Genomic analysis of the Mozambique strain of *Vibrio cholerae* O1 reveals the origin of El Tor strains carrying classical CTX prophage

Shah M. Faruque*†, Vincent C. Tam‡, Nityananda Chowdhury*, Pornphan Diraphat§, Michelle Dziejman‡¶, John F. Heidelberg‖, John D. Clemens**, John J. Mekalanos‡, and G. Balakrish Nair*†

*Molecular Genetics Laboratory, International Centre for Diarrhoeal Disease Research, Bangladesh, Dhaka-1212, Bangladesh; §Department of Microbiology, Faculty of Public Health, Mahidol University, Bangkok 73170, Thailand; ¶School of Medicine and Dentistry, University of Rochester, 601 Elmwood Avenue, Box 672, Rochester, NY 14642; ‖The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850; **International Vaccine Institute, Kwanak, P.O. Box 14, Seoul 151-600, Korea; and ‡Department of Microbiology and Molecular Genetics, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115

Contributed by G. Balakrish Nair, January 17, 2007 (sent for review January 5, 2007)

Cholera outbreaks in subSaharan African countries are caused by strains of the El Tor biotype of toxigenic *Vibrio cholerae* O1. The El Tor biotype is the causative agent of the current seventh cholera pandemic, whereas the classical biotype, which was associated with the sixth pandemic, is now extinct. Besides other genetic differences the CTX prophages encoding cholera toxin in the two biotypes of *V. cholerae* O1 have distinct repressor (*rstR*) genes. However, recent incidences of cholera in Mozambique were caused by an El Tor biotype *V. cholerae* O1 strain that, unusually, carries a classical type (CTX^{class}) prophage. We conducted genomic analysis of the Mozambique strain and its CTX prophage together with chromosomal phage integration sites to understand the origin of this atypical strain and its evolutionary relationship with the true seventh pandemic strain. These analyses showed that the Mozambique strain carries two copies of CTX^{class} prophage located on the small chromosome in a tandem array that allows excision of the prophage, but the excised phage genome was deficient in replication and did not produce CTX^{class} virion. Comparative genomic microarray analysis revealed that the strain shares most of its genes with the typical El Tor strain N16961 but did not carry the TLC gene cluster, and RS1 sequence, adjacent to the CTX prophage. Our data are consistent with the Mozambique strain's having evolved from a progenitor similar to the seventh pandemic strain, involving multiple recombination events and suggest a model for origination of El Tor strains carrying the classical CTX prophage.

cholera | CTX phage | TLC element | evolution

Cholera caused by toxigenic *Vibrio cholerae* is a major public health problem confronting many developing countries, where outbreaks occur frequently and are particularly associated with poverty and poor sanitation (1, 2). The occurrence of seven distinct pandemics of cholera have been recorded since the beginning of the first pandemic in 1817. The current seventh pandemic, which originated in Indonesia in 1961, is the most extensive in geographic spread and duration, and the causative agent is *V. cholerae* O1 of the El Tor biotype. The sixth pandemic and presumably the earlier pandemics were caused by *V. cholerae* O1 of the classical biotype. These two biotypes of *V. cholerae* O1 differ in certain phenotypic and genetic characteristics (2). In toxigenic *V. cholerae*, the genes encoding cholera toxin (*ctxAB*) are part of the CTX prophage (3). Besides other genetic differences between the two biotypes of *V. cholerae* O1, the CTX prophages in these two biotypes are also distinct in their sequence of the repressor gene (*rstR*), although most of the phage genomes are similar in the two biotypes (4, 5). The factors that led to the extinction of the classical biotype and its replacement by the El Tor biotype of *V. cholerae* O1 as a causative agent of cholera are not clearly known.

The seventh pandemic of cholera reached sub-Saharan West Africa in the early 1970s and caused explosive outbreaks with a high case fatality, mainly because of a lack of background immunity in the population and inadequacies in the health care infrastructure (2). In this epidemic, cholera spread along the coast and into the interior through waterways and further disseminated into the interior of the Sahelian states by land travel fostered by nomadic tribes (1, 2). Cholera outbreaks caused by the El Tor biotype of *V. cholerae* O1 continue to occur frequently in many sub-Saharan African countries. However, cholera surveillance in Beira, the second largest city of Mozambique, in early 2004, revealed the presence of a distinct El Tor strain associated with incidences of cholera (6). In preliminary analyses, this Mozambique strain was found to display most of the typical traits of the El Tor biotype, but, interestingly, the resident CTX prophage in the strain is of the classical type ($CTX^{Class}\Phi$). In view of rapidly growing genetic diversity among toxigenic *V. cholerae* strains with epidemic potential, we conducted detailed analyses of the Mozambique strain and its CTX prophage to understand the origin of this strain. Here, we show that the Mozambique strain represents a distinct branch in the evolutionary pathway of pathogenic *V. cholerae* strains with epidemic potential.

Results

Structure of the CTX Prophage Array in the Mozambique Strain. Previous studies have described the existence of at least three widely diverse repressor genes (*rstR* genes) carried by different CTX phages, i.e., $CTX^{ET}\Phi$, $CTX^{Class}\Phi$, and $CTX^{Calc}\Phi$ (4, 5, 7). This diversity of *rstR* constitutes the molecular basis for heteroimmunity among CTX phages. We examined the CTX prophage in 18 clinical isolates from Mozambique using specific probes for the different repressor genes *rstR* as well as probes for the *ctxA* and *zot* genes [see supporting information (SI) Fig. 5]. All isolates hybridized with the $rstR^{Class}$ probe, and not with other *rstR* probes, and the restriction patterns derived from respective CTX phage genes were identical for all of the isolates tested. The deduced CTX prophage array structure based on the observed restriction patterns of the *ctxA*, *zot*, and *rstR* genes in the

Author contributions: S.M.F., J.D.C., J.J.M., and G.B.N. designed research; S.M.F., V.C.T., N.C., P.D., M.D., and J.F.H. performed research; S.M.F., V.C.T., N.C., P.D., M.D., J.F.H., and J.J.M. analyzed data; and S.M.F. and J.J.M. wrote the paper.

