanzania. To understand the epidemiology and to guide control, we used genomics to study *V. cholerae* O1 isolated in Tanzania during the past three decades. Tanzanian cholera outbreaks were caused by the T5, T10 and T13 sub-lineages of the 7PET lineage of *V. cholerae* O1 with some strains showing an unusual 100-bp deletion on the CTX prophage. From 1993 to 2017, most sub-lineages found in patients were also found in the aquatic environment and the close phylogenetic relationships between strains from the two niches suggest that the African Great Lakes may act as a reservoir for cholera outbreak strains. Moreover, we reported clonal transmission at regional and global scale favored by population displacements. Regional collaborative efforts are advised for effective cholera control.

Introduction

In 1974, cholera reached Tanzania on the shores of Lake Nyasa bordering Malawi [1], and has since caused recurrent outbreaks of varying magnitudes almost every year resulting in over 250,000 reported cases and 13,078 deaths until 2018 [2,3]. In Africa, the different epidemics could all be traced back to a single lineage from South Asia, which has been introduced at least 11 times since the first epidemic in the 1970s [4]. The ongoing seventh cholera pandemic is characterized by multiple waves of *V. cholerae* O1 strains associated with various genotypic markers mainly variations in the *ctx*B gene on the CTX prophage [4,5]. To understand the evolution of *V. cholerae* O1 requires genome-wide analyses at national and regional scales [6].

Previous analysis of *V. cholerae* O1 from the 2015 cholera outbreak in Tanzania revealed that strains involved in initial outbreaks around refugee camps formed two distinct genetic lineages both different from other strains associated with the countrywide outbreak occurring later in the same year [7]. This indicates the occurrence of heterogeneous *V. cholerae* O1 through introductions of different sub-lineages into the country at different time points. Studies have also indicated aquatic environments as a potential source for cholera outbreak strains in Tanzania [8,9].

Here, we analyze 22 *V. cholerae* O1 from the 2015–2017 cholera outbreak in Tanzania in a national and global context along with strains recovered from Lake Victoria aiming to investigate their evolution, including determinants of pathogenicity and antimicrobial resistance. Lessons learnt from these past outbreak strains provide evidence of cross-border spread of *V. cholerae* O1 in the East African region and call for integrated collaborations of the different concerned health authorities to proactively establish joint control strategies to circumvent future cholera epidemics in the region.

Material and methods

Study area and strains collection

The United Republic of Tanzania is an East African country and part of the African Great Lakes Region [10]. We studied clinical *V. cholerae* O1 strains and publicly available genomes of *V. cholerae* O1 from eleven regions of mainland Tanzania and Zanzibar originating between

1993 and 2017 (Fig 1, S1 Table). *V. cholerae* O1 isolated between 2015 and 2017 from cholera patients in Ruvuma, Songwe, Dar es Salaam, Morogoro, Mwanza, Mbeya, Kigoma and Tanga were obtained from the National Health Laboratory Quality Assurance and Training Centre of the Ministry of Health in Dar es Salaam (Fig 1). *V. cholerae* O1 isolated during the 2016–2017 cholera outbreak from Zanzibar were obtained from Mnazi Mmoja Hospital of the Ministry of Health and Social Affairs. Overall, two strains per region from mainland Tanzania and six strains from Zanzibar resulting in 22 strains in total were confirmed as *V. cholerae* O1 and subjected to antimicrobial susceptibility testing as previously described [9], and whole genome sequencing (WGS). Public genomes of clinical *V. cholerae* O1 isolated between 1993 and 2015 (n = 23) [4,7] and recent environmental *V. cholerae* strains from Lake Victoria, Tanzania (n = 9) [9] were obtained from the Genbank and the European Nucleotide Archives (ENA) and included in the phylogenetic analyses (S1 Table).

DNA extraction, whole genome sequencing and genome assembly

DNA from the 22 *V. cholerae* O1 isolates was extracted using the automated Maxwell DNA extraction machine (Promega Maxwell RSC, Wisconsin, USA) and sequencing was performed on a Miseq (Illumina, Inc., San Diego, CA, USA) as previously described [9] at the University of Copenhagen, Denmark. Raw sequences were submitted to ENA (Accession number PRJEB30604). Reads were assembled using SPAdes v. 3.9 [11] and assemblies were annotated using Prokka (v. 1.12-beta) with default settings, using barrnap 0.7 for rRNA prediction [12].

Characterization of *V. cholerae* O1 from Tanzania

Sequenced strains were analyzed using the online tools from the CGE platform (https://cge.cbs.dtu.dk/services/cge/) with default settings as previously described [9]. This included identification of *V. cholerae* serogroup-specific genes (*rfbV*-O1, *wbfZ*-O139), biotype-specific genes (*ctx*B, *rst*R, *tcp*A), major virulence genes, and VC2346 specific for the seventh cholera pandemic [13,14]. Detection of genomic islands of *V. cholerae* VPI-1, VPI-2, VSP-1, VSP-2 and the Type VI secretion system (T6SS) proteins was carried out using MyDbFinder 1.2. Furthermore, MyDbFinder 1.2 [14] coupled with nucleotides BLAST served for genotyping of the strains based on the *ctx*B of the CTX prophage that they carried. The *ctx*B of *V. cholerae* N16961 (AE003852) served as reference for *ctx*B3 to search for prototype El Tor strains. The

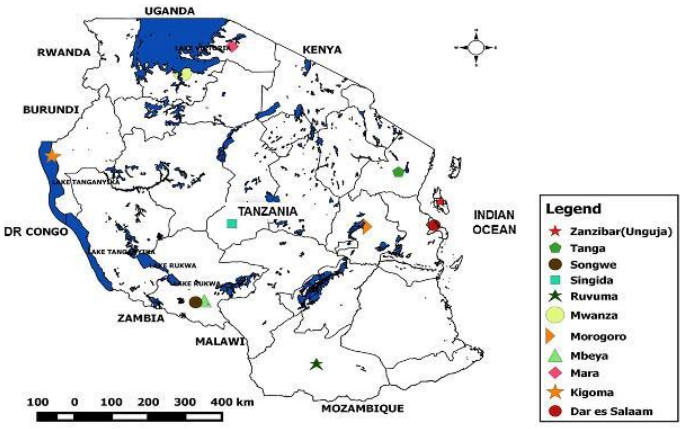


Fig 1. Sampling area. The *V. cholerae* O1 strains analyzed originated from regions listed in the legend box of the map. Map constructed with QGIS version 2.12.3 (https://www.qgis.org) using the GPS coordinates recorded from our sampling sites and Tanzanian country shape files obtained from DIVA-GIS (http://www.diva-gis.org/gdata).

https://doi.org/10.1371/journal.pntd.0007934.g001

*ctx*B1 of *V. cholerae* O395 (CP001235) was used to identify altered El Tor strains of the early third wave of the seventh pandemic, whereas the point mutation (C to A) at position 58 in the *ctx*B1 making it *ctx*B7 [15] served to identify strains of genotype *ctx*B7 of to the current third wave of the seventh pandemic. ResFinder 3.1 [16] with default options assessed acquired antimicrobial resistance (AMR) genes. MyDbFinder 1.2 [17] was used to detect the SXT integrative conjugative element, class 1 integrons, and the presence of mutations in the DNA gyrase (*gyr*A gene) and in the DNA topoisomerase IV (*par*C gene) [14]. Search for plasmids was conducted using PlasmidFinder 1.3, MyDbFinder 1.2 tools, with cryptic plasmid replicons [9] and Blast atlas using GView (https://server.gview.ca/) to assess occurrence of plasmid replicons in the sequences. *In-silico* MLST was performed [13] based on internal fragments of the seven housekeeping genes: *adk*, *gyrB*,*metE*, *mdh*, *pntA*, *purM*, and *pyrC* using MLST 2.0 [17]. The included public available genomes have previously been reported [4,7,9] and were included for comparative analysis. The sequence types of the already published genomes were originally not reported [4,7], but we determined these using MLST 2.0 [17]. We localized resistance genes on plasmids from the public available genomes containing the IncA/C2 plasmid [4] using Blast Atlas in GView (https://server.gview.ca/). Likewise, we analyzed clinical *V. cholerae* O1 from previous studies [4,7] for deletions on the CTX prophage and the SXT conjugative elements by mapping the reads against the reference *V. cholerae* 2010EL-1786. We searched antimicrobial resistance genes and did *ctx*B genotyping and analysis of all major virulence genes as described above in the clinical *V. cholerae* O1 strains reported by Kachwamba et al [7], as they did not report such characteristics. The environmental strains were characterized and reported in a previous study [9] but were used in the present study for pangenomic comparison with the 2015–2017 outbreak strains for in-depth genomic analyses and for the overall phylogenetic evolution of Tanzanian *V. cholerae* since 1993 through 2017.

Phylogenetic and pan-genome analyses

The phylogenetic relationship between *V. cholerae* O1 that caused different outbreaks in Tanzania from 1993 to 2017 was assessed along with strains recovered from the environment using raw reads and trimmed assemblies in CSIPhylogeny version 1.4 with default options for a local single nucleotide polymorphism (SNP) analysis [18]. All Tanzanian strains were then placed in a global phylogenetic context of 590 genomes of the seventh cholera pandemic to identify the global genetic relatedness and diversity of the Tanzanian strains. The pre-seventh pandemic *V. cholerae* O1 strain M66-2 was used to root the trees. The newick files obtained in CSIPhilogeny 1.4 were annotated and visualized in iTOL [19].

We conducted a pangenome analysis for a genome-wide comparison between selected *V. cholerae* strains obtained from Lake Victoria (n = 9), Tanzania [9] and the clinical strains that caused cholera in 2015 to 2017 (n = 22). Annotated .gff files were used as an input to Roary (v. 3.7.0) pangenome analysis tool [20]. The binary presence/absence data of accessory genes produced in Roary was used to calculate the associations between all genes in the accessory genome and the selected traits of the isolates by employing the Scoary (v. 1.6.11) tool [21]. The accessory genome tree was visualized in phandango [22].

Results and discussion

Genomic characteristics, local phylogeny and pan-genome analysis of Tanzanian *V. cholerae* O1

V. cholerae associated with cholera in Tanzania, belong to serogroup O1, as they possess the *rfb*-O1 gene (Table 1, S1 Table). All Tanzanian strains, including isolates from Lake Victoria

Table 1. Genome characteristics of *V. cholerae* O1 isolated in Tanzania from 1993 to 2017.

Isolation year	Biotype	MLST	*ctx*B (Wave)	Resistance (SXT/R391)	T sub-linages	References
1993–1997	Prototype El Tor	ST69	*ctx*B3 (wave 1)	ICEVchHai1_*del (floR,strA/B,sul2) + 400bp gap in floR	T5	[4]
1998–2009	Atypical El Tor	ST69	*ctx*B1 (early wave 3)	ICEVchHai1	T10	[4,5]
2011–2012	Atypical El Tor	ST69	*ctx*B1 (early wave3)	ICEVchHai1_del(floR,strA/B,sul2) + 400bp gap in floR	T10	[7]
2015 (Kigoma January)	Atypical El Tor	ST515	*ctx*B1 (early wave 3)	ICEVchHai1	T10	[7]
2015–2017	Atypical El Tor	ST69	*ctx*B7 (current wave 3)	ICEVchHai1_del(floR,strA/B,sul2)	T13	This study
2017 (P2, P3, F2, F4, W1, W3)	Atypical El Tor	ST69	*ctx*B7 (current wave 3)	ICEVchHai1_del(floR,strA/B,sul2)	T13	[9]
2017 (F1, F3, W2) + Kigoma May 2015	Atypical El Tor	ST69	*ctx*B1 (early wave 3)	ICEVchHai1_del(floR,strA/B,sul2)	T10	[7,9]

*del: deletions in specified genes.

https://doi.org/10.1371/journal.pntd.0007934.t001

are of the seventh pandemic lineage 7PET, possessed the seventh pandemic-specific gene (VC2346) and differ from the reference seventh pandemic El Tor strain N16961 with a maximum of 160 SNPs (S2 Table, sheet 1). In agreement with previous reports, we confirmed that strains from 1993 through 1997 were all the prototype El Tor biotype (*ctxB3*) *V. cholerae* of sub-lineage T5 [4]. Cholera outbreaks occurring from 1998 until 2017 were caused by strains of the atypical El Tor biotype carrying either *ctx*B1 or *ctx*B7, while having *rst*R and *tcp*A of the typical El Tor biotype (Fig 2, Table 1). This coincides with the period of emergence of the hybrid biotype conferred by *ctx*B1 genes and associated with cholera outbreaks, which has since replaced the typical El Tor biotype in recent outbreaks [4,5,23]. These hybrid strains are

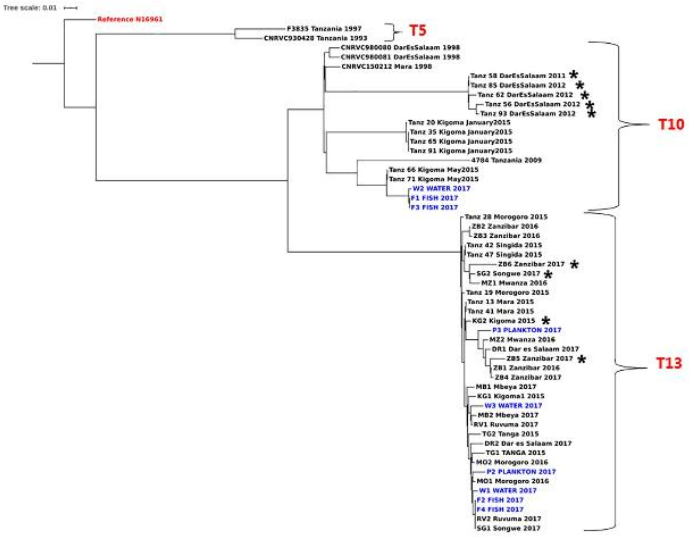


Fig 2. Maximum likelihood tree of *V. cholerae* O1 isolated in Tanzania from 1993 to 2017 along with strains from Lake Victoria (in blue). The reference strain *V. cholerae* N16961 was used to root the tree. Strains with the 100bp deletion in *ctx*A are marked with a star (*) and the T sub-lineages [4] of each phylogenetic cluster are indicated in brackets.

https://doi.org/10.1371/journal.pntd.0007934.g002

known for their ability to produce more cholera toxin than the prototype El Tor biotype strains causing a more severe diarrhea [24]. The 1993 and 1997 strains belong to the first wave of the seventh cholera pandemic and the T5 sub-lineage of 7PET. Clinical strains from 1998 up to 2012 and strains F1, F3 and W2 isolated from Lake Victoria in 2017 belong to the early part of wave III (*ctx*B1) of the seventh pandemic and part of the African T10 cluster, confirming previous reports [4]. The Kigoma refugee camos outbreak of May 2015 [25] also belong to this cluster.

V. cholerae O1 strains from 2013 were not available and there was no cholera reported in Tanzania in 2014 [2]. Strains isolated in outbreaks occurring after 2014 except for the those responsible for the Kigoma outbreaks in Janurary and May 2015, contain *ctx*B7 of the current third wave within the seventh cholera pandemic and belong to the T13 sub-lineage. Compared to T5 and T10 strains that occurred in the previous years, this shows significant genomic diversity of *V. cholerae* responsible for outbreaks in Tanzania overtime in line with the variation previously reported across the continent [4,5]. T13 strains are responsible for the ongoing cholera outbreak in Eastern Africa and Yemen [26,27]. Strains of the T13 sub-lineage formed a separate cluster on the local phylogenetic tree (Fig 2) and seem to occur in Tanzania after 2014, a time that corresponds with the global emergence of this sub-lineage [26]. Our clinical samples isolated between 2015 and 2017 are most closely related to *V. cholerae* O1 isolated in Lake Victoria in 2017 with as low as three SNPs difference and the environmental isolates also containing *ctx*B7 and being part of the T13 sub-lineage (Fig 2). This confirms our previous findings [9] and suggests a connection between environmental and outbreak strains where the isolates from the Lake could be either outbreak strains released into the environment through fecal contamination, e.g. sewage or they could be the source of the outbreak suggesting an environmental reservoir of *V. cholerae* O1 as described in Thailand, Cameroon and previously in Tanzania [8,13,14]. Isolates F1, F3 and W2 isolated in 2017 from Lake Victoria were revealed to belong to the sub-lineages T10 and are genetically related to pandemic strains circulating in the country since 1998 until 2015. This suggests an environmental survival of the strains even when outbreaks have ceased in people, favoring resurgence of epidemics overtime, with Lake Victoria serving as a reservoir as is also the case for Lake Chad [9,13]. Of the twenty-two 2015–2017 strains sequenced in this study, none was T10so their presence in the lake could not be directly linked to the discharge of urban sewage emanating from the ongoing outbreaks and the environment could remain a potential reservoir for resurgence of toxigenic *V. cholerae* O1. However, since only a few samples were sequenced in this study from the 2015–2017 outbreak, we cannot rule out the possible presence of T10 sub-lineage in the outbreak and their subsequent discharge in the lake justifying the close genetic relatedness between our environmental isolates F1, F3 and W2 and the clinical T10 strains from the country (Fig 2, Fig 3C). Moreover, despite the well-described environmental reservoirs for *V. cholerae* [28–30], and the evidence of different sub-lineages of the seventh pandemic strains in the aquatic environment, it remains unclear if patients or the Lake Victoria was the original source of the isolates.

Most Tanzanian *V. cholerae* O1 strains isolated after 2014 are T13 and belong to the common MLST type ST69 [14]. Nevertheless, a group of T10 strains caused an outbreak in the city of Kigoma in January 2015 [7] belonging to ST515; a type that had not occurred before in Tanzania and which formed a separate cluster in the phylogenetic tree (Fig 2) within the T10 cluster. These strains belonged to a separate genotype when previously compared by MLVA typing with other genomes from late 2015 [7]. We found that the T10 strains of ST515 most likely originated from the neighboring Democratic Republic of Congo (DRC), a country known for recurrent cholera outbreaks [31] and other neighboring countries where they have caused outbreaks between 2012 and 2013 (Panel C, Fig 3).

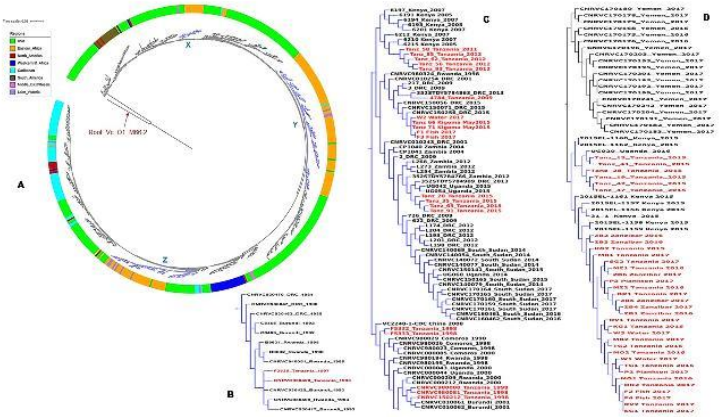


Fig 3. SNP-tree showing global phylogenetic relationships of *V. cholerae* O1 genomes by regions. The blue clades labeled X, Y and Z in panel A indicate Tanzanian strains within the T5, T10 and T13 transmission events, respectively. Panel B is a zoom into the clade X showing the Tanzanian T5 strains. Panel C is the Y clade of Tanzanian strains within a T10 cluster. Panel D displays the clade Z indicating T13 strains including Tanzanian strains. In panels B, C and D, the Tanzanian clinical and environmental strains are highlighted in red.

https://doi.org/10.1371/journal.pntd.0007934.g003

Although T10 strains have been occurring in Tanzania since 1998 through 2015, the sequence type ST515 of the Kigoma strains was different and can be distinguished from the circulating Tanzanian T10 (ST69) (Panel C, Fig 3). However, in May 2015 the refugee camp outbreak in Kigoma caused by the locally circulating T10 ST69 strains only occurred around the refugee camp and could be attributed to the regional spread of this genotype likely favored by population displacement and refugees that fled in at the time due to conflicts in Burundi [7,25]. ST515 has been circulating in DRC before its occurrence in Kigoma in January 2015, thus, the presence of refugee camp in the area and the interaction between local fishermen and refugees from DRC and Burundi could have favored the introduction of T10 ST515 into Tanzania since this type occurred only around Kigoma near the DRC border. T10 were not related to the T13 *V. cholerae* O1 strains (at least 108 SNPs apart) involved in the countrywide cholera outbreaks later in the same year [7] (Fig 2). It was not possible to identify any genome sequences of *V. cholerae* associated with outbreaks in Burundi between 2010 and 2015, a period where most refugees fled into Tanzania. The observed regional transmission is consistent with cholera outbreaks in Tanzania being caused by diverse strains even within the same year and underlines that regional collaborative efforts are required for effective cholera control in countries located around the African Great Lakes.

The occurrence of virulence-associated genes and pathogenicity islands among the *V. cholerae* O1 sequenced in this study was similar to that of strains from previous studies [4,7,9] (S1 Table). Major virulence-associated genes such as *ctxA*, *ctxB*, *zot*, *ace*, *tcpA*, *hlyA*, *mshA*, *rtxA*, *ompU*, and *toxR*, as well as *Vgr*G, *Vas*, *Tsi* proteins of the type VI secretion system, glucose metabolism genes, *als* and the flagella-mediated cytotoxin gene *mak*A were present in all sequenced strains. Moreover, our sequences contained Vibrio Pathogenicity Islands mainly VPI-1 and VPI-2 as well as VSP-1 and VSP-2 normally found in strains of the seventh pandemic.

Nevertheless, a 100-bp nucleotide deletion was observed in the cholera enterotoxin gene (*ctx*A) between positions 1042170 and 1042270 in strains Kg2, Sg2, Zb5 and Zb6 isolated between 2015 and 2017 (S1 Fig) as well as in the published genomes of *V. cholerae* O1 isolated in 2011 and 2012 [7]. To confirm this, we repeated DNA extraction from fresh cultures of the

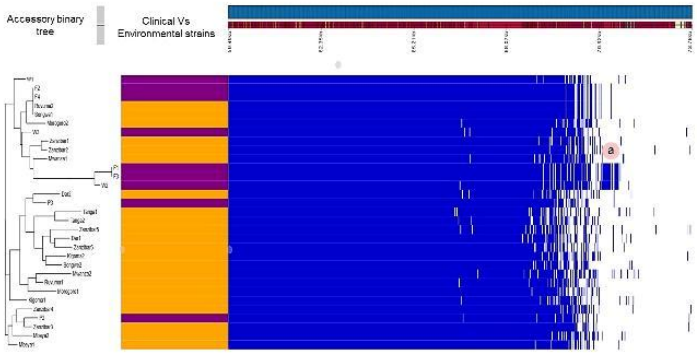


Fig 4. Accessory genome content of pandemic *V. cholerae* from Tanzania (2015–2017 in orange) versus *V. cholerae* O1 isolated in Lake Victoria (purple). The tree at the left shows the accessory binary tree of the accessory genome indicating that clinical strains F2 and F4 are identical to environment strains Rv2 and Sg1. The blue boxes mark presence of genes and white gaps represent absence of gene products. The label (a) shows strains W2, F1 and F3 containing a unique region of proteins from the VSP-2 genomic island like the murin DD-endopeptidase *Mep*M that are absent in other strains.

https://doi.org/10.1371/journal.pntd.0007934.g004

four mentioned outbreak strains and re-sequenced them with results remaining the same. The 100-bp deletion was also confirmed by mapping the reads to the reference *V. cholerae* 2010EL-1786. The concerned strains were negative for *ctx*A in PCR, although the strains originated from stool samples of cholera patients. It remains to be shown how this deletion affects cholera enterotoxin production. These deletions are however, not monophyletic because they are found in strains belonging to two separate clusters from T10 and T13 (Fig 2) suggesting that they could be involved in recombination events since the deletions are occurring within known mobile elements and such events have been reported to affect the structure of *V. cholerae* populations [32]. Differences in the clinical relevance of these recombined strains compared to other strains can however not be demonstrated with the current data. Studies in Mozambique [33] and Mexico [34] have reported outbreak strains of *V. cholerae* O1 lacking *ctx*A. Moreover in Bangladesh, *V. cholerae* isolated from a cholera patient lacked the entire CTX bacteriophage encoding *ctxAB* genes where toxigenic *ctx*A-positive strains co-infected the same individual at the same time [35,36]. The phylogenetic difference between the two strains in that patient suggests that different populations of *V. cholerae* can occur in the same patient at a given time.

When *V. cholerae* O1 strains isolated from Lake Victoria [9] were compared to the latest outbreak strains using a genome-wide approach, we observed that the clinical and environmental isolates share a core genome of 3,321 genes, being the number of genes common to all 31 analyzed strains, out of a total pan-genome size of 3,687 (90.07%) (Fig 4). As shown in the core genome phylogeny where clinical and strains from the Lake were highly related with as low as 3 SNPs apart, the accessory genome also shows that two fish isolates (F2 and F4) are identical to two isolates from patients (Fig 4), confirming the connection between isolates from the environment and from patients. This finding supports our initial argument of an environmental reservoir for *V. cholerae* as a potential source of outbreaks [9] and persistence of pandemic strains in the environment confirming why *V. cholera* O1 has persisted across the three major niche dimensions namely space, time, and habitat [37]. Nevertheless, we still cannot be conclusive on the direction of contamination between the environment and patients. The core-genome is made amongst others of the outer membrane protein genes, the kinase two-component signal transduction histidine-proteins, the chemotaxis proteins and

corroborate previous findings that define species-specific genes of *V. cholerae* supporting environmental adaptation [38]. The accessory genome of the analyzed genomes is however, organized in two main clusters of 110 genes (Fig 4). Between clinical strains and those recovered from the environment, no gene from the accessory genome showed a significant predilection to either of the niches (Benjamini p-value >0.05), substantiating a strong genetic relatedness even at accessory genome level between clinical and environmental *V. cholerae* O1 in Tanzania. This finding is however contrary to previous studies that reported a clear difference between clinical and environmental *V. cholerae* O1 primarily due to lack of virulence-associated genes in most environmental strains [38].

The accessory genome of the analyzed strains essentially constitutes of genomic islands mainly the Vibrio Pathogenicity Islands, toxin co-regulated pilus biosynthesis proteins, the CTX prophage, and resistance genes on the SXT integrative conjugative element (Fig 4). These findings corroborates previous finding [37,39] and confirms that the CTX prophage is not part of the core genome of *V. cholerae* O1. The accessory binary trees (Fig 4) shows a distinct cluster of four non-T13 strains (W2, F1 and F3), with a significantly different accessory genome content (Benjamini p-value< 0.05). The accessory genome of strains recovered from the environment reveal that they are characterized by the presence of bicyclomycin resistance proteins encoded by genes acquired by horizontal gene transfer [39]. Strains W2, F1 and F3 harbored proteins belonging to the genomic island of VSP-2 like the murein DD-endopeptidase *Mep*M, that were absent in remaining strains (Fig 4, label a).

Determinants of antimicrobial resistance

Our sequenced strains showed phenotypic resistance to streptomycin, amoxicillin-clavulanic acid and ampicillin as well as nalidixic acid. Resistance to nalidixic acid was confirmed by the presence of amino acid substitutions in *gyr*A (Ser83-Ile) and *par*C (Ser85Leu). Strains were, however, susceptible to several antimicrobials including gentamicin, ciprofloxacin, ceftazidime, tetracycline, cefotaxime and chloramphenicol. All *V. cholera*e O1 genomes contained resistance genes for chloramphenicol (*cat*B9) and trimethoprim (*dfr*A1/15) with the latter gene being part of the SXT element, but our sequenced strains were susceptible to chloramphenicol in phenotypic tests. Such discrepancy between phenotypic and genotypic profiles have been reported previously [40]. Moreover it has already been reported that the presence of catB9 is not associated with resistance [4].

In accordance with characterization of previous *V cholerae* O1 strains [4,7], our strains contained the SXT integrative conjugative element with genetic similarity to that of *V. cholerae* ICEVchHai1 and harbor the specific integrase genes of the class 1 integron, (*int*I gene). Blast Atlas analysis revealed that strains from 2015 to 2017 have approximately 10-kb nucleotide deletions on the SXT element especially in *flo*R (bp 99050 to 99200), *str*A/B (bp 100350 to 100600; 100800 to 100900; 101600 to 101850) and *sul*2 (bp 102300 to 102450) (S1 Fig) most likely resulting in phenotypic susceptibility to phenicols and sulphonamide. These deletions are characteristic for the T13 sub-lineage of *V. cholerae* O1 El Tor found in the current third wave of the seventh pandemic and have been previously reported in Cameroon [13] and Yemen [26]. These deletions in the ICE fragment may have caused the strains to be less resistant to antimicrobials as compared to the clinical T5 strains isolated in 1993 and 1997, which harbor conjugative IncA/C2 plasmids as reported elsewhere [4] with additional beta-lactam (bla_{CARB-4}), and tetracycline (*tet*B) resistance. No strains isolated after 1998 contained conjugative plasmids. It seems that *V. cholerae* O1 clones of the third wave have lost the *Inc*A/C plasmids over the years [4,26,41].

V. cholerae O1 from Tanzanian outbreaks in a global context

In the global context of the seventh pandemic, Tanzanian strains are located on three time-separated clusters (Panel A, Fig 3). The T5 prototype El Tor strains from 1993 and 1997 are located in a cluster of closely related genomes from India, Bangladesh and China isolated between the 1970´s and the 1990's (Panel B, Fig 3). These strains have been circulating for nearly 20 years in Africa revealing decades long transmission chain between African countries [4]. Their relatedness to strains from Asia shown in our analysis (Panel B, Fig 3) reiterates the Asian origin of initial cholera outbreaks in Tanzania and in Africa [4]. The T10 strains isolated between 1998 and 2012, including the 2015 strains from Kigoma formed a regional cluster (Panel C, Fig 3), confirming spread of *V. cholerae* O1 between Tanzania and other Eastern African countries like Rwanda, Burundi, Kenya, Uganda, DRC, South Sudan, Comoros, and Zambia [4,7,27]. *V. cholerae* O1 isolated in Tanzania during the 2015–2017 outbreak clustered with strains from East Africa mainly the 2015 and 2016 outbreak strains from Kenya and Uganda with a maximum of 50 SNPs difference (Panel D, Fig 3 and S2 Table, sheet 2). The fact that these three neighboring countries that have Lake Victoria in common experienced outbreaks during the same period with genetically closely related strains, also found in the lake, underlines the need for regional collaboration for cholera control and the inclusion of environmental surveillance in control strategies. Moreover, all *V. cholerae* O1 strains isolated after 2014 until 2017 are closely related to *V. cholerae* O1 that caused the devastating 2016–2017 outbreaks in Yemen (Panel D, Fig 3) confirming previous reports on potential human-mediated transmission around the globe [26,42].

In conclusion, genomic analyses of *V. cholerae* O1 responsible for various outbreaks in Tanzania between 1993 and 2017 confirmed that the seventh pandemic El Tor strains caused all outbreaks. This lineage however has undergone significant genetic changes over time. The year 2015 for instance shows the diversity of strains causing various outbreaks in Tanzania because in that year the January outbreaks were caused by T10 ST515 strains, while in May the outbreak in the same city was caused by T10 ST69 and from August 2015 the Kigoma strains were T13. We have confirmed spread within the Eastern African countries notably between Tanzania, the Democratic Republic of Congo, Kenya and Uganda, Rwanda, Burundi, Zambia, South Sudan and Comoros, as well as a global spread between East African countries and Yemen for T10 and T13 strains. Tanzanian older epidemics clones of T5 sub-lineage however most likely originated from India, Bangladesh or China. These findings are consistent with human-mediated spread of cholera around the globe. We have documented potential aquatic environmental reservoir for *V. cholerae* O1 strains, which are closely related to epidemic clones with similar accessory-genome contents. Different sub-lineages of epidemic strains mainly T10 and T13 have been found in the lake substantiating survival, persistence from the lake and favor further human exposure. Tanzanian *V. cholerae* O1 strains show limited antimicrobial resistance and some present nucleotide deletions on the CTX prophage. The observed regional spread calls for well-coordinated cholera control efforts including environmental monitoring of *V. cholerae* O1 in the African Great Lakes regions, which is currently the main cholera hotspot on the African continent. We propose initiation of vaccination programs in countries whose neighbors declare cholera epidemics.

Limitations of the study

In the present study only a limited number (n = 22) of *V. cholerae* O1 isolates collected between 2015 and 2017 have been analyzed from an outbreak that caused over 30, 000 reported cases between August 2015 and early 2018. Considering this limited sample size, it is difficult to rule out the possibility of occurrence of more recent T10 isolates collected in humans during

the outbreaks around Lake Victoria justifying their clustering with our environmental F1, F3 and W2 isolates. Furthermore, the data presented in this study provided evidence of phylogenetic relatedness between clinical and environmental isolates of *V. cholerae* O1 in Tanzania but cannot indicate the direction of pathogen transfer and original source. Moreover, the identification of imported strains of *V. cholerae* through refugees and the occurrence of different sub-lineages over time in Tanzania and beyond in the Great Lakes region cannot effectively guide cholera control without parallel epidemiological studies and interventions from decision makers. The tools used in this study and the available data are not able to predict the next potential sub-lineages to emerge in future epidemics and their clinical relevance in order to proactively propose solutions. Furthermore, the current data does not allow to conclude on the epidemiological relevance of the identified *V. cholerae* O1 from cholera patients containing deletions on the *ctx*A gene, the main virulence factor for cholera toxin production.

Supporting information

S1 Fig. Nucleotide deletions in *ctx*A and on the SXT fragment of *V. cholerae* O1 genomes from Tanzania sequenced in this study. Observed gaps represent the areas of missing nucleotides in strains indicated in the color legend.
(TIF)

S1 Table. Genomic sequence data, virulence profile and occurrence of antimicrobial resistance genes in Tanzanian *V. cholerae* O1 strains.
(XLSX)

S2 Table. Pairwise SNP differences for local and global phylogeny of 589 strains used in the global seventh pandemic tree.
(XLSX)

Acknowledgments

Authors extend their gratitude to Mr Salum Nyanga and Victor Muchunguzi from the Tanzanian National Health Laboratory, Quality Assurance and Training Centre in Dar es Salaam, Tanzania, as well as Mr Mwinyi Msellen, the Director of Training and Research in Mnazi mmoja hospital, Zanzibar, Tanzania for facilitating access to archive *V. cholerae* O1 strains from the 2015–2017 outbreaks.

Author Contributions

Conceptualization: Yaovi Mahuton Gildas Hounmanou, Robinson H. Mdegela, Rene S. Hendriksen, John Elmerdahl Olsen, Anders Dalsgaard.

Data curation: Yaovi Mahuton Gildas Hounmanou.

Formal analysis: Yaovi Mahuton Gildas Hounmanou.

Funding acquisition: Yaovi Mahuton Gildas Hounmanou, Robinson H. Mdegela.

Investigation: Yaovi Mahuton Gildas Hounmanou.

Methodology: Yaovi Mahuton Gildas Hounmanou, Pimlapas Leekitcharoenphon, Egle Kudirkiene, Rene S. Hendriksen, Anders Dalsgaard.

Project administration: Robinson H. Mdegela, John Elmerdahl Olsen, Anders Dalsgaard.

Resources: Yaovi Mahuton Gildas Hounmanou, Pimlapas Leekitcharoenphon, Egle Kudirkiene, Robinson H. Mdegela, Rene S. Hendriksen, John Elmerdahl Olsen, Anders Dalsgaard.

Software: Yaovi Mahuton Gildas Hounmanou, Pimlapas Leekitcharoenphon, Egle Kudirkiene, Rene S. Hendriksen.

Supervision: Robinson H. Mdegela, Rene S. Hendriksen, John Elmerdahl Olsen, Anders Dalsgaard.

Validation: Pimlapas Leekitcharoenphon, Egle Kudirkiene, Robinson H. Mdegela, Rene S. Hendriksen, John Elmerdahl Olsen, Anders Dalsgaard.

Visualization: Yaovi Mahuton Gildas Hounmanou, Pimlapas Leekitcharoenphon, Egle Kudirkiene, Rene S. Hendriksen, John Elmerdahl Olsen, Anders Dalsgaard.

Writing – original draft: Yaovi Mahuton Gildas Hounmanou.

Writing – review & editing: Yaovi Mahuton Gildas Hounmanou, Pimlapas Leekitcharoenphon, Egle Kudirkiene, Robinson H. Mdegela, Rene S. Hendriksen, John Elmerdahl Olsen, Anders Dalsgaard.

References

1. Mbwette TS. Cholera outbreaks in Tanzania. J R Soc Health. 1987; 107: 134–136. https://doi.org/10.1177/146642408710700407 PMID: 3116246
2. Lessler J, Moore SM, Luquero FJ, McKay HS, Grais R, Henkens M, et al. Mapping the burden of cholera in sub-Saharan Africa and implications for control: an analysis of data across geographical scales. The Lancet. 2018; 391: 1908–1915. https://doi.org/10.1016/S0140-6736(17)33050-7 PMID: 29502905
3. WHO. Cholera–United Republic of Tanzania. In: WHO [Internet]. 2018 [cited 15 Oct 2018]. Available: http://www.who.int/csr/don/12-january-2018-cholera-tanzania/en/
4. Weill F-X, 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: 785–789. https://doi.org/10.1126/science.aad5901 PMID: 29123067
5. Mutreja A, Kim DW, Thomson N, Connor TR, Lee JH, Kariuki S, et al. Evidence for multiple waves of global transmission within the seventh cholera pandemic. Nature. 2011; 477: 462–465. https://doi.org/10.1038/nature10392 PMID: 21866102
6. Rashid M, Rashed SM, Islam T, Johura F-T, Watanabe H, Ohnishi M, et al. CtxB1 outcompetes CtxB7 in Vibrio cholerae O1, Bangladesh. J Med Microbiol. 2016; 65: 101–103. https://doi.org/10.1099/jmm.0.000190 PMID: 26487638
7. Kachwamba Y, Mohammed AA, Lukupulo H, Urio L, Majigo M, Mosha F, et al. Genetic Characterization of Vibrio cholerae O1 isolates from outbreaks between 2011 and 2015 in Tanzania. BMC Infectious Diseases. 2017; 17. https://doi.org/10.1186/s12879-017-2252-9 PMID: 28219321
8. Dalusi L, Saarenheimo J, Lyimo TJ, Lugomela C. Genetic relationship between clinical and environmental Vibrio cholerae isolates in Tanzania: A comparison using repetitive extragenic palindromic (REP) and enterobacterial repetitive intergenic consensus (ERIC) fingerprinting approach. African Journal of Microbiology Research. 2015; 9: 455–462. https://doi.org/10.5897/AJMR2014.7307
9. Hounmanou YMG, Leekitcharoenphon P, Hendriksen RS, Dougnon TV, Mdegela RH, Olsen JE, et al. Surveillance and Genomics of Toxigenic Vibrio cholerae O1 From Fish, Phytoplankton and Water in Lake Victoria, Tanzania. Frontiers in Microbiology. 2019; 10. https://doi.org/10.3389/fmicb.2019.00901 PMID: 31114556
10. Nkoko D, Giraudoux P, Plisnier P-D, Tinda A, Piarroux M, Sudre B, et al. Dynamics of Cholera Outbreaks in Great Lakes Region of Africa, 1978–2008. Emerging Infectious Diseases. 2011; 17. https://doi.org/10.3201/eid1711.110170 PMID: 22099090
11. 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. Journal of Computational Biology. 2012; 19: 455–477. https://doi.org/10.1089/cmb.2012.0021 PMID: 22506599
12. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014; 30: 2068–2069. https://doi.org/10.1093/bioinformatics/btu153 PMID: 24642063

13. Kaas RS, Ngandjio A, Nzouankeu A, Siriphap A, Fonkoua M-C, Aarestrup FM, et al. The Lake Chad Basin, an Isolated and Persistent Reservoir of Vibrio cholerae O1: A Genomic Insight into the Outbreak in Cameroon, 2010. Zhou D, editor. PLOS ONE. 2016; 11: e0155691. https://doi.org/10.1371/journal.pone.0155691 PMID: 27191718

14. Siriphap A, Leekitcharoenphon P, Kaas RS, Theethakaew C, Aarestrup FM, Sutheinkul O, et al. Characterization and Genetic Variation of Vibrio cholerae Isolated from Clinical and Environmental Sources in Thailand. Murthy AK, editor. PLOS ONE. 2017; 12: e0169324. https://doi.org/10.1371/journal.pone.0169324 PMID: 28103259

15. Naha A, Pazhani GP, Ganguly M, Ghosh S, Ramamurthy T, Nandy RK, et al. Development and Evaluation of a PCR Assay for Tracking the Emergence and Dissemination of Haitian Variant ctxB in Vibrio cholerae O1 Strains Isolated from Kolkata, India. Journal of Clinical Microbiology. 2012; 50: 1733–1736. https://doi.org/10.1128/JCM.00387-12 PMID: 22357499

16. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. Journal of Antimicrobial Chemotherapy. 2012; 67: 2640–2644. https://doi.org/10.1093/jac/dks261 PMID: 22782487

17. 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: 1355–1361. https://doi.org/10.1128/JCM.06094-11 PMID: 22238442

18. Kaas RS, Leekitcharoenphon P, Aarestrup FM, Lund O. Solving the Problem of Comparing Whole Bacterial Genomes across Different Sequencing Platforms. PLOS ONE. 2014; 9: e104984. https://doi.org/10.1371/journal.pone.0104984 PMID: 25110940

19. Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res. 2016; 44: W242–245. https://doi.org/10.1093/nar/gkw290 PMID: 27095192

20. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015; 31: 3691–3693. https://doi.org/10.1093/bioinformatics/btv421 PMID: 26198102

21. Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. Genome Biology. 2016; 17: 238. https://doi.org/10.1186/s13059-016-1108-8 PMID: 27887642

22. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, Harris SR. Phandango: an interactive viewer for bacterial population genomics. Bioinformatics. 2018; 34: 292–293. https://doi.org/10.1093/bioinformatics/btx610 PMID: 29028899

23. Kim EJ, Lee D, Moon SH, Lee CH, Kim SJ, Lee JH, et al. Molecular Insights Into the Evolutionary Pathway of Vibrio cholerae O1 Atypical El Tor Variants. PLoS Pathog. 2014; 10. https://doi.org/10.1371/journal.ppat.1004384 PMID: 25233006

24. Ghosh-Banerjee J, Senoh M, Takahashi T, Hamabata T, Barman S, Koley H, et al. Cholera Toxin Production by the El Tor Variant of Vibrio cholerae O1 Compared to Prototype El Tor and Classical Biotypes. Journal of Clinical Microbiology. 2010; 48: 4283–4286. https://doi.org/10.1128/JCM.00799-10 PMID: 20810767

25. ReliefWeb. Burundi/Tanzania: Cholera Outbreak—May 2015. In: ReliefWeb [Internet]. 2015 [cited 26 Oct 2019]. Available: https://reliefweb.int/disaster/ep-2015-000058-tza

26. Weill F-X, Domman D, Njamkepo E, Almesbahi AA, Naji M, Nasher SS, et al. Genomic insights into the 2016–2017 cholera epidemic in Yemen. Nature. 2019; 565: 230. https://doi.org/10.1038/s41586-018-0818-3 PMID: 30602788

27. Bwire G, Sack DA, Almeida M, Li S, Voeglein JB, Debes AK, et al. Molecular characterization of Vibrio cholerae responsible for cholera epidemics in Uganda by PCR, MLVA and WGS. PLOS Neglected Tropical Diseases. 2018; 12: e0006492. https://doi.org/10.1371/journal.pntd.0006492 PMID: 29864113

28. Hounmanou YMG, Mdegela RH, Dougnon TV, Madsen H, Withey JH, Olsen JE, et al. Tilapia (Oreochromis niloticus) as a Putative Reservoir Host for Survival and Transmission of Vibrio cholerae O1 Biotype El Tor in the Aquatic Environment. Front Microbiol. 2019; 10. https://doi.org/10.3389/fmicb.2019.01215 PMID: 31214149

29. Islam MS, Zaman MH, Islam MS, Ahmed N, Clemens JD. Environmental reservoirs of Vibrio cholerae. Vaccine. 2019 [cited 9 Jul 2019]. https://doi.org/10.1016/j.vaccine.2019.06.033 PMID: 31285087

30. Lutz C, Erken M, Noorian P, Sun S, McDougald D. Environmental reservoirs and mechanisms of persistence of Vibrio cholerae. Front Microbiol. 2013; 4. https://doi.org/10.3389/fmicb.2013.00375 PMID: 24379807

31. Ingelbeen B, Hendrickx D, Miwanda B, van der Sande MAB, Mossoko M, Vochten H, et al. Recurrent Cholera Outbreaks, Democratic Republic of the Congo, 2008–2017. Emerging Infectious Diseases. 2019; 25: 856–864. https://doi.org/10.3201/eid2505.181141 PMID: 31002075

32. Keymer DP, Boehm AB. Recombination Shapes the Structure of an Environmental Vibrio cholerae Population. Appl Environ Microbiol. 2011; 77: 537–544. https://doi.org/10.1128/AEM.02062-10 PMID: 21075874

33. Garrine M, Mandomando I, Vubil D, Nhampossa T, Acacio S, Li S, et al. Minimal genetic change in Vibrio cholerae in Mozambique over time: Multilocus variable number tandem repeat analysis and whole genome sequencing. Dunachie SJ, editor. PLOS Neglected Tropical Diseases. 2017; 11: e0005671. https://doi.org/10.1371/journal.pntd.0005671 PMID: 28622368

34. Choi SY, Rashed SM, Hasan NA, Alam M, Islam T, Sadique A, et al. Phylogenetic Diversity of Vibrio cholerae Associated with Endemic Cholera in Mexico from 1991 to 2008. mBio. 2016; 7: e02160–15. https://doi.org/10.1128/mBio.02160-15 PMID: 26980836

35. Domman D, Chowdhury F, Khan AI, Dorman MJ, Mutreja A, Uddin MI, et al. Defining endemic cholera at three levels of spatiotemporal resolution within Bangladesh. Nat Genet. 2018; 50: 951–955. https://doi.org/10.1038/s41588-018-0150-8 PMID: 29942084

36. Kendall EA, Chowdhury F, Begum Y, Khan AI, Li S, Thierer JH, et al. Relatedness of Vibrio cholerae O1/O139 Isolates from Patients and Their Household Contacts, Determined by Multilocus Variable-Number Tandem-Repeat Analysis. Journal of Bacteriology. 2010; 192: 4367–4376. https://doi.org/10.1128/JB.00698-10 PMID: 20585059

37. Dutilh BE, Thompson CC, Vicente AC, Marin MA, Lee C, Silva GG, et al. Comparative genomics of 274 Vibrio cholerae genomes reveals mobile functions structuring three niche dimensions. BMC Genomics. 2014; 15. https://doi.org/10.1186/1471-2164-15-654 PMID: 25096633

38. Vesth T, Wassenaar TM, Hallin PF, Snipen L, Lagesen K, Ussery DW. On the Origins of a Vibrio Species. Microb Ecol. 2010; 59: 1–13. https://doi.org/10.1007/s00248-009-9596-7 PMID: 19830476

39. Robins WP, Mekalanos JJ. Genomic Science in Understanding Cholera Outbreaks and Evolution of Vibrio cholerae as a Human Pathogen. Curr Top Microbiol Immunol. 2014; 379: 211–229. https://doi.org/10.1007/82_2014_366 PMID: 24590676

40. Hossain ZZ, Leekitcharoenphon P, Dalsgaard A, Sultana R, Begum A, Jensen PKM, et al. Comparative genomics of Vibrio cholerae O1 isolated from cholera patients in Bangladesh. Lett Appl Microbiol. 2018. https://doi.org/10.1111/lam.13046 PMID: 29981154

41. Spagnoletti M, Ceccarelli D, Rieux A, Fondi M, Taviani E, Fani R, et al. Acquisition and Evolution of SXT-R391 Integrative Conjugative Elements in the Seventh-Pandemic Vibrio cholerae Lineage. mBio. 2014; 5: e01356–14. https://doi.org/10.1128/mBio.01356-14 PMID: 25139901

42. Hendriksen RS, Price LB, Schupp JM, Gillece JD, Kaas RS, Engelthaler DM, et al. Population Genetics of Vibrio cholerae from Nepal in 2010: Evidence on the Origin of the Haitian Outbreak. mBio. 2011; 2. https://doi.org/10.1128/mBio.00157-11 PMID: 21862630

Manuscript D: "Tilapia (*Oreochromis niloticus*) as a Putative Reservoir Host for Survival and Transmission of *Vibrio cholerae* O1 Biotype El Tor in the Aquatic Environment"

Published in ***Frontiers in Microbiology***

Hounmanou YMG, Mdegela RH, Dougnon TV, Madsen H, Withey JH, Olsen JE and Dalsgaard A (2019). Front. Microbiol. 10:1215. doi: 10.3389/fmicb.2019.01215

frontiers in Microbiology

ORIGINAL RESEARCH
published: 31 May 2019
doi: 10.3389/fmicb.2019.01215

Tilapia (*Oreochromis niloticus*) as a Putative Reservoir Host for Survival and Transmission of *Vibrio cholerae* O1 Biotype El Tor in the Aquatic Environment

Yaovi Mahuton Gildas Hounmanou[1], Robinson H. Mdegela[2], Tamegnon Victorien Dougnon[3], Henry Madsen[1], Jeffrey H. Withey[4], John E. Olsen[1] and Anders Dalsgaard[1]*

[1] Department of Veterinary and Animal Sciences, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark, [2] Department of Veterinary Medicine and Public Health, College of Veterinary Medicine and Biomedical Sciences, Sokoine University of Agriculture, Morogoro, Tanzania, [3] Research Unit in Applied Microbiology and Pharmacology of Natural Substances, Research Laboratory in Applied Biology, Polytechnic School of Abomey-Calavi, University of Abomey-Calavi, Abomey-Calavi, Benin, [4] Department of Biochemistry, Microbiology, and Immunology, Wayne State University School of Medicine, Detroit, MI, United States

Studies have reported the occurrence of *Vibrio cholerae* in fish but little is known about the interaction between fish and toxigenic *V. cholerae* as opposed to phytoplankton, which are well-established aquatic reservoirs for *V. cholerae*. The present study determined the role of tilapia (*Oreochromis niloticus*) as a reservoir host for survival and transmission of *V. cholerae* in aquatic environments. Three experiments were performed with one repetition each, where *O. niloticus* (~2 g) kept in beakers were inoculated with four *V. cholerae* strains (5×10^7 cfu/mL). Firstly, infected tilapia were kept in stagnant water and fed live brine shrimp (*Artemia salina*) larvae daily. Secondly, infected tilapia were kept without feeding and water was changed every 24 h. Thirdly, infected tilapia were fed and water was renewed daily. Infected tilapia and non-infected controls were sacrificed on days 1, 2, 3, 7, and 14 post-inoculation and *V. cholerae* were enumerated in intestinal content and water. Another experiment assessed the transmission of *V. cholerae* from infected to non-infected tilapia. The study revealed that El Tor biotype *V. cholerae* O1 and *V. cholerae* non-O1 colonized tilapia intestines and persisted at stable concentrations during the second week of the experiment whereas the Classical biotype was undetectable after 1 week. In stagnant water with feeding, *V. cholerae* counts dropped to 10^5 cfu/ml in water and from 10^7 to 10^4 cfu/intestine in fish after 14 days. When water was renewed, counts in water decreased from 10^7 to 10^3 cfu/ml and intestinal counts went from 10^6 to 10^2 cfu/intestine regardless of feeding. All strains were transmitted from infected to naïve fish after 24 h of cohabitation. Tilapia

OPEN ACCESS

Edited by:
Francesca Leoni, Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche, Italy

Reviewed by:
Diane McDougald, University of Technology Sydney, Australia
Thandavarayan Ramamurthy, Translational Health Science and Technology Institute, India
Taviani Elisa, Universidade Eduardo Mondlane, Mozambique

***Correspondence:**
Yaovi Mahuton Gildas Hounmanou
gil@sund.ku.dk;
gilmahu@yahoo.fr

Specialty section:
This article was submitted to Infectious Diseases, a section of the journal Frontiers in Microbiology

Received: *20 February 2019*
Accepted: *15 May 2019*
Published: *31 May 2019*

Citation:
Hounmanou YMG, Mdegela RH, Dougnon TV, Madsen H, Withey JH, Olsen JE and Dalsgaard A (2019) Tilapia (Oreochromis niloticus) as a Putative Reservoir Host for Survival and Transmission of Vibrio cholerae O1 Biotype El Tor in the Aquatic Environment. Front. Microbiol. 10:1215. doi: 10.3389/fmicb.2019.01215

like other fish may play an essential role in the survival and dissemination of *V. cholerae* O1 in aquatic environments, e.g., the seventh pandemic strains mostly. In this study, tilapia were exposed to high concentrations of *V. cholerae* to ensure initial uptake and follow-up studies with lower doses resembling natural concentrations of *V. cholerae* in the aquatic environment are needed to confirm our findings.

Keywords: *Vibrio cholerae*, tilapia, cholera transmission, microbial ecology, reservoirs

INTRODUCTION

Vibrio cholerae is one of the longest recognized human infectious pathogens, yet there is still much to clarify on the emergence and transmission of cholera, the disease for which *V. cholerae* is the causative agent. *V. cholerae* O1 and O139 are the only serogroups causing cholera, with the leading strains being the toxigenic *V. cholerae* O1 El Tor and Classical biotypes (Dalsgaard et al., 2001). The Classical biotype, however, has not been implicated in cholera outbreaks for several decades and has become extremely rare, if not extinct, in the aquatic environment since the beginning of the seventh cholera pandemic (Safa et al., 2006; Nag et al., 2018). The main virulence factor in humans expressed by all biotypes is cholera toxin. The intestinal colonization of *V. cholerae* in humans requires production of the cholera toxin co-regulated pilus (TCP), whose main transcription activator is ToxT (Faruque et al., 1998; Sanchez and Holmgren, 2011). *V. cholerae* O1 biotype El Tor that lacks active *tox*T can therefore be regarded as non-toxigenic. *V. cholerae* non-O1/O139 strains are ubiquitous in aquatic environments and rarely produce cholera toxin, but can cause sporadic diarrhea (Hounmanou et al., 2016). Most *V. cholerae* non-O1 do not contain *tcp*A (regulated by ToxT), but if present, and the role of *tox*T remains the same which is to regulate the transcription of TCP. In this study, a *tox*T mutant of *V. cholerae* O1 El Tor served to assess whether the lack of transcription of *tcp*A (regulated by ToxT) would affect colonization in tilapia.

Fish are potential carriers for *V. cholerae* and the occurrence of toxigenic and non-toxigenic strains of *V. cholerae* in tilapia (*Oreochromis niloticus*) and African catfish (*Clarias gariepinus*) has been reported during non-cholera outbreak periods (Hounmanou et al., 2016). *V. cholerae* non-O1 has also been isolated in carps (*Rastrineobola argentea*) from Lake Tanganyika (Nyambuli et al., 2018). A study in Bangladesh indicated that fish could serve as a potential vehicle for *V. cholerae* transmission to humans (Hossain et al., 2018), and genetic analysis of *V. cholerae* from a cholera outbreak in Zanzibar suggested that marine fish were implicated in pathogen transmission (Rabia et al., 2017).

In sub-Saharan Africa, where the seventh cholera pandemic has a high impact in terms of morbidities and mortalities (Mengel et al., 2014; Weill et al., 2017), frequent epidemics occur around the African Great Lakes Region (AGLR), where fishing, and fish processing represent an essential socio-economic activity (Nkoko et al., 2011; Plisnier et al., 2015; Ajayi and Smith, 2018). Association between cholera and the aquatic environment is well-established in the AGLR (Urassa et al., 2009; Reyburn et al., 2011). Well-recognized aquatic reservoirs of *V. cholerae* include phytoplanktons, zooplanktons, algae, and cyanobacteria. The role of fish as a reservoir host, however, remain speculative because the mere presence of *V. cholerae* in fish is not sufficient to confirm that fish is a true reservoir host providing multiplication and persistence of the pathogen (Tamplin et al., 1990; Halpern et al., 2004; Islam et al., 2015; Halpern and Izhaki, 2017).

Studies conducted using zebrafish as models indicated that they can be colonized by *V. cholerae* and can transmit the bacteria to naïve zebrafish via excretion (Runft et al., 2014; Mitchell et al., 2017). This therefore calls for further studies to explore the fate of *V. cholerae* in wild caught fish like tilapia.

In the present experimental study, we worked with one of the common edible fish species around the AGLR, namely tilapia (*O. niloticus*). The aim was to determine the role of tilapia in the survival of *V. cholerae* in the aquatic environment and in transmitting the pathogen. The results suggest that *V. cholerae* O1 El Tor (causing the seventh cholera pandemic) do survive in tilapia, which may have important implications in the epidemiology of the ongoing cholera epidemic around the AGLR.

MATERIALS AND METHODS

Bacterial Strains Used in Tilapia Experiments

Vibrio cholerae O1 El Tor (strain E7946) and Classical (strain O395) biotypes used in this study were clinical strains obtained from a laboratory collection (**Table 1**). They are all streptomycin resistant, which allowed for the use of selective isolation procedures during the experiment. An environmental *V. cholerae* non-O1 was isolated from carps (*Rastrineobolla agentea*) in Lake Victoria, Tanzania, and confirmed streptomycin resistant by disc diffusion following standard methods of isolation and identification of *V. cholerae* (Hounmanou et al.,

TABLE 1 | Streptomycin-resistant *V. cholerae* strains used in this study.

Strain ID	Strain characteristics	Source
O395	*V. cholerae* O1, classical biotype	Runft et al., 2014
E7946	*V. cholerae* O1, El Tor biotype	Runft et al., 2014
JW612	*V. cholerae* O1 Δ*toxT* (mutant of E7946)	Runft et al., 2014
V. cholerae non-O1	*Ctx*-negative, *V. cholerae* non-O1 strain isolated from carps in Lake Victoria	This study

2016). The actual serotype of this non-O1 strain was not determined. The strain was included in the study to assess whether colonization in tilapia differs between clinical strains of serogroup O1 and the environmental *V. cholerae* non-O1 strains. *V. cholerae* O1 strain JW612, a *tox*T mutant was included to test whether the lack of transcription of *tcp*A (regulated by *tox*T) would affect colonization of the tilapia (**Table 1**).

Elimination of Natural *V. cholerae* in Tilapia Prior to Experiment

We used tilapia (*O. niloticus*) juveniles (approximately 2 g) obtained from a hatchery at the Sokoine University of Agriculture in Morogoro, Tanzania where the experiments were performed. About 20 tilapia juveniles were placed in beakers (2,000 mL) containing autoclaved tap water with constant aeration. To ensure that the tilapia juveniles did not contain environmental *V. cholerae*, intestinal samples of five tilapia per beaker were plated on thiosulfate-citrate-bile salts-sucrose agar (TCBS) (Oxoid Limited, Hampshire, United Kingdom) after 24, 48, and 72 h.. This was followed by a species-specific PCR for the *omp*W gene (Nandi et al., 2000) using DNA from isolates recovered on TCBS agar and from samples enriched in APW. Water samples from the beakers were also cultured for *V. cholerae*. No *V. cholerae* colonies were detected in any water and fish gut samples and none of the samples produced the 588 bp band expected for the *omp*W gene. Tilapia juveniles confirmed negative for *V. cholerae* were used in the experiment.

Exposure of Tilapia to *V. cholerae* Inoculated in Water

The four *V. cholerae* strains used in the experiments (**Table 1**) were first grown in Luria-Bertani (LB) broth (Difco, Becton Dickinson and company, Maryland, United States) at 37°C for 24 h with agitation. The growth curves of the strains revealed that the overnight cultures (24 h at 37°C) of all the strains reached on average 10^{10} cfu/mL (data not shown). Following procedures described for zebrafish models (Runft et al., 2014; Mitchell et al., 2017), the *V. cholerae* strains were grown overnight in tubes containing 10 ml LB. Tubes with the overnight bacterial cultures were centrifuged at 5,000 *g* for 2–3 min and the pellets with bacteria were washed twice in normal saline solution (sterile water + 0.9% NaCl). Bacterial cells collected from 10 LB tubes were suspended in one mL normal saline/tube and added to beakers (aquarium) containing 2,000 mL of autoclaved tap water and 20 tilapia juveniles. Thus, the count of each *V. cholerae* strain was about 5×10^7 cfu/mL in each beaker, similar to concentrations used in the zebrafish models (Runft et al., 2014). This concentration is high compared to the expected concentration of *V. cholerae* in the natural aquatic environment. However, because gavage is not a natural route of administration in fish, we exposed tilapia to the test strains of *V. cholerae* by immersion in the beakers; a high dose of inoculum was therefore needed to ensure an uptake of the test strains. The beakers were kept at room temperature (25°C) for 2 weeks with constant aeration by an air pump. The experiment included four beakers with the *V. cholerae* strains (**Table 1**) and another beaker containing tilapia juveniles were given 1 mL sterile normal saline solution with no *V. cholerae* (control). Three tilapia juveniles were sacrificed 1, 2, 3, 7, and 14 days after inoculation of the *V. cholerae* strains. Fish were infected similarly, in all experiments.

Experimental Design

We conducted three different experiments with two replicates of each. Each experiment was repeated 2 weeks after the end of the first experiment to have a biological replicate that ensures the validity of the study. In each experiment, tilapia were inoculated with 5×10^7 cfu/mL of various *V. cholerae* strains (**Table 1**) by immersion in beakers as previously described (Mitchell et al., 2017). Tilapia (~2 g) were sacrificed on days 1, 2, 3, 7, and 14 post-exposure and the concentration of *V. cholerae* in water and intestinal content was enumerated as was the water turbidity based on OD_{600} water measurements. The beakers with tilapia, but no *V. cholerae* served as controls and one mL of sterile normal saline solution was added at the onset of the experiment.

In Experiment 1, tilapia were starved for 24 h before the experiment. Two hours after the inoculation of the *V. cholerae* strains, feeding was initiated with fresh hatched brine shrimp (*Artemia salina*) (JBL GmbH & Co., Neuhofen, Germany) served twice a day until termination of the experiment. The brine shrimp were hatched from dry eggs in sterile water and did not contain *V. cholerae* as shown by lack of yellow bacterial colonies when grown on TCBS agar plates. Water in the beakers remained unchanged until the experiment was terminated. This experiment aimed to determine the colonization, the survival and shedding of the *V. cholerae* test strains in tilapia in stagnant water. The control for this experiment was a beaker with tilapia fed with the same brine shrimp and inoculated with sterile normal saline. We chose live feed because of the size of the juveniles being used in the study and also to avoid commercial feeds which may increase water fouling in the beakers.

In Experiment 2, water in the beakers was replaced daily (every 24 h) with fresh sterile water of the same volume (2,000 mL) and fish were not fed. When the water was changed, fish were removed from the initial beaker then washed twice in sterile tap water to remove external *V. cholerae* by rubbing their surface before placing them in a new beaker with sterile water. Infection procedures were the same as in Experiment 1 but this experiment aimed to assess the impact of feeding and water renewal on the survival and excretion of *V. cholerae* in tilapia in the absence of feeding.

In Experiment 3, water was changed like in Experiment 2, but fish were fed brine shrimp free of *V. cholerae* like in the Experiment 1. Tilapia was exposed to *V. cholerae* by immersion as in the other experiments. This experiment aimed to differentiate between the impact of feeding and water exchange on the survival and excretion of *V. cholerae* in tilapia.

Enumeration of *V. cholerae*

Fish were collected every morning and sacrificed. In Experiments 2 and 3 (where water was renewed), fish, and water samples were taken from the beakers before water renewal. After the tilapia juveniles (three fish were collected per time point) were sacrificed, the intestinal content was aseptically removed using sterile scissors and metal blades, and the content placed into 1 mL of sterile normal saline. After homogenization of the intestinal contents by crushing and shaking, serial decimal dilutions of the homogenate were made and 10 μL of dilutions were subsequently spread onto Luria-Bertani Agar (LA) (Difco) plates containing streptomycin (100 μg/mL). One mL water samples from the beakers were collected at each time point and diluted and plated on LA as described for the intestinal content samples. When we obtained 30 or less colonies on LA plates containing streptomycin, all isolates were re-streaked on TCBS agar for confirmation. With higher colony numbers, the identity of at least 30 colonies appearing on the LA plates was confirmed on TCBS agar. Selected El Tor and Classical isolates were collected from TCBS agar plates and confirmed as *ctxA*-positive by PCR to verify that they did not lose their virulence. In Experiments 2 and 3, colony counts were lower due to the daily change of water and, therefore, 100 μL of each dilution was plated on LA. Colony forming units per fish intestine or mL of water were calculated using counts from all plates based on dilutions with valid counts divided by the sum of offset values. Thus in Experiment 1, the detection limit for one sample (fish intestine or water sample) was 900 cfu (2.95 on a logarithmic scale with base 10) per intestine or per mL of water sample and 150 cfu per intestine for six fish samples combined and 450 cfu pr mL for two water samples. For Experiments 2 and 3 where 100 μL was plated, the detection limit for an individual sample was 9 cfu (1.96 on Log_{10} scale). At each time point intestines from three tilapia were analyzed individually for *V. cholerae*.

The use of high concentrations of *V. cholerae* could increase stress in the fish causing excretion of more waste particles which was evaluated by measuring the optical density values of the water. Like in the zebrafish models (Mitchell et al., 2017), the optical density (OD_{600}) of 1 mL water sample was read at 600 nm in a spectrophotometer using normal saline as the blank. Optical density was measured also in the control groups and at different time points during the 14 days of experiment. OD values of the beaker water were measured at each time point along with *V. cholerae* counts in intestine and water samples.

Assessing Transmission of *V. cholerae* Within Tilapia Populations

Ten tilapia juveniles of ∼4 g were exposed to fresh overnight cultures of the *V. cholerae* strains (**Table 1**) in beakers containing 1 L autoclaved tap water as described above (approximately 5×10^7 cfu/mL). After 6 h of exposure, a time which previously was found sufficient to allow colonization of *V. cholerae* in zebrafish (Runft et al., 2014), the tilapia juveniles were washed in autoclaved tap water twice to remove any external *V. cholerae* present on the fish body. The juveniles were then placed in another beaker with sterile water containing ten naive tilapia juveniles of ∼2 g with the smaller size allowing differentiation from the larger 4 g fish. After 24 h of cohabitation, four of the naïve tilapia juveniles were sacrificed per beaker and intestinal *V. cholerae* populations were enumerated as described above.

Statistical Analyses

Vibrio cholerae counts (x) were calculated as total count per fish intestine (fish samples) or total count per mL (water samples) for the two repeated trials. Comparisons of bacterial counts in fish samples [$\log_{10}(x+1)$] between strains and over time was done using multiple linear regression where also the interaction between strain and time was assessed. Repetition was not a significant predictor of bacterial counts neither when tested alone nor in the full model and therefore it was left out of the final analysis. *V. cholerae* counts in water samples were not compared in a similar manner as there was only one sample for each repetition, strain and time point, but those counts were correlated with average counts in the fish using linear regression. Model assumptions were verified using normal probability plot of standardized residuals and histogram of residuals. Homoskedatiscity was checked using rvf-plot plus Breusch-Pagan/Cook-Weisberg test for heteroskedasticity. *P*-values < 0.05 were taken to indicate significant differences in Stata (Version 12, StataCorp, College Station, TX, United States). Bacterial counts from transmission experiments enumerated 24 h post exposure were only strain dependent. Therefore, one-way ANOVA was performed to compare mean counts for the two trials among strains.

RESULTS

Survival of *V. cholerae* in Tilapia Kept in Stagnant Water and Given Live Feed

In Experiment 1, tilapia were exposed to 5×10^7 cfu/mL of *V. cholerae* as described in the Section "Materials Methods." They fish were kept in stagnant water and fed brine shrimp for the 2 weeks duration of the experiment. Tilapia juveniles exposed to *V. cholerae* were found to be colonized by *V. cholerae* within 24 h after exposure, with average intestinal counts for all test strains varying between 10^7 cfu/intestine on day 1 post infection to 10^5 cfu/intestine 14 days after infection. *V. cholerae* counts declined over time for all four strains (**Figure 1A**), with a significant interaction seen between time and strain ($p < 0.001$).There was no difference in concentration of different test *V. cholerae* strains for the first 2 days. However, 3 and 7 days after exposure, counts were lower for the Classical biotype than those of the *V. cholerae* non-O1 strain ($p < 0.05 - p < 0.001$ depending on the day). One fish exposed to the *V. cholerae* O1 Classical biotype did not contain *V. cholerae* on day 7 (i.e., below the detection limit of 2.95 on a $\log_{10}$-scale). After 14 days, all fish exposed to the *V. cholerae* O1 Classical biotype no longer contained detectable levels of this strain. Both *V. cholerae* O1 El Tor and a Δ*tox*T *V. cholerae* O1 El Tor, had lower counts

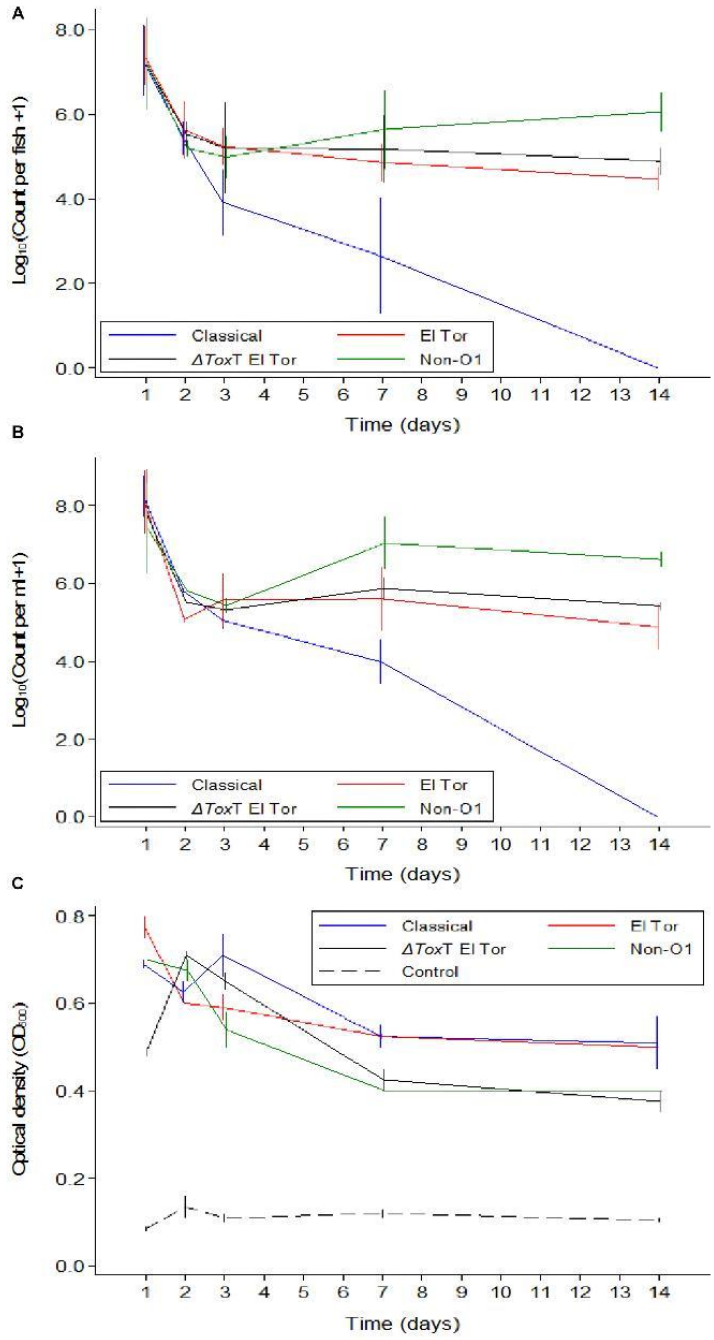


FIGURE 1 | Mean counts of *V. cholerae* from two repetitions in tilapia intestines **(A)** count per fish, in water **(B)** count per mL, and absorbance of water **(C)** over time when tilapia were kept in the same water and given live feed. Error bars indicate 95% CL. Each strain is slightly off its exact *x*-value to allow distinction of the error bars. Relevant statistical differences between strains and time points are indicated in the text.

than the *V. cholerae* non-O1 strain ($p < 0.05$). Except for the Classical biotype strain, strains survived in tilapia for 2 weeks at constant levels. No *V. cholerae* was detected in the uninfected control tilapia for the duration of the experiment.

Vibrio cholerae numbers in water were similar to numbers in fish intestinal content. The Classical biotype was not detected in water 1 week after exposure (**Figure 1B**). Counts in water correlated with average counts in the fish intestine (**Figure 2**). However, the average *V. cholerae* counts in water varied between 10^8 cfu/mL on day 1 to 10^6 cfu/mL 14 days after inoculation of test strains and was higher than those observed in fish intestines. No *V. cholerae* was isolated from water in the control beakers.

Optical density of water from beakers with fish exposed to *V. cholerae* strains was considerably higher than the density in aquaria with unexposed fish, even for the Classical biotype strain after it was no longer detectable. Significant predictors of optical density when analyzing only the four *V. cholerae* strains were strain ($p < 0.05$) and time ($p < 0.001$). The Classical and El Tor biotypes did not differ when adjusting for time, while the Δ*toxT* mutant of El Tor, and the non-O1 strains both differed from the Classical strain but did not differ between themselves. Optical water densities were lower at day 7 and 14 than during day 1 (**Figure 1C**).

Survival of *V. cholerae* in Tilapia When Water Was Changed Daily and Tilapia Were Not Fed

In Experiment 2, fish were exposed to *V. cholerae* as described above. However, in contrast to Experiment 1, the tilapia were not fed and the water was changed daily. Thus, after the initial infection only *V. cholerae* that multiplied in the intestine and excreted by the fish would be detected in the water. With daily water exchange and absence of feeding, tilapia were still colonized by all strains of *V. cholerae* 24 h post infection with average counts around 10^6 cfu/intestine (**Figure 3A**). Up to 2 days post infection, there was no difference in colonization levels between strains ($p > 0.05$). However, from day 3 post infection, the concentration of the Classical biotype strain decreased significantly in the fish intestines and was undetectable after 1 week ($p < 0.001$). Despite the absence of feeding and with the constant daily water exchange, tilapia remained colonized with the three other strains of *V. cholerae* until the end of the 1 weeks, but the counts dropped significantly from day to day, most significantly during the first 7 days ($p < 0.05$). No *V. cholerae* growth was detected in the uninfected control group. During the second week, 5–15% mortality was recorded in all beakers probably due to starvation.

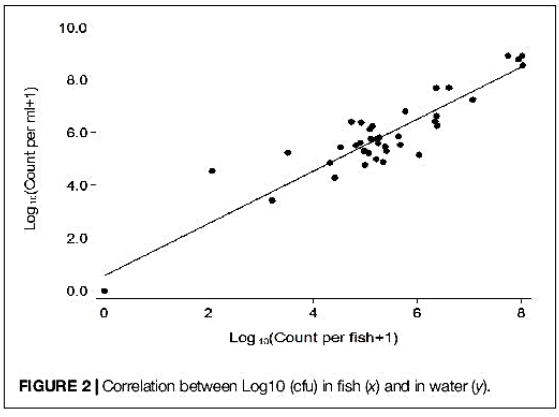


FIGURE 2 | Correlation between Log10 (cfu) in fish (*x*) and in water (*y*).

In Experiment 2, *V. cholerae* concentrations in water were similar to those in the intestine, varying from 10^7 cfu/mL on day one post infection to 10^3 cfu/mL 14 days after inoculation, with significant daily decreases ($p < 0.05$). Like in the fish intestine, the Classical biotype strain could no longer be detected in the water after 1 week (**Figure 3B**). The other three strains remained present in the water despite the daily water replacement with fresh sterile water, suggesting continuous *V. cholerae* multiplication, and excretion by the fish. No *V. cholerae* was detected in the uninfected fish from the control beakers.

Overall, when tilapia were starved and water was exchanged daily, the optical density of water in beakers containing infected tilapia remained significantly higher than in the control beakers where fish were not infected ($p < 0.05$). The difference was more pronounced in the first week ($p < 0.001$); however, in the second week of the experiment, excretion levels decreased as the OD values from infected groups became statistically similar to the OD values of the uninfected control even though the numbers were higher than that of the control ($p > 0.05$). There was no significant difference between the four strains in terms of excretion at any time point ($p > 0.05$). Despite the absence of the Classical biotype strain in the second week, the OD values in that aquarium remained similar to that of the other strains (**Figure 3C**).

Survival of *V. cholerae* in Tilapia When Water Was Changed Daily and Tilapia Were Given Live Feed

In Experiment 3, fish were again infected with *V. cholerae* and water was renewed daily. In contrast to Experiment 2, these fish were also fed daily. *V. cholerae* counts in tilapia intestines, counts in water, and the water OD measurements showed similar values and trends as those recorded in Experiment 2 where tilapia were starved (**Figures 4A–C**). However, there was no fish mortality in Experiment 3 as compared to the Experiment 2 where fish were not fed.

Comparison of *V. cholerae* Counts in Fish Guts and in Water Between the Three Experiments

A comparison between the three experiments was made to distinguish if feeding or water renewal influenced the survival of *V. cholerae* in fish. In stagnant water (Experiment 1), *V. cholerae* counts dropped from 10^7 to 10^5 cfu/mL and from 10^7 to 10^4 cfu/intestine in fish. However, in Experiments 2 and 3 where water was changed, *V. cholerae* in water decreased from 10^7 to 10^3 cfu/mL and gut counts ranged between 10^6 and

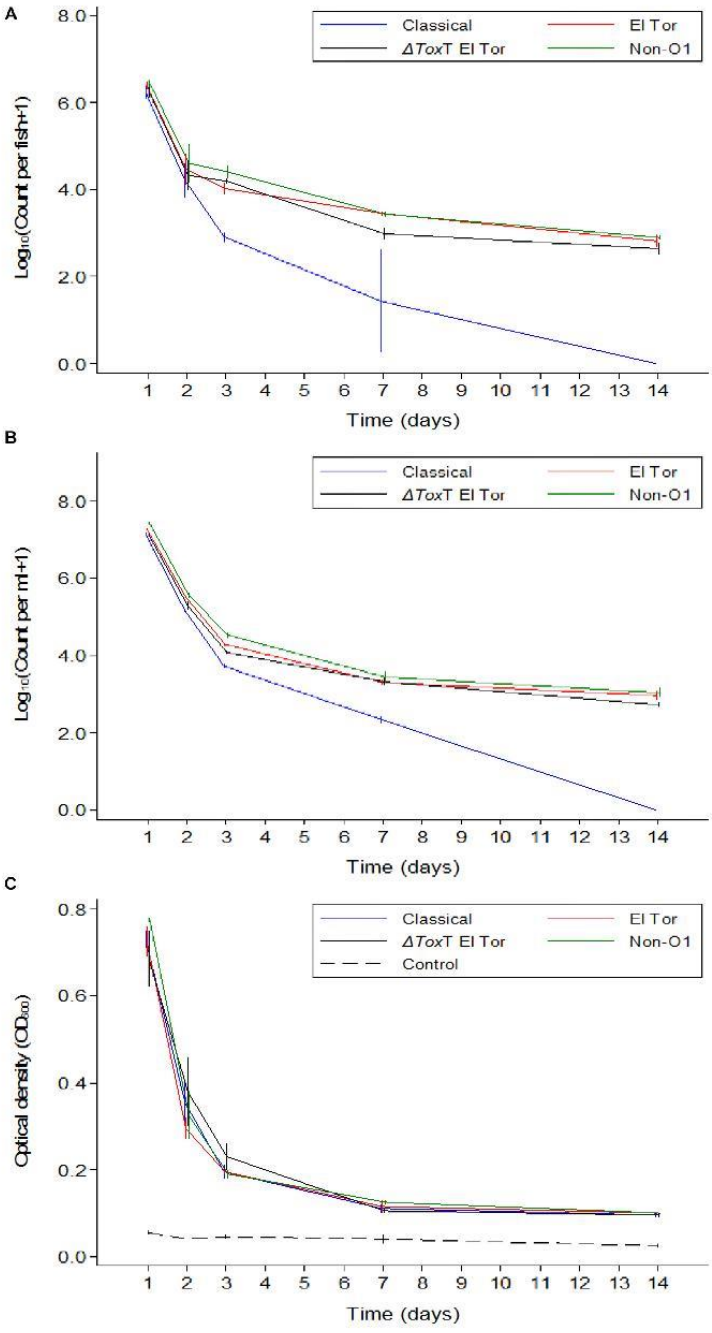


FIGURE 3 | Mean counts of *V. cholerae* from two repetitions in tilapia intestines **(A)** count per fish, in water **(B)** count per mL, and absorbance of water **(C)** over time when aquarium water was changed daily and tilapia were not fed. Error bars indicate 95% CL. Each strain is slightly off its exact *x*-value to allow distinction of the error bars. Relevant statistical differences between strains and time points are indicated in the text.

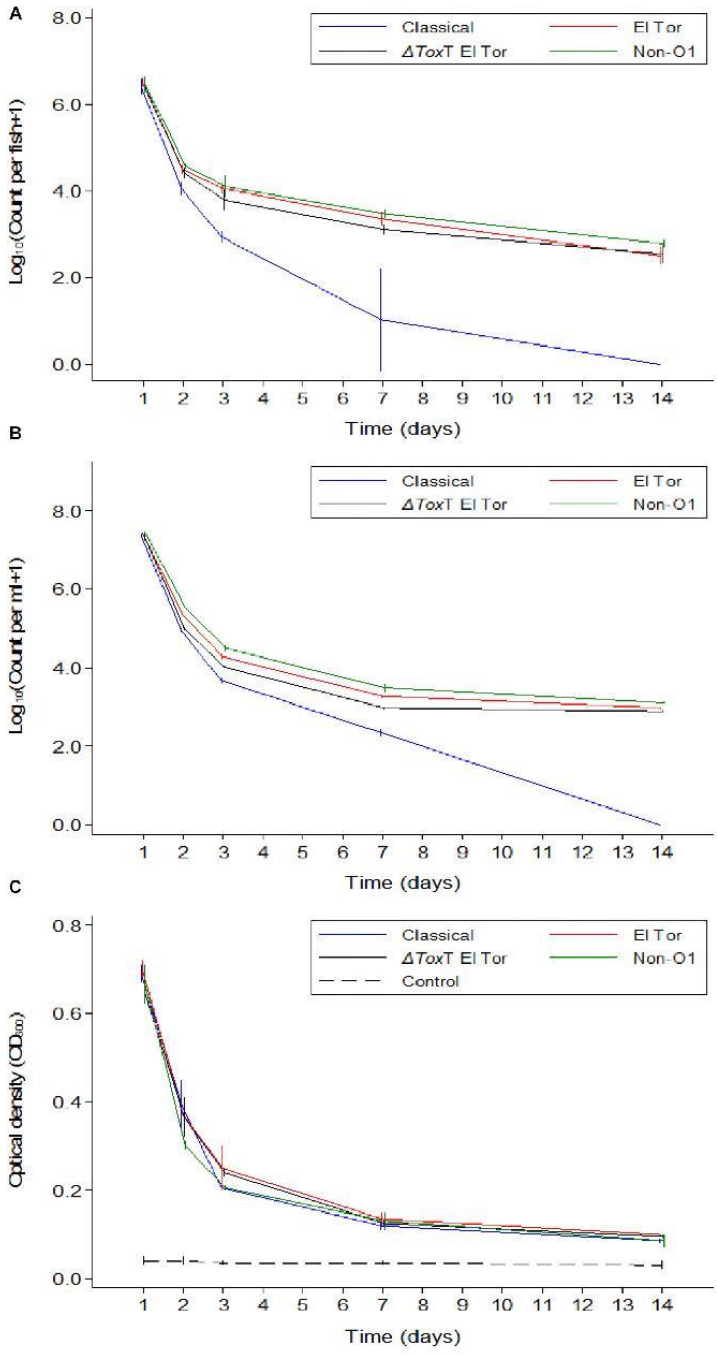


FIGURE 4 | Mean counts of *V. cholerae* from two repetitions in tilapia intestines **(A)** count per fish, in water **(B)** count per mL, and absorbance of water **(C)** over time when aquarium water was changed daily and tilapia are given live feed. Error bars indicate 95% CL. Each strain is slightly off its exact *x*-value to allow distinction of the error bars. Relevant statistical differences between strains and time points are indicated in the text.

10^2 cfu/intestine, with significant daily decreases ($p < 0.05$). When water was replaced daily with fresh sterile tap water in Experiments 2 and 3, *V. cholerae* counts were statistically similar both in water and intestines regardless of presence or absence of feeding ($p > 0.05$; **Figures 5A,B**). This indicates that in both water and tilapia intestines, feeding did not have any impact on concentrations of *V. cholerae* ($p > 0.05$, **Figure 5**). In contrast, higher *V. cholerae* counts were recovered in water and intestine when water was not changed, i.e., comparing Experiment 1 with Experiments 2 and 3 ($p < 0.05$; **Figures 5A,B**). This significant variation between the three experiments was due to water renewal.

Throughout the experimental period, *V. cholerae* counts in tilapia intestines were not statistically different between Experiments 2 and 3 ($p > 0.05$); however, these counts were significantly lower when compared with Experiment 1 at each time point ($p < 0.01$). Furthermore, in water, the comparison between experiments revealed that there was no significant difference between the three experiments during the two first days, but from day 3 to day 14, the change of water significantly influenced bacterial counts ($p < 0.001$).

The OD values of water was higher when the experiment was done in the same water as compared to when water was changed daily ($p < 0.05$, **Figure 5C**). Apart from day one post infection when OD values from the three experiments were statistically similar, the absorbance of water differed significantly in Experiment 1 when compared to Experiments 2 and 3 ($p < 0.001$), i.e., OD values in Experiments 2 and 3 were not statistically different ($p > 0.05$). We conclude that feeding had no significant influence on the OD values but the change of water did reduce the water turbidity overtime as the concentration of *V. cholerae* decreased.

Transmission of *V. cholerae* Within Tilapia Populations

As we observed a stable survival of *V. cholerae* over 2 weeks in infected tilapia compared to the uninfected controls, a question emerged whether the bacteria could be transmitted from infected to naïve tilapia. After 24 h of cohabitation with infected tilapia (without feeding), average *V. cholerae* counts of 10^5 cfu/intestine were observed in naïve fish that were initially tested free of *V. cholerae*. The concentration of *V. cholerae* found in naïve tilapia were similar for all four test strains ($p > 0.05$, **Figure 6**).

DISCUSSION

Experimental exposure model studies with *V. cholerae* in zebrafish suggest that *V. cholerae* can colonize fish guts and be transmitted among zebrafish populations (Runft et al., 2014; Mitchell et al., 2017). Our results in tilapia are consistent with the observations in zebrafish. High counts of the different strains of *V. cholerae* in tilapia intestines were observed 24 h after exposure as a sign of colonization. The sharp drop in *V. cholerae* counts in fish and water between 24 and 48 h post inoculum in all three colonization experiments could be due to natural shock of the *V. cholerae* bacterial cells attributable to changed environments, which may lead to a dormant state also known as viable but not culturable (Kamruzzaman et al., 2010). However, after day 2 to 3 post infection, concentrations of *V. cholerae* remained more or less stable, demonstrating adaptation and survival in the tilapia. Non-toxigenic strains, notably Δ*toxT* of *V. cholerae* O1 El Tor and the environmental non-O1 strain, were also able to colonize fish and persist over time. Intestinal colonization of *V. cholerae* in humans requires production of the cholera TCP, whose main virulence transcription activator is ToxT (Faruque et al., 1998; Sanchez and Holmgren, 2011). The fact that *tox*T mutants and non-O1 strains of *V. cholerae* were found in tilapia intestine over time is in accordance with observations in zebrafish (Runft et al., 2014) and suggests that TCP is not essential for *V. cholerae* colonization of fish. Moreover, studies have discovered a novel flagella-mediated cytotoxin MakA which is proposed to be involved in *V. cholerae* intestinal colonization in zebrafish (Dongre et al., 2018). *V. cholerae* strains used in our study are wild types and possess a flagellum, so the secretion of MakA protein associated with flagella could be involved in the colonization of tilapia.

In Experiment 1, where fish were kept in stagnant water and fed live brine shrimp (*A. salina*) free of *V. cholerae*, there were high *V. cholerae* counts in intestines and water together with high OD values. Since the infection dose (5×10^7) was the same in all experiments, the higher concentrations of *V. cholerae* observed in Experiment 1 compared to findings in Experiments 2 and 3, in which water was changed daily, were thought to be associated with the continuous provision of brine shrimp, as their presence could enhance attachment and multiplication of *V. cholerae* in the fish intestine. The ADP-ribosylating cholix toxin in *V. cholerae* has been shown to play an important role in the survival of the organism in the aquatic environment and facilitates its attachment to crustaceans, notably the brine shrimp. Moreover, *V. cholerae* are known for their ability to attach to chitin exoskeletons of shrimp, copepods and other crustaceans that serve as substrate for their survival and multiplication (Huq et al., 1983; Tamplin et al., 1990; Hood and Winter, 1997; Patra and Mohamed, 2003). However, results in Experiment 3 where water was replaced on a daily basis and fish were fed with the same live feed rejected the hypothesis that brine shrimp could enhance colonization, because the change of water was found to be the only significant variable associated with *V. cholerae* concentrations.

Comparison of the three experiments shows that irrespective of feeding and water exchange, tilapia were colonized by environmental non-O1 *V. cholerae* as well as *V. cholerae* seventh pandemic El Tor (7PET) strains and were isolated beyond 2 weeks. This strongly suggests that in natural aquatic environments, where fish can live in stagnant or running water with presence of various feed items, tilapia may constitute a reservoir of toxigenic, and non-toxigenic strains of *V. cholerae*. It is worth noticing that the concentration of *V. cholerae*, i.e., 10^7 cfu per ml water in the beakers, was higher as compared to concentrations that can be expected in natural aquatic environments during non-cholera outbreak periods (Senderovich et al., 2010). In contrast, there are little data available about the actual concentration of *V. cholerae* O1 in such environments. In

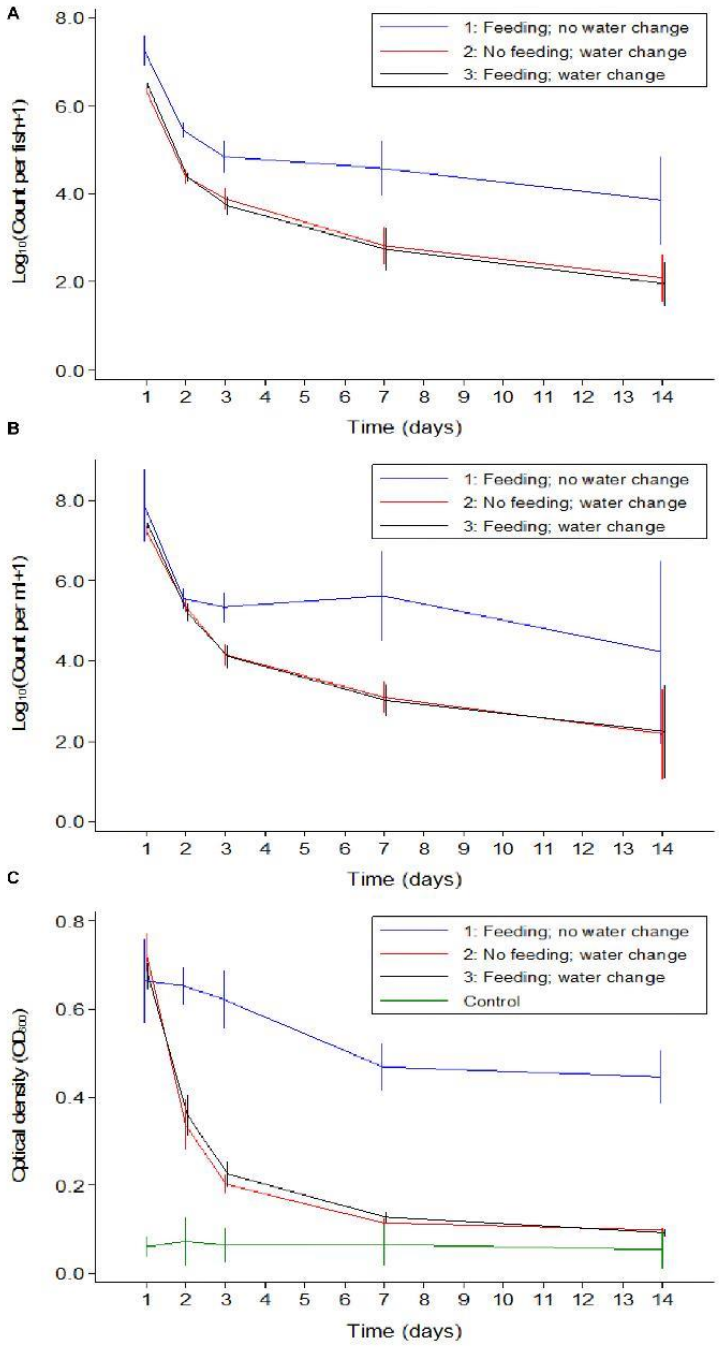


FIGURE 5 | Comparison of *V. cholerae* counts (all strains combined) in tilapia guts **(A)**, in water **(B)**, and absorbance **(C)** between the three experiments overtime. Relevant statistical differences between experiments and time points are indicated in the text.

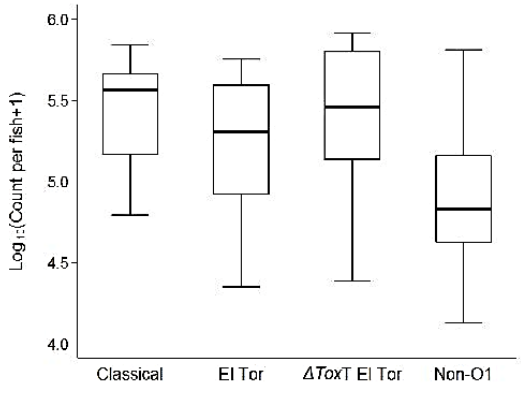


FIGURE 6 | Transmission of *V. cholerae* from infected to naïve tilapia via excretions 24 h post cohabitation using box and whiskers plot. The thick line is the median and the box is the interquartile range and whiskers the range. Counts are from the two repetitions.

a previous study, we did report that tilapia were able to live in raw sewage and were found to carry *V. cholerae* O1 (Hounmanou et al., 2016). Moreover, tilapia in our experiment were infected by immersion and we therefore used a similar dose of 10^7 cfu per ml as in the zebrafish experiments (Runft et al., 2014; Mitchell et al., 2017) to ensure an uptake of the test strains. It is not known how many *V. cholerae* cells fish ingest in natural water systems. The high inoculum of *V. cholerae* used in this study could enhance colonization and may represent a limitation of our study. Nevertheless, even at low concentrations of *V. cholerae* similar to natural conditions, like from day 7 in all our experiments (10^3 CFU/intestine), tilapia remained colonized, with the O1 El Tor biotype maintaining the highest numbers. Furthermore, the transmission experiment with naïve tilapia placed in beakers with tilapia carrying *V. cholerae* in the intestine showed that when the naïve tilapia were exposed to about $10^{5.5}$ cfu per mL water they became infected and had similar bacterial concentrations in their intestine after 24 h (**Figure 6**). It should be noted that we did not determine the *V. cholerae* concentration in the water in the transmission experiment (**Figure 6**) and that the stated expected concentration in the water of $10^{5.5}$ cfu per mL is based on values found at day 2 in Experiment 2 (**Figure 3B**), where infected tilapia were washed and transferred to beakers with fresh water. Overall, the results indicate that irrespective of the initial concentration of *V. cholerae* in water, tilapia can become colonized with *V. cholerae* and act as a reservoir for transmission and long-term survival.

Findings from Experiments 2 and 3 show that the concentration in water after 1 week was around 10^4 to 10^3 cfu/mL similar to what has been reported in the natural environmental waters (Senderovich et al., 2010). One week later, the concentration of *V. cholerae* was around 10^3 to 10^2 cfu/mL. Despite a low concentration of *V. cholerae* in water seen during the last 7 days and the continuous daily water renewal, the OD values of the water in the beakers remained higher than the OD values in the beakers of the control fish. This suggests that the increased OD water values were due to excreted material from the tilapia, i.e., stress-related discharges, probably due to the initial high concentration of *V. cholerae* in the water. *V. cholerae* are able to colonize tilapia over an extended time span, multiply in the intestine, and be excreted into the aquatic environment but a high initial concentration could be stressful for the fish. We therefore suggest further studies to explore lower infection doses administered possibly by gavage to ensure sufficient uptake. Moreover, even when the Classical biotype of *V. cholerae* was no longer detectable in tilapia intestines and in water, the optical density of water in those beakers remained higher than in the control beakers. This is consitent with observations in zebrafish that heat-killed *V. cholerae* still induced mild diarrhea in zebrafish (Mitchell et al., 2017). This further suggests that the discharges and water turbidity provoked by *V. cholerae* in tilapia is neither due to cholera toxin genes nor to viability or biotype of *V. cholerae* but probably caused by the stress generated by the high initial infection dose of *V. cholerae*. Furthermore, the flagella-mediated secretion of MakA cytotoxin was suggested as a source of toxicity and death in zebrafish infected with wild-type *V. cholerae* (Dongre et al., 2018).

The absence of Classical biotype *V. cholerae* after 1 week and the persistence of the seventh cholera pandemic biotype El Tor *V. cholerae* O1 strains is similar to findings in zebrafish (Runft et al., 2014) and consistent with the rare isolation or extinction of the Classical biotype in the ongoing cholera pandemic (Echenberg, 2011; Weill et al., 2017). The El Tor biotype and the non-O1 serogroup strains seem more fit in the fish gut (aquatic environment) than the Classical O1 biotype strains which may explain the increasing recovery of these strains in most contemporary environmental studies (Hounmanou et al., 2016; Bwire et al., 2018). The persistence of *V. cholerae* O1 biotype El Tor in tilapia and water is of public health relevance as it provides evidence of environmental survival of the current pandemic El Tor biotype strains where they can emerge from and cause epidemics. Furthermore, *in vitro* and *in vivo* experiments have demonstrated that in the presence of glucose, *V. cholerae* of the Classical biotype generates organic acids that inhibit their growth, while the growth of El Tor biotype is enhanced due to their ability to produce acetoin (2,3-butanediol), a neutral fermentation end product (Yoon and Mekalanos, 2006; Sengupta et al., 2017; Nag et al., 2018). It could therefore, be that carbohydrates present in the water of the beakers, e.g., droppings from the tilapia, did facilitate glucose metabolism of *V. cholerae,* resulting in loss of viability of the Classical biotype and survival of El Tor biotype. Such unfavorable conditions may also cause the strains, especially the Classical biotypes, to enter a dormant state, known as viable but not culturable (VBNC) (Bari et al., 2013; Xu et al., 2018), which may be one explanation as to why the Classical biotype strains were not detected after 1 week of the experiments.

The concentrations of *V. cholerae* in tilapia intestines correlated with those in the water. As the number of fish in the beakers decreased over time, *V. cholerae* counts in water decreased. This strong correlation ($p < 0.0001$) between

counts in water and fish indicates that *V. cholerae* have reduced ability to multiply in clean water (Colwell et al., 2003) as compared to the intestine and is continuously excreted by the fish host. Thus, the main significant predictor of *V. cholerae* concentrations in water was the concentration of *V. cholerae* in the fish intestine. Moreover, the observed similar *V. cholerae* counts in water and in fish intestines in all experiments is likely attributable to the fact that the beakers used as aquarium provided a restrained space to the fish in a low volume of water (2 L). When *V. cholerae*-free tilapia cohabitated with infected tilapia in sterile water overnight, their intestine was also colonized, providing evidence of transmission. The transmission was not strain-dependent, as all the four test strains had similar transmission rates, demonstrating that toxigenic and nontoxigenic strains can equally be transmitted between tilapia populations and that *V. cholerae* can survive and amplify in tilapia, and also be disseminated from tilapia. Our findings are similar to reports in zebrafish models (Runft et al., 2014). Furthermore, the fact that naïve tilapia became infected in the transmission experiments substanciates again that tilapia were effectively and stably colonized by *V. cholerae* that then was disseminated between fish populations. Although only intestines were studied in our experiments, other organs like gills and skin could also play a significant role in the transmission process as they would provide nutrients favoring colonization of *V. cholerae*. However, previous studies have demonstrated that intestines are the main factors involved in colonization and transmission of *V. cholerae* in fish (Runft et al., 2014) and fish were washed twice with sterile saline by rubbing before the transmission experiment to remove external bacteria. The potential epidemiological importance of this study is that during a cholera outbreak, while all efforts are deployed toward containing the epidemics at human level, fish may serve as vehicle of dissemination of the bacteria in other areas, which may subsequently be hit by the same outbreak even when human patients are quarantined in the initial outbreak settings. The possible role of fish in transmitting *V. cholerae* is further supported by the findings that fish eating birds such as Great cormorants have been found to carry and disperse *V. cholerae* in space and time as they feed on infected fish and get colonized by *V. cholerae* (Laviad -Shitrit et al., 2017).

In summary, we have demonstrated that toxigenic *V. cholerae* O1 biotype El Tor and non-toxigenic strains of *V. cholerae* colonized the intestines of tilapia and were transmitted to naïve tilapia. This study provides answers to a hypothesis posed in a previous study (Halpern and Izhaki, 2017) that fish can be colonized by *Vibrio cholerae* and subsequently horizontally transfer *V. cholerae* to other fish within the same species, and probably to other fish species. This suggests that *V. cholerae* colonizes and persists in fish, and is transmitted between fish in aquatic environments, which may influence the epidemiology of cholera. Tilapia and other fish are potential reservoir hosts involved in the survival, excretion and transmission of *V. cholerae* in time and space. Cholera surveillance strategies may need to be updated accordingly including analysis of fish for the presence of *V. cholerae* O1 in aquatic environments. We furthermore suggest further studies to confirm the role of tilapia as an environmental reservoir host of *V. cholerae* O1 biotype El Tor using lower infection doses administered possibly by gavage to ensure sufficient uptake and limit stress to the fish.

DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without under reservation, to any qualified researcher.

ETHICS STATEMENT

The present study was approved by the ethical review board through the ethical clearance certificate for conducting animal related research in Tanzania with ethical approval number SUA/CVMBS/018/07 (File submitted to editorial). Euthanasia of fish during the study and disposal of waste materials were performed according to instructions of the above-mentioned ethical approval.

AUTHOR CONTRIBUTIONS

YH designed the study, carried out the experiments in the laboratory, analyzed the results, and drafted the manuscript. RM supervised the experiments and critically reviewed and edited the original draft of the manuscript. TD participated in critical reviewing and editing of the manuscript. HM contributed to statistical analysis and data interpretation. JW provided the test strains and critical revision of the manuscript. JO and AD conceived the study and reviewed the manuscript. AD validated the data and supervised the study. All authors read and approved the final manuscript.

FUNDING

The study was supported by the Danish International Development Assistance (Danida) – funded project "Innovations and Markets for Lake Victoria Fisheries" (IMLAF), DFC File No. 14-P01-TAN and the International Foundation for Sciences (IFS), Stockholm, Sweden, through the grant A6100-1 to YH. JW was funded by NIH grant R01AI127390 from the United States Public Health Service.

ACKNOWLEDGMENTS

The authors extend their gratitude to the funders.

REFERENCES

Ajayi, A., and Smith, S. I. (2018). Recurrent cholera epidemics in Africa: which way forward? A literature review. *Infection* doi: 10.1007/s15010-018-1186-5 [Epub ahead of print].

Bari, S. M. N., Roky, M. K., Mohiuddin, M., Kamruzzaman, M., Mekalanos, J. J., and Faruque, S. M. (2013). Quorum-sensing autoinducers resuscitate dormant *Vibrio cholerae* in environmental water samples. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9926–9931. doi: 10.1073/pnas.1307697110

Bwire, G., Debes, A. K., Orach, C. G., Kagirita, A., Ram, M., Komakech, H., et al. (2018). Environmental surveillance of *Vibrio cholerae* O1/O139 in the five african great lakes and other major surface water sources in Uganda. *Front. Microbiol.* 9:1560. doi: 10.3389/fmicb.2018.01560

Colwell, R. R., Huq, A., Islam, M. S., Aziz, K. M. A., Yunus, M., Khan, N. H., et al. (2003). Reduction of cholera in Bangladeshi villages by simple filtration. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1051–1055. doi: 10.1073/pnas.0237386100

Dalsgaard, A., Serichantalergs, O., Forslund, A., Lin, W., Mekalanos, J., Mintz, E., et al. (2001). Clinical and environmental isolates of *Vibrio cholerae* serogroup O141 Carry the CTX phage and the genes encoding the toxin-coregulated pili. *J. Clin. Microbiol.* 39, 4086–4092. doi: 10.1128/JCM.39.11.4086-4092.2001

Dongre, M., Singh, B., Aung, K. M., Larsson, P., Miftakhova, R., Persson, K., et al. (2018). Flagella-mediated secretion of a novel *Vibrio cholerae* cytotoxin affecting both vertebrate and invertebrate hosts. *Commun. Biol.* 1:59. doi: 10.1038/s42003-018-0065-z

Echenberg, M. J. (2011). *Africa in the Time of Cholera: a History of Pandemics from 1817 to the Present*. New York, NY: Cambridge University Press.

Faruque, S. M., Albert, M. J., and Mekalanos, J. J. (1998). Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiol. Mol. Biol. Rev.* 62, 1301–1314.

Halpern, M., Broza, Y. B., Mittler, S., Arakawa, E., and Broza, M. (2004). Chironomid egg masses as a natural reservoir of *Vibrio cholerae* Non-O1 and Non-O139 in freshwater habitats. *Microbial. Ecol.* 47, 341–349. doi: 10.1007/s00248-003-2007-2006

Halpern, M., and Izhaki, I. (2017). Fish as hosts of *Vibrio cholerae*. *Front. Microbiol.* 8:282. doi: 10.3389/fmicb.2017.00282

Hood, M. A., and Winter, P. A. (1997). Attachment of *Vibrio cholerae* under various environmental conditions and to selected substrates. *FEMS Microbiol. Ecol.* 22, 215–223. doi: 10.1111/j.1574-6941.1997.tb00373.x

Hossain, Z. Z., Farhana, I., Tulsiani, S. M., Begum, A., and Jensen, P. K. M. (2018). Transmission and toxigenic potential of *Vibrio cholerae* in Hilsha Fish (*Tenualosa ilisha*) for human consumption in Bangladesh. *Front. Microbiol.* 9:222. doi: 10.3389/fmicb.2018.00222

Hounmanou, Y. M. G., Mdegela, R. H., Dougnon, T. V., Mhongole, O. J., Mayila, E. S., Malakalinga, J., et al. (2016). Toxigenic *Vibrio cholerae* O1 in vegetables and fish raised in wastewater irrigated fields and stabilization ponds during a non-cholera outbreak period in Morogoro, Tanzania: an environmental health study. *BMC Res. Notes* 9:466. doi: 10.1186/s13104-016-2283-2280

Huq, A., Small, E. B., West, P. A., Huq, M. I., Rahman, R., and Colwell, R. R. (1983). Ecological relationships between *Vibrio cholerae* and planktonic crustacean copepods. *Appl. Environ. Microbiol.* 45, 275–283.

Islam, M. S., Islam, M. S., Mahmud, Z. H., Cairncross, S., Clemens, J. D., and Collins, A. E. (2015). Role of phytoplankton in maintaining endemicity and seasonality of cholera in Bangladesh. *Trans. R. Soc. Trop. Med. Hyg.* 109, 572–578. doi: 10.1093/trstmh/trv057

Kamruzzaman, M., Udden, S. M. N., Cameron, D. E., Calderwood, S. B., Nair, G. B., Mekalanos, J. J., et al. (2010). Quorum-regulated biofilms enhance the development of conditionally viable, environmental *Vibrio cholerae*. *Proc. Natl. Acad. Sci.* 107, 1588–1593. doi: 10.1073/pnas.0913404107

Laviad -Shitrit, S., Lev-Ari, T., Katzir, G., Sharaby, Y., Izhaki, I., and Halpern, M. (2017). Great cormorants (*Phalacrocorax carbo*) as potential vectors for the dispersal of *Vibrio cholerae*. *Sci. Rep.* 7:7973. doi: 10.1038/s41598-017-08434-8438

Mengel, M. A., Delrieu, I., Heyerdahl, L., and Gessner, B. D. (2014). "Cholera outbreaks in africa," in *Cholera Outbreaks*, eds G. B. Nair and Y. Takeda (Berlin: Springer), 117–144. doi: 10.1007/82_2014_369

Mitchell, K. C., Breen, P., Britton, S., Neely, M. N., and Withey, J. H. (2017). Quantifying *Vibrio cholerae* enterotoxicity in a Zebrafish infection model. *Appl. Environ. Microbiol.* 83, e783–e717. doi: 10.1128/AEM.00783-717

Nag, D., Breen, P., Raychaudhuri, S., and Withey, J. H. (2018). Glucose metabolism by *Escherichia coli* inhibits *Vibrio cholerae* intestinal colonization of Zebrafish. *Infect. Immun.* 86, e486–e418. doi: 10.1128/IAI.00486-418

Nandi, B., Nandy, R. K., Mukhopadhyay, S., Nair, G. B., Shimada, T., and Ghose, A. C. (2000). Rapid method for species-specific identification of *Vibrio cholerae* using primers targeted to the gene of outer membrane protein OmpW. *J. Clin. Microbiol.* 38, 4145–4151.

Nkoko, D., Giraudoux, P., Plisnier, P.-D., Tinda, A., Piarroux, M., Sudre, B., et al. (2011). Dynamics of cholera outbreaks in great lakes region of Africa, 1978–2008. *Emerg. Infect. Dis.* 17, 2026–2034. doi: 10.3201/eid1711.110170

Nyambuli, S., Mhongole, O. J., Katakweba, A. A., Dalsgaard, A., and Mdegela, R. H. (2018). Prevalence, pathogenic markers and antibiotic susceptibility of *Vibrio cholerae* in Sardines, Water and phytoplankton in lake Tanganyika, Tanzania. *Int. J. Agri., Forest. Fish.* 6:29.

Patra, S. K., and Mohamed, K. S. (2003). Enrichment of *Artemia nauplii* with the probiotic yeast *Saccharomyces boulardii* and its resistance against a pathogenic Vibrio. *Aqua. Int.* 11, 505–514. doi: 10.1023/B:AQUI.0000004193.40039.54

Plisnier, P.-D., Poncelet, N., Cocquyt, C., De Boeck, H., Bompangue, D., Naithani, J., et al. (2015). *Cholera Outbreaks at Lake Tanganyika Induced by Climate Change?*. Brussels: Belgian Science Policy.

Rabia, A., Wambura, P., Misinzo, G., Kimera, S., Mdegela, R., Mzula, A., et al. (2017). Molecular Epidemiology of *Vibrio cholerae* recovered from sewage drains, captured Fish and humans in 2015/16 cholera outbreak in Zanzibar, Tanzania. *J. Adv. Microbiol.* 5, 1–11. doi: 10.9734/JAMB/2017/36036

Reyburn, R., Kim, D. R., Emch, M., Khatib, A., von Seidlein, L., and Ali, M. (2011). Climate variability and the outbreaks of cholera in Zanzibar, East Africa: a time series analysis. *Am. J. Trop. Med. Hyg.* 84, 862–869. doi: 10.4269/ajtmh.2011.10-0277

Runft, D. L., Mitchell, K. C., Abuaita, B. H., Allen, J. P., Bajer, S., Ginsburg, K., et al. (2014). Zebrafish as a natural host Model for *Vibrio cholerae* colonization and transmission. *Appl. Environ. Microbiol.* 80, 1710–1717. doi: 10.1128/AEM.03580-3513

Safa, A., Bhuyian, N. A., Nusrin, S., Ansaruzzaman, M., Alam, M., Hamabata, T., et al. (2006). Genetic characteristics of Matlab variants of *Vibrio cholerae* O1 that are hybrids between classical and El Tor biotypes. *J. Med. Microbiol.* 55, 1563–1569. doi: 10.1099/jmm.0.46689-46680

Sanchez, J., and Holmgren, J. (2011). Cholera toxin - a foe & a friend. *Indian J. Med. Res.* 133, 153–163.

Senderovich, Y., Izhaki, I., and Halpern, M. (2010). Fish as reservoirs and vectors of *Vibrio cholerae*. *PLoS One* 5:e8607. doi: 10.1371/journal.pone.0008607

Sengupta, C., Ekka, M., Arora, S., Dhaware, P. D., Chowdhury, R., and Raychaudhuri, S. (2017). Cross feeding of glucose metabolism byproducts of *Escherichia coli* human gut isolates and probiotic strains affect survival of *Vibrio cholerae*. *Gut Pathog.* 9:3. doi: 10.1186/s13099-016-0153-x

Tamplin, M. L., Gauzens, A. L., Huq, A., Sack, D. A., and Colwell, R. R. (1990). Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Appl. Environ. Microbiol.* 56, 1977–1980.

Urassa, W., Mhando, Y., Mhalu, F., and Mgonja, S. (2009). Antimicrobial susceptibility pattern of *Vibrio cholerae* 01 strains during two cholera outbreaks in Dar Es Salaam, Tanzania. *East Afr. Med. J.* 77, 350–353. doi: 10.4314/eamj.v77i7.46661

Weill, F.-X., Domman, D., Njamkepo, E., Tarr, C., Rauzier, J., Fawal, N., et al. (2017). Genomic history of the seventh pandemic of cholera in Africa. *Science* 358, 785–789. doi: 10.1126/science.aad 5901

Xu, T., Cao, H., Zhu, W., Wang, M., Du, Y., Yin, Z., et al. (2018). RNA-seq-based monitoring of gene expression changes of viable but non-culturable state of *Vibrio cholerae* induced by cold seawater. *Environ. Microbiol. Rep.* 10, 594–604. doi: 10.1111/1758-2229.12685

Yoon, S. S., and Mekalanos, J. J. (2006). 2,3-Butanediol Synthesis and the Emergence of the *Vibrio cholerae* El Tor Biotype. *Infect. Immun.* 74, 6547–6556. doi: 10.1128/IAI.00695-696

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Chapter 4. General discussion and limitations of the Thesis

4.1. Dynamics of *V. cholerae* O1 and cholera in Tanzania

The evolution of cholera epidemics for the past decades in Tanzania as described in **Manuscript A** and supported by a previous study on Zanzibar islands (Bi et al., 2018) indicates that the country can be considered a cholera endemic area, as they reported cholera cases almost every year since 1974 (Lessler et al., 2018). Tanzania is actually one of the 48 cholera endemic countries where cholera is still a serious public health concern and where the WHO aims to reduce deaths by 90% or eliminate the disease by 2030 (WHO, 2017). The geographical distribution of cases throughout the various outbreaks fluctuates significantly, although some high-risk areas were identified. Areas like the central regions and the Lake Zones of the country have experienced more outbreaks or higher magnitude outbreaks. These regions, especially the Dodoma capital region, had the highest cholera incidence in the past decade, and were detected in the hotspot analyses as high-risk areas requiring priority for interventions towards cholera control (**Manuscript A**). This points to the role of urbanization and population displacement on cholera incidence (Phelps et al., 2017, 2018; Sasaki et al., 2008). As described in **Manuscript A**, the increasing number of cholera cases between 2015 and 2017 in Dodoma coincided with the period where the Tanzanian government decided to move offices to Dodoma. This was therefore associated with movements of government employees and affiliated business from Dar es Salaam, an endemic cholera city, as well as a number of people working with the construction of new government office buildings (Lugongo, 2019). Increased risk of transmission of *V. cholerae* due to human movements has been previously reported in different parts of the world, like in the 2010 Haitian outbreak (Hendriksen et al., 2011) and in Yemen in 2016-2017 (Weill et al., 2019).

Between 1998 and 2012 for instance, cholera outbreaks were reported in different and distant regions of Tanzania but they were all caused by a single genotype *ctx*B1, of sequence type ST69

belonging to the T10 sub-lineage within the 7PET lineage of *V. cholerae* O1. This substantiates the implication of human movements as a driver for the spread of *V. cholerae* to cause epidemics across the country (**Manuscript C).** Similar situations are recorded for the countrywide outbreak that occurred between August 2015 and December 2017, where all regions were affected by an outbreak caused by a single genotype of *V. cholerae ctx*B7, sub-lineage T13. Human-mediated transmission of identical sub-lineages of *V. cholerae* is further observed, with evidence of cross-border spread of *V. cholerae* between countries of the AGLR, such as Tanzania, Kenya, Uganda, DRC, Rwanda, Burundi etc., that experienced cholera outbreaks caused by identical sub-lineages of *V. cholerae* O1 in the same period. Besides human transport, the spread of *V. cholerae* across countries or continents can also be attributed to other factors such as migratory water-birds that can move pandemic strains from one location to the next (Laviad - Shitrit et al., 2017; Laviad-Shitrit et al., 2019). This hypothesis is supported, for instance, by the fact that Kenya, Tanzania and Uganda that have Lake Victoria in common experienced cholera outbreaks during the same period with genetically identical strains that could be disseminated by water-birds. The increasing use of antimicrobials in cholera treatment regimens in the early 1990s led to a high prevalence of antimicrobial resistance in T5 and T10 sub-lineages causing outbreaks in Tanzania and its neighboring countries until 2012 (**Manuscript C**). These sub-lineages are mainly marked by the presence of Inc/AC plasmids in their genomes, carrying beta-lactam resistance genes (for T5 strains) and the SXT-ICE element, encoding resistance to sulfamethoxazole and trimethoprim in most T10 strains. The current T13 sub-lineage of 7PET, however, show limited antimicrobial resistance as a result of about 10-kb nucleotide deletions on the SXT-ICE. The fact that these genetically related strains with similar resistance pattern are involved in different epidemics across the region support our initial argument of cross-border transmission of cholera in East Africa, calling for collaborative control efforts. The nucleotide deletions on the CTX prophage encoding the main virulence factor, cholera enterotoxin, in some T10 and T13

strains in Tanzania (**Manuscript C**) prompts the need for further studies on their clinical relevance even though co-infection of *ctxA*-positive and *ctxA*-negative strains can occur in some patients (Domman et al., 2018).

Although seasonality is normally known to affect cholera incidence (Bi et al., 2018; Emch et al., 2008; Lemaitre et al., 2019), our findings in **Manuscript A** showed no positive correlation between rainfall and risk for cholera in Tanzania. It follows that cholera outbreaks in Tanzania may not really follow regular seasonal fluctuations. Furthermore, there was no genetic diversity between *V. cholerae* strains involved in outbreaks from dry and rainy seasons in the same year. Nevertheless, we observed two different sub-lineages causing cholera in Tanzania in 2015. The first, from January to May 2015, occurred in and around Kigoma refugee camps in Tanzania and were caused by T10 strains of sequence type ST515 likely originating from DRC and Zambia because of refugees fleeing into Tanzania from Burundi and DRC at the time (**Manuscript C**). Later in the same year, i.e. from August 2015, it was the T13 ST69 strains already circulating in the region that were involved and probably until now. Apart from anthropogenic activities, environmental factors also influence the dynamics of cholera epidemics.

4.2. Genomic and epidemiological evidence that recurrent cholera epidemics in Tanzania could be triggered by the aquatic environment

Epidemiological data in **Manuscript A** suggested that regions surrounding Lake Victoria, Lake Tanganyika and Lake Nyaza were cholera hotspots in Tanzania and this was supported by regression analyses showing that every 100 Km of water perimeter in a region increased the cholera incidence by 1.5%. Based on these findings, we concluded that living near lakes had an implication on increasing cholera incidence and such correlation has been reported in many cholera endemic settings including DRC (Bompangue et al., 2008; Islam et al., 2019; Nkoko et al., 2011). Moreover, findings in **Manuscript B**, where toxigenic *V. cholerae* O1 isolated from Lake Victoria were phylogenetically identical to those

causing cholera outbreaks in the AGLR, further illustrates the association between the aquatic environment and recurrence of cholera outbreaks. The genetic relatedness was observed both in the core genomes as well as in the accessory genomes of the clinical strains and the environmental *V. cholerae* O1 recovered from Lake Victoria (**Manuscript C**). The genetic content of our environmental strains, including within pathogenicity islands and antimicrobial resistance determinants, was similar to that of clinical outbreak strains, showing a strong connection between the two niches as opposed to previous studies, which suggested that environmental strains usually lack some pathogenic traits (Chun et al., 2009; Li et al., 2019). In **Manuscript D,** we further confirmed our hypothesis with the persistence of up to 10^3 cfu/ml of 7PET *V. cholerae* in water and fish gut for more than two weeks after infection. The rampant status of cholera epidemics in Tanzania and its neighboring countries is therefore not only facilitated by human movements but is also due to occurrence and persistence of pandemic lineages of *V. cholerae* O1 in the aquatic environment, including in fish, phytoplankton and water in Lake Victoria during both epidemic and non-epidemic periods (**Manuscript B**). These environmental reservoirs provide room for toxigenic strains to persist in the environment and maintain the endemicity of cholera in human populations through continuous exposure to new progenitors of toxigenic *V. cholerae* (Lutz et al., 2013). The outcomes from these genomic and epidemiological data can facilitate intervention strategies in Tanzania and East Africa since the potential reservoirs supporting the survival of *V. cholerae* in the environment are identified. To meet the objectives of the global roadmap to cholera control by 2030, findings from this study can help local and specific interventions in Tanzania and the countries surrounding the African Great Lakes. Since the implication of aquatic environments on cholera is proven, it is essential to describe how *V. cholerae* survive in the aquatic environments of Tanzania during and between epidemic periods.

4.3. Intrinsic and external factors supporting persistence of pandemic *V. cholerae* O1 in Lake Victoria and subsequent feeding of the epidemic cycle

The aquatic environment can be harsh for the survival of most bacteria but it is the normal habitat for non-pathogenic *V. cholerae* (non-O1, non-O139) (Alam et al., 2006; Islam et al., 2019). Toxigenic *V. cholerae* O1 of outbreak potential are however not commonly found in the aquatic environment especially during non-outbreaks periods (Faruque et al., 1998; Islam et al., 2019; Lutz et al., 2013). Recovery of such strains in this study (**Manuscript B**) more than a year after cholera has ceased in the Tanzanian basin of Lake Victoria suggested the existence of internal and external factors supporting their survival. All isolates recovered from the environment were of the 7PET lineage and it has been proposed that genes encoded in the seventh pandemic island may function in persistence in the aquatic environment (Chun et al., 2009). These authors further suggested that the absence of the seventh pandemic islands in the Classical biotype could explain its extinction and this is in line with our experimental findings in **Manuscript D** where the Classical biotypes were undetectable in fish guts and in aquarium water after one week. Survival and maintenance of pathogenicity in 7PET *V. cholerae* in the aquatic environment during and between outbreaks favoring resurgence of outbreaks in humans can be attributed to the presence of genetic elements that allow these strains to withstand nutrient depletion or other environmental stresses as well as interactions with aquatic reservoirs (Chun et al., 2009).

Strains recovered from the Lake present molecular adaptation machinery for attachment, survival and defense that supports their environmental survival. The adaptation machinery includes the magnesium and cobalt efflux protein (*Cor*C), the cobalt-zinc-cadmium resistance protein (*Czc*D) and the multidrug efflux pump component (*Mtr*F) conversed in their defense system (**Manuscript B**) that enable strains to persist in the Lake against toxic compounds and heavy metal residues (Gong et al., 2018; Kishe and Machiwa, 2003; Ogwok et al., 2009). Moreover, it is proposed that vibriophages are an integral part of

the natural drop of cholera epidemic curves when lytic phages from the environment and the gut begin to infect and kill outbreak strains (Faruque and Mekalanos, 2012). In order to persist in the aquatic environment and provide progenitors for resurgence of future outbreaks, environmental strains recovered in Lake Victoria do not contain in their chromosome the PICI-like elements, which offer a sequence-specific area for invading phage DNA (Seed et al., 2013). This phage resistance/tolerance phenotype can play a role in the environmental fitness of the strains found in Lake Victoria. The pandemic strains recovered in the Lake also contained the two-component response regulator proteins, histidine kinase, Vibrio Polysaccharides (VPS) biosynthesis proteins and autoinducers (AI-2 *Lux*P and *Lux*Q) involved in quorum sensing and biofilm formation that can serve as one of the internal factors for environmental survival (Bari et al., 2013; Kamruzzaman et al., 2010). Furthermore, as for many other *Vibrio* species, the strains recovered from Lake Victoria showed their ability to survive in the Lake in a viable but not culturable state without losing their toxin genes, as testified by the significantly higher recovery of *ctxA*-positive samples in our direct PCR method compared to the full-culture technique (**Manuscript B**).

With respect to external factors maintaining the persistence of toxigenic *V. cholerae* in Lake Victoria, we reported throughout the dry and rainy seasons very optimum physico-chemical parameters including temperature, pH and salinity, which are well-established conditions for the survival of toxigenic *V. cholerae* in the aquatic environment (Hounmanou et al., 2016). Moreover, the strains demonstrated their attachment to environmental reservoirs such as fish and plankton composed of more than 40 different taxa and belonging to three major phytoplankton groups, including green algae, cyanobacteria, and diatoms. These phytoplankton species have been described to serve as reservoirs for survival of *V. cholerae* in aquatic environments (Islam et al., 2019). Since the epidemiological importance of environmental reservoirs of *V. cholerae* is crucial, we needed to explore the under studied-reservoirs, which are fish that are directly involved in the food chain. Findings in **Manuscript D**

confirmed the hypothesis in **Manuscript B** that fish are not just contaminated by *V. cholerae* due to feeding on phytoplankton, but they can actually serve as reservoir hosts for environmental survival and subsequent transmission of 7PET *V. cholerae* in the aquatic environment. Knowing that fish-eating migratory birds can shed *V. cholerae* obtained from fish for more than 72 h across countries/continents (Laviad-Shitrit et al., 2019), the role of fish in the maintenance of cholera endemicity in Tanzania, which is until now underestimated, should be taken more seriously in interventions. On the other hand, *Oreochromis niloticus* (tilapia) and *Rastrineobola argentea* (carps) shown in this study as reservoirs of 7PET *V. cholerae* (**Manuscript B** and **D**) are popular fish species in the AGLR. Their involvement as external factors and reservoirs favoring persistence of toxigenic *V. cholerae* O1 in the region can not only help to tailor control measures in the lake zones but can also lead to serious economic consequences for people whose livelihood depend on the fisheries sector due to the ban that may follow this activity during outbreaks. It is reported that the stigma related to cholera leads to various economic losses from trade and tourism and this has caused significant under-reporting of cholera in many endemic countries (Lonappan et al., 2019); a situation that cannot help the effective control of cholera. The potential negative economic effects of cholera are also one of the reasons why some countries often report major outbreaks as “acute watery diarrhea” without stating which pathogen caused it.

4.4. Limitations of the Thesis and lessons learnt

The studies included in this Thesis present some limitations that do not impede the conclusions but can provide room for future research.

i. In **Manuscript A**, the weakness of the cholera surveillance system in Tanzania to provide disaggregated data available weekly and at district levels limited the study to identify specific hotspots to serve as data for action to support the roadmap for cholera control. This can be fixed by proposing detailed district-based reporting in the country. The study was however able to determine some useful hotspots that corroborates the genomic findings.

ii. There is a relatively lower incidence of cholera in Tanzania based on reported data and this could be attributed to the fact that reporting cholera has negative economic impacts on affected countries, leading to underreporting (**Manuscript A)**. To alleviate such situation, Tanzanian authorities need to be informed on the importance of accurate and exhaustive reporting for effective control of the disease.

iii. The actual sub-lineages of four environmental 7PET *V. cholerae* strains in **Manuscript B** (P1, F1, F3 and W2) were not accurately identified until further analysis in **Manuscript C** due to limitations of the initially used bioinformatics pipelines. A new harmonized and updated database specific for *V. cholerae* called CholeraeFinder is being constructed in conjunction with the Centre for Genomic Epidemiology at the Technical University of Denmark (DTU) to allow reproducible and accurate analyses of *V. cholerae* genomes.

iv. It could have been interesting to perform whole genome sequencing on DNA directly obtained from enriched samples representing viable but not culturable strains, but the quality of such a complex DNA specimen may not be acceptable for whole genome sequencing as it might contain

Eukaryotic and Prokaryotic genomes and will not be purely and solely *V. cholerae* (**Manuscript B**).

v. Despite the evidence of phylogenetic relatedness between clinical and environmental isolates of *V. cholerae* O1 in Tanzania (**Manuscripts B** and **C**), the current data still cannot indicate the direction of pathogen transfer and the original source. The study, however, identified potential environmental reservoirs that harbor pandemic strains favoring continuous human exposure to the pathogen and maintenance of the rampant status of cholera in the study area.

vi. The currently available methods and knowledge, including those generated in this thesis, are not able to predict the next sub-lineages of *V. cholerae* that can emerge in future epidemics and their clinical relevance (**Manuscripts B** and **C)**. The clinical relevance and epidemiological importance of the currently circulating sub-lineages are however, well described in the Thesis to facilitate interventions.

vii. Despite the presence of some clinical strains with deletions on the CTX prophage, the current data do not indicate the virulence potential of such strains (**Manuscript C**). We however reported that these deletions are not monophyletic and strains harboring them could be recombined, hence more strains with the *ctx* deletion could occur.

viii. In the experimental studies, although fish remained colonized with about 10^3 cfu/intestine of *V. cholerae* until the end of two weeks, the initial infection dose used to ensure uptake of the test strains was higher than the naturally occurring levels of *V. cholerae* in the aquatic environment (**Manuscript D**). We however proposed that further studies need to be conducted with a lower infection dose to confirm the findings.

ix. We provided evidence of horizontal transfer of *V. cholerae* between fish, but it remains to demonstrate the possibilities of vertical transmission through fish eggs and subsequent dissemination by migratory fish-eating birds found around the AGLR (**Manuscript D**).

Chapter 5. Conclusion, recommendations and future research needs

5.1. Conclusion

In sum, the studies carried out during this PhD demonstrated that cholera is endemic in Tanzania because of the recurrent and rampant outbreaks in the Eastern African region supported by existence of environmental reservoirs for pandemic *V. cholerae* in Lake Victoria favoring persistence and resurgence of the pathogen throughout the year. The dynamics of cholera in Tanzania is not always seasonal but mostly depends on anthropogenic activities mainly human movements within the AGLR due to the cross-border transmission of the pathogen. We also reported that cholera high-risk populations in Tanzania include those living near lakes and central regions, being areas where intervention should be prioritized in order to control further spread within the country. We documented that cholera outbreaks in Tanzania for the past three decades are caused by the single lineage 7PET *V. cholerae* O1 occurring in three time-separated sub-linages mainly T5, T10 and T13 over time with introductions of new sequence types from neighboring countries. Antimicrobial resistance was surprisingly not found to be a big problem in the studied strains, especially not among the T13 sub-lineages causing recent epidemics.

Moreover, the studies provide scientific evidence based on genomic and epidemiological data that recurrent cholera epidemics in Tanzania are triggered by the aquatic environment as testified by the genome-wide close phylogenetic relationship between environmental strains recovered from Lake Victoria and clinical strains causing cholera outbreaks in Tanzania and its neighboring countries. Existence of waterbodies in a region was also found to increase the cholera incidence in that region.

Several intrinsic and external factors were pointed out as potentially supporting the persistence of the pandemic strains in the Lake. These include the presence of rampant epidemic foci in the AGLR leading to continuous discharge of epidemic clones in the water and the existence of genes encoding the

seventh pandemic islands in these 7PET strains favoring their survival in aquatic niches. Resistance to toxic compounds and heavy metals from the Lake as well as mechanisms for biofilm formation and phage resistance phenotypes were also identified as factors supporting the persistence of strains in the aquatic environment. On the other hand, the Lake provides optimum growth conditions for *V. cholerae* with its tropical temperature and alkaline pH. The presence of various taxa of phytoplankton in the Lake offering symbiotic life to the *V. cholerae* strains contributed to their survival in the Lake even during non-outbreak periods. Last but not least, Lake Victoria and possibly other Lakes in the AGLR represent underestimated reservoirs for toxigenic *V. cholerae* via the various fish species that they contain because tilapia is demonstrated in this study to play an essential role in the survival and transmission of 7PET *V. cholerae* O1 in aquatic environments.

5.2 Recommendations

Based on the findings of this project we propose the following:

i. A policy brief to health authorities in Tanzania about the hotspot areas identified in this study where priority should be directed for effective control measures against cholera in the country
ii. Routine sanitation and hygiene awareness campaigns in all areas, notably around the Lake zones where small but perpetual epidemic foci exist and feed the aquatic environment with toxigenic strains capable to persist in various reservoirs and reemerge
iii. Multi-sectoral and cross-border collaboration within the Eastern African nations, which currently represent the main cholera hotspot on the continent In view of the cross-border transmission of the 7PET *V. cholerae* in the AGLR and the presence of reservoirs in the Lakes.
iv. The initiation of vaccination in neighboring countries whenever one declares an epidemic since the same sub-lineages are found to be causing outbreaks in countries around the AGLR at the same time
v. The encouragement of proper reporting of cholera cases by health authorities at district levels to strengthen the cholera surveillance system in Tanzania and allow more specific analyses of the data to guide decision on effective control despite the negative connotation of cholera reporting
vi. Increased efforts by the WHO towards cholera control and eradication not only in Africa, but also in Asia where most strains originate, given that all strains causing cholera in Tanzania are from the same 7PET lineage emanating from Asia.

5.3. Future research needs

a. Future research in Tanzania should focus on the clinical relevance and toxin production ability of *V. cholerae* strains presenting nucleotide deletions on the *ctx*A.

b. Genomic surveillance studies should be carried out to determine the epidemiological importance of non-O1/non-O139 strains, which could be undermined sources of cholera.

c. Studies on fish as a reservoir for *V. cholerae* in the aquatic environment need further experiments with lower initial infectious doses closer to natural levels in the aquatic environment. Experiments should also be performed to assess the possibilities of vertical transfer of *V. cholerae* from fish to their offspring.

d. In the AGLR there are various species of migratory birds that feed on fish and other aquatic organisms and the role of these birds in the dissemination of cholera across the region needs to be assessed using genomics and experimental tools.

e. The role of domestic animals in the maintenance of toxigenic *V. cholerae* O1 within households during and between epidemics needs to be investigated.

f. Since the human-mediated transmission of *V. cholerae* seems very prominent in the AGLR, future quantitative and qualitative microbiome studies should focus on the possibilities of *V. cholerae* carriage among healthy human populations during inter-epidemic periods.

g. Like the existing cholera outbreak prediction models, there is a need to propose new models that can be based on the molecular-clock of *V. cholerae* to predict up-coming genotypes of *V. cholerae*, their virulence and epidemiological importance.

References

Alam, M., Sultana, M., Nair, G. B., Sack, R. B., Sack, D. A., Siddique, A. K., et al. (2006). Toxigenic *Vibrio cholerae* in the aquatic environment of Mathbaria, Bangladesh. *Applied and Environmental Microbiology* 72, 2849–2855. doi:10.1128/AEM.72.4.2849-2855.2006.

Ali, M., Nelson, A. R., Lopez, A. L., and Sack, D. A. (2015). Updated global burden of cholera in endemic countries. *PLOS Neglected Tropical Diseases* 9, e0003832. doi:10.1371/journal.pntd.0003832.

Aliabad, N. H., Bakhshi, B., Pourshafie, M. R., Sharifnia, A., and Ghorbani, M. (2012). Molecular diversity of CTX prophage in *Vibrio cholerae*: CTX prophage in *Vibrio cholerae*. *Letters in Applied Microbiology* 55, 27–32. doi:10.1111/j.1472-765X.2012.03253.x.

Armitage, P. D., Pinder, L. C., and Cranston, P. eds. (1995). *The Chironomidae: Biology and ecology of non-biting midges*. Springer Netherlands Available at: https://www.springer.com/gp/book/9780412452604 [Accessed July 22, 2019].

Azman, A. S., Moore, S. M., and Lessler, J. (2019). Surveillance and the global fight against cholera: Setting priorities and tracking progress. *Vaccine*. doi:10.1016/j.vaccine.2019.06.037.

Azurin, J. C., Kobari, K., Barua, D., Alvero, M., Gomez, C. Z., Dizon, J. J., et al. (1967). A long-term carrier of cholera: Cholera Dolores. *Bulletin of the World Health Organization* 37, 745–749. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2554929/ [Accessed July 23, 2019].

Bakhshi, B., Pourshafie, M. R., Navabakbar, F., and Tavakoli, A. (2008). Genomic organisation of the CTX element among toxigenic *Vibrio cholerae* isolates. *Clinical Microbiology and Infection* 14, 562–568. doi:10.1111/j.1469-0691.2008.01976.x.

Bari, S. M. N., Roky, M. K., Mohiuddin, M., Kamruzzaman, M., Mekalanos, J. J., and Faruque, S. M. (2013). Quorum-sensing autoinducers resuscitate dormant *Vibrio cholerae* in environmental water samples. *Proceedings of the National Academy of Sciences* 110, 9926–9931. doi:10.1073/pnas.1307697110.

Bayliss, S. C., Verner-Jeffreys, D. W., Bartie, K. L., Aanensen, D. M., Sheppard, S. K., Adams, A., et al. (2017). The Promise of whole genome pathogen sequencing for the molecular epidemiology of emerging aquaculture pathogens. *Frontiers in Microbiology* 8. doi:10.3389/fmicb.2017.00121.

Beyhan, S., Tischler, A. D., Camilli, A., and Yildiz, F. H. (2006). differences in gene expression between the Classical and El Tor biotypes of *Vibrio cholerae* O1. *Infection and Immunity* 74, 3633–3642. doi:10.1128/IAI.01750-05.

Bi, Q., Abdalla, F. M., Masauni, S., Reyburn, R., Msambazi, M., Deglise, C., et al. (2018). The epidemiology of cholera in Zanzibar: implications for the Zanzibar comprehensive cholera elimination plan. *Journal of Infectious Diseases*. doi:10.1093/infdis/jiy500.

Bompangue, D., Giraudoux, P., Handschumacher, P., Piarroux, M., Sudre, B., Ekwanzala, M., et al. (2008). Lakes as source of cholera outbreaks, Democratic Republic of Congo. *Emerging Infectious Diseases* 14, 798–800. doi:10.3201/eid1405.071260.

Broza, M., Gancz, H., Halpern, M., and Kashi, Y. (2005). Adult non-biting midges: possible windborne carriers of *Vibrio cholerae* non-O1 non-O139. *Environmental Microbiology* 7, 576–585. doi:10.1111/j.1462-2920.2005.00745.x.

Brumfield, K. D., Carignan, B. M., and Son, M. S. (2018). "Genotypic and phenotypic assays to distinguish *Vibrio cholerae* biotype," in *Vibrio Cholerae* Methods in Molecular Biology. (Humana Press, New York, NY), 11–28. doi:10.1007/978-1-4939-8685-9_2.

Bwire, G., Debes, A. K., Orach, C. G., Kagirita, A., Ram, M., Komakech, H., et al. (2018). Environmental surveillance of *Vibrio cholerae* O1/O139 in the five African Great Lakes and other major surface water sources in Uganda. *Frontiers in Microbiology* 9. doi:10.3389/fmicb.2018.01560.

Chin, C.-S., Sorenson, J., Harris, J. B., Robins, W. P., Charles, R. C., Jean-Charles, R. R., et al. (2011). The origin of the Haitian cholera outbreak strain. *New England Journal of Medicine* 364, 33–42. doi:10.1056/NEJMoa1012928.

Chun, J., Grim, C. J., Hasan, N. A., Lee, J. H., Choi, S. Y., Haley, B. J., et al. (2009). Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. *PNAS* 106, 15442–15447. doi:10.1073/pnas.0907787106.

Colwell, R. R., Huq, A., Islam, M. S., Aziz, K. M. A., Yunus, M., Khan, N. H., et al. (2003). Reduction of cholera in Bangladeshi villages by simple filtration. *Proc Natl Acad Sci U S A* 100, 1051–1055. doi:10.1073/pnas.0237386100.

Colwell, R. R., Kaper, J., and Joseph, S. W. (1977). *Vibrio cholerae*, *Vibrio parahaemolyticus*, and other vibrios: occurrence and distribution in Chesapeake Bay. *Science* 198, 394–396.

Constantin de Magny, G., and Colwell, R. R. * (2009). Cholera and climate: A demonstrated relationship. *Transactions of the American Clinical and Climatological Association* 120, 119–128. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2744514/ [Accessed July 27, 2019].

Dalsgaard, A., Serichantalergs, O., Forslund, A., Lin, W., Mekalanos, J., Mintz, E., et al. (2001). Clinical and environmental isolates of *Vibrio cholerae* serogroup O141 carry the CTX phage and the genes encoding the Toxin-Coregulated Pili. *Journal of Clinical Microbiology* 39, 4086–4092. doi:10.1128/JCM.39.11.4086-4092.2001.

Dalusi, L., Lyimo, T. J., Lugomela, C., Hosea, K. M. M., and Sjöling, S. (2015). Toxigenic *Vibrio cholerae* identified in estuaries of Tanzania using PCR techniques. *FEMS Microbiology Letters* 362. doi:10.1093/femsle/fnv009.

Devault, A. M., Golding, G. B., Waglechner, N., Enk, J. M., Kuch, M., Tien, J. H., et al. (2014). Second-pandemic strain of *Vibrio cholerae* from the Philadelphia cholera outbreak of 1849. *New England Journal of Medicine* 370, 334–340. doi:10.1056/NEJMoa1308663.

Domman, D., Chowdhury, F., Khan, A. I., Dorman, M. J., Mutreja, A., Uddin, M. I., et al. (2018). Defining endemic cholera at three levels of spatiotemporal resolution within Bangladesh. *Nature Genetics* 50, 951–955. doi:10.1038/s41588-018-0150-8.

Domman, D., Quilici, M.-L., Dorman, M. J., Njamkepo, E., Mutreja, A., Mather, A. E., et al. (2017). Integrated view of *Vibrio cholerae* in the Americas. *Science* 358, 789–793. doi:10.1126/science.aao2136.

Dutilh, B. E., Thompson, C. C., Vicente, A. C., Marin, M. A., Lee, C., Silva, G. G., et al. (2014). Comparative genomics of 274 *Vibrio cholerae* genomes reveals mobile functions structuring three niche dimensions. *BMC Genomics* 15. doi:10.1186/1471-2164-15-654.

Dziejman, M., Serruto, D., Tam, V. C., Sturtevant, D., Diraphat, P., Faruque, S. M., et al. (2005). Genomic characterization of non-O1, non-O139 *Vibrio cholerae* reveals genes for a type III secretion system. *PNAS* 102, 3465–3470. doi:10.1073/pnas.0409918102.

Echenberg, M. J. (2011). *Africa in the time of cholera: a history of pandemics from 1817 to the present*. New York: Cambridge University Press.

Emch, M., Feldacker, C., Islam, M. S., and Ali, M. (2008). Seasonality of cholera from 1974 to 2005: a review of global patterns. *International Journal of Health Geographics* 7, 31. doi:10.1186/1476-072X-7-31.

Fang, L., Ginn, A. M., Harper, J., Kane, A. S., and Wright, A. C. (2019). Survey and genetic characterization of *Vibrio cholerae* in Apalachicola Bay, Florida (2012-2014). *Journal of Applied Microbiology* 126, 1265–1277. doi:10.1111/jam.14199.

Faruque, S. M., Albert, M. J., and Mekalanos, J. J. (1998). Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiology and. Molecular Biology Reviews* 62, 1301–1314.

Faruque, S. M., Chowdhury, N., Kamruzzaman, M., Ahmad, Q. S., Faruque, A. S. G., Salam, M. A., et al. (2003). Reemergence of epidemic *Vibrio cholerae* O139, Bangladesh. *Emerging Infectious Diseases* 9, 1116–1122. doi:10.3201/eid0909.020443.

Faruque, S. M., and Mekalanos, J. J. (2012). Phage-bacterial interactions in the evolution of toxigenic *Vibrio cholerae*. *Virulence* 3, 556–565. doi:10.4161/viru.22351.

Ghosh-Banerjee, J., Senoh, M., Takahashi, T., Hamabata, T., Barman, S., Koley, H., et al. (2010). Cholera toxin production by the El Tor Variant of *Vibrio cholerae* O1 compared to prototype El Tor and Classical biotypes. *Journal of Clinical Microbiology* 48, 4283–4286. doi:10.1128/JCM.00799-10.

Gong, L., Yu, P., Zheng, H., Gu, W., He, W., Tang, Y., et al. (2018). Comparative genomics for non-O1/O139 *Vibrio cholerae* isolates recovered from the Yangtze river estuary versus *V. cholerae* representative isolates from serogroup O1. *Molecular Genetics Genomics* doi:10.1007/s00438-018-1514-6.

Grad, Y. H., and Waldor, M. K. (2013). Deciphering the origins and tracking the evolution of cholera epidemics with whole-genome-based molecular epidemiology. *mBio* 4, e00670-13. doi:10.1128/mBio.00670-13.

Griffith, D. C., Kelly-Hope, L. A., and Miller, M. A. (2006). Review of reported cholera outbreaks worldwide, 1995-2005. *American Journal of Tropical Medicine and Hygiene* 75, 973–977.

Haley, B. J., Choi, S. Y., Grim, C. J., Onifade, T. J., Cinar, H. N., Tall, B. D., et al. (2014). Genomic and phenotypic characterization of *Vibrio cholerae* non-O1 isolates from a US Gulf Coast cholera outbreak. *PLoS ONE* 9, e86264. doi:10.1371/journal.pone.0086264.

Halpern, M., Broza, Y. B., Mittler, S., Arakawa, E., and Broza, M. (2004). Chironomid egg masses as a natural reservoir of *Vibrio cholerae* non-O1 and non-O139 in freshwater habitats. *Microbial Ecology* 47. doi:10.1007/s00248-003-2007-6.

Halpern, M., and Izhaki, I. (2017). Fish as hosts of *Vibrio cholerae*. *Frontiers in Microbiology* 8. doi:10.3389/fmicb.2017.00282.

Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., et al. (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* 406, 477–483. doi:10.1038/35020000.

Hendriksen, R. S., Bortolaia, V., Tate, H., Tyson, G. H., Aarestrup, F. M., and McDermott, P. F. (2019). Using genomics to track global antimicrobial resistance. *Frontiers in Public Health* 7. doi:10.3389/fpubh.2019.00242.

Hendriksen, R. S., Price, L. B., Schupp, J. M., Gillece, J. D., Kaas, R. S., Engelthaler, D. M., et al. (2011). Population genetics of *Vibrio cholerae* from Nepal in 2010: Evidence on the origin of the Haitian outbreak. *mBio* 2. doi:10.1128/mBio.00157-11.

Hossain, Z. Z., Farhana, I., Tulsiani, S. M., Begum, A., and Jensen, P. K. M. (2018). Transmission and toxigenic potential of *Vibrio cholerae* in Hilsha fish (*Tenualosa ilisha*) for human consumption in Bangladesh. *Frontiers in Microbiology* 9. doi:10.3389/fmicb.2018.00222.

Hounmanou, Y. M. G., Leekitcharoenphon, P., Hendriksen, R. S., Dougnon, T. V., Mdegela, R. H., Olsen, J. E., et al. (2019a). Surveillance and genomics of toxigenic *Vibrio cholerae* O1 from fish, phytoplankton and water in Lake Victoria, Tanzania. *Frontiers in Microbiology* 10. doi:10.3389/fmicb.2019.00901.

Hounmanou, Y. M. G., Mdegela, R. H., Dougnon, T. V., Madsen, H., Withey, J. H., Olsen, J. E., et al. (2019b). Tilapia (*Oreochromis niloticus*) as a putative reservoir host for survival and transmission

of *Vibrio cholerae* O1 biotype El Tor in the aquatic environment. *Frontiers in Microbiology* 10. doi:10.3389/fmicb.2019.01215.

Hounmanou, Y. M. G., Mdegela, R. H., Dougnon, T. V., Mhongole, O. J., Mayila, E. S., Malakalinga, J., et al. (2016). Toxigenic *Vibrio cholerae* O1 in vegetables and fish raised in wastewater irrigated fields and stabilization ponds during a non-cholera outbreak period in Morogoro, Tanzania: an environmental health study. *BMC Research Notes* 9. doi:10.1186/s13104-016-2283-0.

Huq, A., West, P. A., Small, E. B., Huq, M. I., and Colwell, R. R. (1984). Influence of water temperature, salinity, and pH on survival and growth of toxigenic *Vibrio cholerae* serovar 01 associated with live copepods in laboratory microcosms. *Applied and Environmental Microbiology* 48, 420–424.

Ingelbeen, B., Hendrickx, D., Miwanda, B., van der Sande, M. A. B., Mossoko, M., Vochten, H., et al. (2019). Recurrent cholera outbreaks, Democratic Republic of the Congo, 2008–2017. *Emerging Infectious Diseases* 25, 856–864. doi:10.3201/eid2505.181141.

Islam, M. S., Islam, M. S., Mahmud, Z. H., Cairncross, S., Clemens, J. D., and Collins, A. E. (2015). Role of phytoplankton in maintaining endemicity and seasonality of cholera in Bangladesh. *Transactions of The Royal Society of Tropical Medicine and Hygiene* 109, 572–578. doi:10.1093/trstmh/trv057.

Islam, M. S., Zaman, M. H., Islam, M. S., Ahmed, N., and Clemens, J. D. (2019). Environmental reservoirs of *Vibrio cholerae*. *Vaccine*. doi:10.1016/j.vaccine.2019.06.033.

Kachwamba, Y., Mohammed, A. A., Lukupulo, H., Urio, L., Majigo, M., Mosha, F., et al. (2017). Genetic characterization of *Vibrio cholerae* O1 isolates from outbreaks between 2011 and 2015 in Tanzania. *BMC Infectious Diseases* 17. doi:10.1186/s12879-017-2252-9.

Kamruzzaman, M., Udden, S. M. N., Cameron, D. E., Calderwood, S. B., Nair, G. B., Mekalanos, J. J., et al. (2010). Quorum-regulated biofilms enhance the development of conditionally viable, environmental *Vibrio cholerae*. *Proceedings of the National Academy of Sciences* 107, 1588–1593. doi:10.1073/pnas.0913404107.

Katz, L. S., Petkau, A., Beaulaurier, J., Tyler, S., Antonova, E. S., Turnsek, M. A., et al. (2013). Evolutionary dynamics of *Vibrio cholerae* O1 following a single-source introduction to Haiti. *mBio* 4, e00398-13. doi:10.1128/mBio.00398-13.

Kim, E. J., Lee, C. H., Nair, G. B., and Kim, D. W. (2015). Whole-genome sequence comparisons reveal the evolution of *Vibrio cholerae* O1. *Trends in Microbiology* 23, 479–489. doi:10.1016/j.tim.2015.03.010.

Kim, E. J., Lee, D., Moon, S. H., Lee, C. H., Kim, S. J., Lee, J. H., et al. (2014). Molecular insights into the evolutionary pathway of *Vibrio cholerae* O1 atypical El Tor variants. *PLoS Pathogens* 10, e1004384. doi:10.1371/journal.ppat.1004384.

Kishe, M. A., and Machiwa, J. F. (2003). Distribution of heavy metals in sediments of Mwanza Gulf of Lake Victoria, Tanzania. *Environment International* 28, 619–625. doi:10.1016/S0160-4120(02)00099-5.

Laviad -Shitrit, S., Lev-Ari, T., Katzir, G., Sharaby, Y., Izhaki, I., and Halpern, M. (2017). Great cormorants (*Phalacrocorax carbo*) as potential vectors for the dispersal of *Vibrio cholerae*. *Scientific Reports* 7. doi:10.1038/s41598-017-08434-8.

Laviad-Shitrit, S., Izhaki, I., and Halpern, M. (2019). Accumulating evidence suggests that some waterbird species are potential vectors of *Vibrio cholerae*. *PLOS Pathogens* 15, e1007814. doi:10.1371/journal.ppat.1007814.

Lemaitre, J., Pasetto, D., Perez-Saez, J., Sciarra, C., Wamala, J. F., and Rinaldo, A. (2019). Rainfall as a driver of epidemic cholera: Comparative model assessments of the effect of intra-seasonal precipitation events. *Acta Tropica* 190, 235–243. doi:10.1016/j.actatropica.2018.11.013.

Lessler, J., Moore, S. M., Luquero, F. J., McKay, H. S., Grais, R., Henkens, M., et al. (2018). Mapping the burden of cholera in sub-Saharan Africa and implications for control: an analysis of data across geographical scales. *The Lancet* 391, 1908–1915. doi:10.1016/S0140-6736(17)33050-7.

Li, Z., Pang, B., Wang, D., Li, J., Xu, J., Fang, Y., et al. (2019). Expanding dynamics of the virulence-related gene variations in the toxigenic *Vibrio cholerae* serogroup O1. *BMC Genomics* 20. doi:10.1186/s12864-019-5725-y.

Lipp, E. K., Huq, A., and Colwell, R. R. (2002). Effects of global climate on infectious disease: the cholera model. *Clinical Microbiology Reviews* 15, 757–770. doi:10.1128/CMR.15.4.757-770.2002.

Lonappan, S., Golecha, R., and Balakrish Nair, G. (2019). Contrasts, contradictions and control of cholera. *Vaccine*. doi:10.1016/j.vaccine.2019.08.022.

Lugongo, B. (2019). Tanzania: government move to Dodoma now at 86 per cent. *Tanzania Daily News (Dar es Salaam).* Available at: https://allafrica.com/stories/201902060401.html [Accessed July 8, 2019].

Lutz, C., Erken, M., Noorian, P., Sun, S., and McDougald, D. (2013). Environmental reservoirs and mechanisms of persistence of *Vibrio cholerae*. *Frontiers in Microbiology* 4. doi:10.3389/fmicb.2013.00375.

Mbwette, T. S. (1987). Cholera outbreaks in Tanzania. *Journal of the Royal Society of Health* 107, 134–136.

Mengel, M. A., Delrieu, I., Heyerdahl, L., and Gessner, B. D. (2014). "Cholera outbreaks in Africa," in *Cholera Outbreaks*, eds. G. B. Nair and Y. Takeda (Berlin, Heidelberg: Springer Berlin Heidelberg), 117–144. doi:10.1007/82_2014_369.

Mitchell, K. C., Breen, P., Britton, S., Neely, M. N., and Withey, J. H. (2017). Quantifying *Vibrio cholerae* enterotoxicity in a Zebrafish infection model. *Applied and Environmental Microbiology* 83, e00783-17. doi:10.1128/AEM.00783-17.

Mohammadi barzelighi, H., Isfahan, IR Iran, Bakhshi, B., Tehran, IR Iran, Boustanshenas, M., and Tehran, IR Iran (2016). Genetic determinants differences between *Vibrio cholerae* biotypes. *Infection, Epidemiology and Medicine* 2, 26–30. doi:10.18869/modares.iem.2.2.26.

Moore, S., Dongdem, A. Z., Opare, D., Cottavoz, P., Fookes, M., Sadji, A. Y., et al. (2018). Dynamics of cholera epidemics from Benin to Mauritania. *PLOS Neglected Tropical Diseases* 12, e0006379. doi:10.1371/journal.pntd.0006379.

Moore, S., Miwanda, B., Sadji, A. Y., Thefenne, H., Jeddi, F., Rebaudet, S., et al. (2015). Relationship between distinct African cholera epidemics revealed via MLVA haplotyping of 337 *Vibrio cholerae* isolates. *PLOS Neglected Tropical Diseases* 9, e0003817. doi:10.1371/journal.pntd.0003817.

Mutreja, A., and Dougan, G. (2019). Molecular epidemiology and intercontinental spread of cholera. *Vaccine*. doi:10.1016/j.vaccine.2019.07.038.

Mutreja, A., Kim, D. W., Thomson, N., Connor, T. R., Lee, J. H., Kariuki, S., et al. (2011). Evidence for multiple waves of global transmission within the seventh cholera pandemic. *Nature* 477, 462–465. doi:10.1038/nature10392.

Naha, A., Chowdhury, G., Ghosh-Banerjee, J., Senoh, M., Takahashi, T., Ley, B., et al. (2013). Molecular characterization of high-level-cholera-toxin-producing El Tor variant *Vibrio cholerae* strains in the Zanzibar archipelago of Tanzania. *Journal of Clinical Microbiology* 51, 1040–1045. doi:10.1128/JCM.03162-12.

Naha, A., Pazhani, G. P., Ganguly, M., Ghosh, S., Ramamurthy, T., Nandy, R. K., et al. (2012). Development and evaluation of a PCR assay for tracking the emergence and dissemination of Haitian variant *ctx*B in *Vibrio cholerae* O1 strains isolated from Kolkata, India. *Journal of Clinical Microbiology* 50, 1733–1736. doi:10.1128/JCM.00387-12.

Narendrakumar, L., Gupta, S. S., Johnson, J. B., Ramamurthy, T., and Thomas, S. (2019). Molecular adaptations and antibiotic resistance in *Vibrio cholerae*: A Communal Challenge. *Microbial Drug Resistance* 25, 1012–1022. doi:10.1089/mdr.2018.0354.

NBS (2012). Tanzania Census 2012 - National Bureau of Statistics. Available at: http://dataforall.org/dashboard/tanzania/ [Accessed October 29, 2018].

Nelson, E. J., Harris, J. B., Glenn Morris, J., Calderwood, S. B., and Camilli, A. (2009). Cholera transmission: the host, pathogen and bacteriophage dynamic. *Nature Reviews Microbiology* 7, 693–702. doi:10.1038/nrmicro2204.

Nkoko, D., Giraudoux, P., Plisnier, P.-D., Tinda, A., Piarroux, M., Sudre, B., et al. (2011). Dynamics of cholera outbreaks in Great Lakes Region of Africa, 1978–2008. *Emerging Infectious Diseases* 17. doi:10.3201/eid1711.110170.

Ogwok, P., Muyonga, J. H., and Sserunjogi, M. L. (2009). Pesticide residues and heavy metals in Lake Victoria Nile Perch, *Lates niloticus*, belly flap oil. *Bulletin Environmental Contamination and Toxicology* 82, 529–533. doi:10.1007/s00128-009-9668-x.

Olago, D., Marshall, M., Wandiga, S. O., Opondo, M., Yanda, P. Z., Kanalawe, R., et al. (2007). Climatic, socio-economic, and health factors affecting human vulnerability to cholera in the Lake Victoria basin, East Africa. *Ambio* 36, 350–358.

Pal, B. B., Khuntia, H. K., Nayak, S. R., Mohanty, A., and Biswal, B. (2017). *Vibrio cholerae* O1 Ogawa strains carrying the *ctx*B7 Allele caused a large cholera outbreak during 2014 in the tribal areas of Odisha, India. *Japanese Journal of Infectious Diseases* 70, 549–553. doi:10.7883/yoken.JJID.2016.585.

Paz, S. (2009). Impact of temperature variability on cholera incidence in southeastern Africa, 1971-2006. *Ecohealth* 6, 340–345. doi:10.1007/s10393-009-0264-7.

Phelps, M. D., Azman, A. S., Lewnard, J. A., Antillón, M., Simonsen, L., Andreasen, V., et al. (2017). The importance of thinking beyond the water-supply in cholera epidemics: A historical urban case-study. *PLOS Neglected Tropical Diseases* 11, e0006103. doi:10.1371/journal.pntd.0006103.

Phelps, M., Perner, M. L., Pitzer, V. E., Andreasen, V., Jensen, P. K. M., and Simonsen, L. (2018). Cholera epidemics of the past offer new insights into an old enemy. *The Journal of Infectious Diseases* 217, 641–649. doi:10.1093/infdis/jix602.

Plisnier, P.-D., Poncelet, N., Cocquyt, C., De Boeck, H., Bompangue, D., Naithani, J., et al. (2015). *Cholera outbreaks at Lake Tanganyika induced by Climate Change?*.

Rabia, A., Wambura, P., Misinzo, G., Kimera, S., Mdegela, R., Mzula, A., et al. (2017). Molecular epidemiology of *Vibrio cholerae* recovered from sewage drains, captured fish and humans in 2015/16 cholera outbreak in Zanzibar, Tanzania. *Journal of Advances in Microbiology* 5, 1–11. doi:10.9734/JAMB/2017/36036.

Rahaman, M. H., Islam, T., Colwell, R. R., and Alam, M. (2015). Molecular tools in understanding the evolution of *Vibrio cholerae*. *Frontiers in Microbiology* 6. doi:10.3389/fmicb.2015.01040.

Rajasingham, A., Hardy, C., Kamwaga, S., Sebunya, K., Massa, K., Mulungu, J., et al. (2019). Evaluation of an emergency bulk chlorination project targeting drinking water vendors in cholera-affected wards of Dar es Salaam and Morogoro, Tanzania. *American Journal of Tropical Medicine and. Hygiene* doi:10.4269/ajtmh.18-0734.

Ramamurthy, T., Mutreja, A., Weill, F.-X., Das, B., Ghosh, A., and Nair, G. B. (2019). Revisiting the global epidemiology of cholera in conjunction with the genomics of *Vibrio cholerae*. *Frontiers in Public Health* 7. doi:10.3389/fpubh.2019.00203.

Rashed, S. M., Iqbal, A., Mannan, S. B., Islam, T., Rashid, M., Johura, F., et al. (2013). *Vibrio cholerae* O1 El Tor and O139 Bengal strains carrying *ctx*BET, Bangladesh. *Emerging Infectious Diseases* 19, 1713–1715. doi:10.3201/eid1910.130626.

Rashid, M., Rashed, S. M., Islam, T., Johura, F.-T., Watanabe, H., Ohnishi, M., et al. (2016). *Ctx*B1 outcompetes *Ctx*B7 in *Vibrio cholerae* O1, Bangladesh. *Journal of Medical Microbiology* 65, 101–103. doi:10.1099/jmm.0.000190.

Ratchford, C., and Wang, J. (2019). Modeling cholera dynamics at multiple scales: environmental evolution, between-host transmission, and within-host interaction. *Mathematical and Biosciences Engineering* 16, 782–812. doi:10.3934/mbe.2019037.

Rebaudet, S., Sudre, B., Faucher, B., and Piarroux, R. (2013). Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes. *Journal of Infectious Diseases* 208 Suppl 1, S46-54. doi:10.1093/infdis/jit195.

Reidl, J., and Klose, K. E. (2002). *Vibrio cholerae* and cholera: out of the water and into the host. *FEMS Microbiology Reviews* 26, 125–139.

Reimer, A., Domselaar, G., Stroika, S., Walker, M., Kent, H., Tarr, C., et al. (2011). Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. *Emerging Infectious Diseases* 17. doi:10.3201/eid1711.110794.

Reyburn, R., Kim, D. R., Emch, M., Khatib, A., von Seidlein, L., and Ali, M. (2011). Climate variability and the outbreaks of cholera in Zanzibar, East Africa: a time series analysis. *American Journal of Tropical Medicine and Hygiene* 84, 862–869. doi:10.4269/ajtmh.2011.10-0277.

Rinaldo, A., Bertuzzo, E., Mari, L., Righetto, L., Blokesch, M., Gatto, M., et al. (2012). Reassessment of the 2010-2011 Haiti cholera outbreak and rainfall-driven multiseason projections. *Proceedings of the National Academy of Sciences* 109, 6602–6607. doi:10.1073/pnas.1203333109.

Robins, W. P., and Mekalanos, J. J. (2014). Genomic science in understanding cholera outbreaks and evolution of *Vibrio cholerae* as a human pathogen. *Current Topics in Microbiology and Immunology* 379, 211–229. doi:10.1007/82_2014_366.

Runft, D. L., Mitchell, K. C., Abuaita, B. H., Allen, J. P., Bajer, S., Ginsburg, K., et al. (2014). Zebrafish as a natural host model for *Vibrio cholerae* colonization and transmission. *Applied and Environmental Microbiology* 80, 1710–1717. doi:10.1128/AEM.03580-13.

Safa, A., Nair, G. B., and Kong, R. Y. C. (2010). Evolution of new variants of *Vibrio cholerae* O1. *Trends Microbiol.* 18, 46–54. doi:10.1016/j.tim.2009.10.003.

Sanchez, J., and Holmgren, J. (2011). Cholera toxin - a foe & a friend. *Indian Journal of Medical Research* 133, 153–163.

Sasaki, S., Suzuki, H., Igarashi, K., Tambatamba, B., and Mulenga, P. (2008). Spatial analysis of risk factor of cholera outbreak for 2003-2004 in a peri-urban area of Lusaka, Zambia. *American Journal of Tropical Medicine and Hygiene* 79, 414–421.

Seed, K. D., Lazinski, D. W., Calderwood, S. B., and Camilli, A. (2013). A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. *Nature* 494, 489–491. doi:10.1038/nature11927.

Senderovich, Y., Izhaki, I., and Halpern, M. (2010). Fish as reservoirs and vectors of *Vibrio cholerae*. *PLoS ONE* 5, e8607. doi:10.1371/journal.pone.0008607.

Shimada, T., Arakawa, E., Itoh, K., Okitsu, T., Matsushima, A., Asai, Y., et al. (1994). Extended serotyping scheme for *Vibrio cholerae*. *Current Microbiology* 28, 175–178. doi:10.1007/BF01571061.

Siriphap, A., Leekitcharoenphon, P., Kaas, R. S., Theethakaew, C., Aarestrup, F. M., Sutheinkul, O., et al. (2017). Characterization and genetic variation of *Vibrio cholerae* isolated from clinical and environmental sources in Thailand. *PLOS ONE* 12, e0169324. doi:10.1371/journal.pone.0169324.

Spagnoletti, M., Ceccarelli, D., Rieux, A., Fondi, M., Taviani, E., Fani, R., et al. (2014). Acquisition and evolution of SXT-R391 integrative conjugative elements in the Seventh-Pandemic *Vibrio cholerae* lineage. *mBio* 5, e01356-14. doi:10.1128/mBio.01356-14.

Stoltzfus, J. D., Carter, J. Y., Akpinar-Elci, M., Matu, M., Kimotho, V., Giganti, M. J., et al. (2014). Interaction between climatic, environmental, and demographic factors on cholera outbreaks in Kenya. *Infectious Diseases of Poverty* 3, 37. doi:10.1186/2049-9957-3-37.

Tamplin, M. L., Gauzens, A. L., Huq, A., Sack, D. A., and Colwell, R. R. (1990). Attachment of *Vibrio cholerae* serogroup O1 to zooplankton and phytoplankton of Bangladesh waters. *Applied and*

Environmental Microbiology 56, 1977–1980. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC184543/ [Accessed April 9, 2018].

Trærup, S. L. M., Ortiz, R. A., and Markandya, A. (2011). The Costs of climate change: A study of cholera in Tanzania. *International Journal of Environmental Research and Public Health* 8, 4386–4405. doi:10.3390/ijerph8124386.

Vezzulli, L., Pruzzo, C., Huq, A., and Colwell, R. R. (2010). Environmental reservoirs of *Vibrio cholerae* and their role in cholera: Environmental reservoirs of *V. cholerae*. *Environmental Microbiology Reports* 2, 27–33. doi:10.1111/j.1758-2229.2009.00128.x.

Vital, M., Füchslin, H. P., Hammes, F., and Egli, T. (2007). Growth of *Vibrio cholerae* O1 Ogawa Eltor in freshwater. *Microbiology (Reading, Engl.)* 153, 1993–2001. doi:10.1099/mic.0.2006/005173-0.

Wang, R., Yu, D., Yue, J., and Kan, B. (2016). Variations in SXT elements in epidemic *Vibrio cholerae* O1 El Tor strains in China. *Scientific Reports* 6. doi:10.1038/srep22733.

Weil, A. A., and Ryan, E. T. (2018). Cholera: recent updates. *Current Opinion in Infectious Diseases* doi:10.1097/QCO.0000000000000474.

Weill, F.-X., Domman, D., Njamkepo, E., Almesbahi, A. A., Naji, M., Nasher, S. S., et al. (2019). Genomic insights into the 2016–2017 cholera epidemic in Yemen. *Nature* 565, 230. doi:10.1038/s41586-018-0818-3.

Weill, F.-X., Domman, D., Njamkepo, E., Tarr, C., Rauzier, J., Fawal, N., et al. (2017). Genomic history of the seventh pandemic of cholera in Africa. *Science* 358, 785–789. doi:10.1126/science.aad5901.

Wendel, J. (2015). Climate change predicted to worsen spread of cholera. *Eos*. Available at: https://eos.org/articles/climate-change-predicted-worsen-spread-cholera [Accessed September 23, 2019].

WHO (2017). Ending cholera. A global roadmap to 2030. Available at: http://www.who.int/cholera/publications/global-roadmap/en/ [Accessed October 15, 2018].

WHO (2018a). Cholera – United Republic of Tanzania. *WHO*. Available at: http://www.who.int/csr/don/12-january-2018-cholera-tanzania/en/ [Accessed October 15, 2018].

WHO (2018b). WHO | Weekly epidemiological record: cholera articles. *WHO*. Available at: http://www.who.int/cholera/statistics/en/ [Accessed July 26, 2019].