The authors declare no conflict of interest.

Abbreviation: ER, end repeat.

†To whom correspondence may be addressed. E-mail: faruque@icddrb.org or gbnair@icddrb.org.

This article contains supporting information online at www.pnas.org/cgi/content/full/0700365104/DC1.

Mozambique strain agreed with a previous report (8), suggesting that the strain carries two copies of the CTX^{Class} prophage, arranged in tandem (SI Fig. 5). We attempted to identify the CTX prophage integration site in the Mozambique strain of *V. cholerae* O1 by analyzing PCR products containing the junctions between the prophage and chromosomal DNA. Interestingly, we found that primers specific for the TLC region (9), which is found upstream of the CTX prophage in other *V. cholerae* O1 strains, failed to produce a PCR amplicon, suggesting that the TLC region is absent in the Mozambique strain. This was later confirmed by using genomic microarray and sequence analysis, and both copies of the CTX^{Class} prophage were found to be located on the small chromosome (see below). Most El Tor strains are also known to carry a satellite phage genome originally termed RS1 element, which is related to CTXΦ (10). RS1 encodes an antirepressor rstC that counteracts the activity of the CTX phage repressor rstR and thus promotes expression of genes required for phage production (11). However, analysis of the Mozambique strain for the presence of RS1 showed that, as in classical biotype strains, the RS1 element was also absent in the Mozambique strain.

Excision of the CTX Prophage. The production of extrachromosomal CTXΦ genome (pCTX) by excision of the prophage and subsequent production of phage particles is presumed to depend on the structure of the chromosomal CTX array. Davis and colleagues (4) have proposed that a distinct process, apparently unrelated to phage DNA integration, usually underlies formation of CTXΦ virions. This process, which does not require homologous recombination, depends on the presence of tandem prophages or of a prophage followed by a related genetic element known as RS1. The CTX prophage is known to reside in the chromosome of El Tor biotype strains forming tandem arrays of two prophages or a prophage and an RS1. On the other hand, most classical strains have either a solitary phage or two fused truncated prophages (5), and classical strains are not known to produce pCTX or infectious virions. Because, in the Mozambique strain, the CTX prophage formed an array of two tandem prophages, we assumed that this array structure might support the excision of the CTX prophage to form pCTX in the Mozambique strain. To verify this, we conducted Southern blot analysis of plasmid and phage preparations from all isolates. Although none of the isolates produced phage particles, at least three isolates produced a plasmid band that hybridized with the CTX probe (data not shown). However, repeated attempts to clone the excised CTXΦ genome by inserting a kanamycin-resistance (Kan^R) marker in intergenic regions failed, although the excised phage genome could be cloned into pUC18, suggesting that the excised phage DNA was likely deficient in self-replication. The phage genome from one of these Mozambique isolates designated B-33 was cloned into pUC18 to construct pNSF-1. The insert of pNSF-1 was later sequenced and found to be the complete genome of a classical type CTXΦ (see below).

Comparative Genomic Analysis. The representative Mozambique strain B-33 was analyzed together with various control *V. cholerae* strains (SI Table 2) in our collection for their gene content as compared with strain N16961, a typical El Tor strain whose whole genome has been sequenced (12). The results of microarray analyses of various strains included in the study are summarized in Fig. 1. The array was based on the genomic sequence of N16961 and represented a total of 3,890 genes (13).

A total of 27 genes carried by strain N16961 were either absent or replaced in the Mozambique strain (SI Table 3). Whereas the CTX phage repressor gene *rstR* was of the El Tor type in N1696, the *rstR* gene was of the classical type in the Mozambique strain, because this strain carried the CTX^{class} prophage. In agreement with typical classical strains, the Mozambique strain did not carry the RS1 and was, hence, negative for the *rstC* gene. Among other

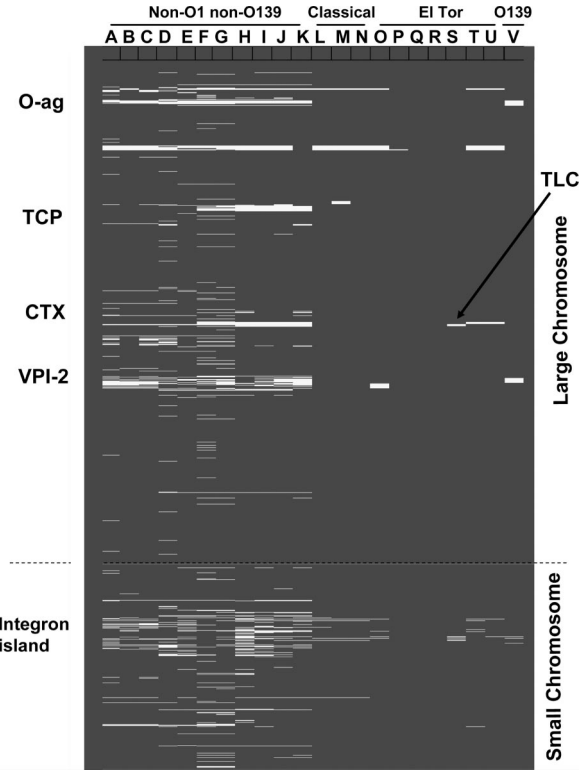

Fig. 1. Results of microarray-based comparative genomic analyses of the Mozambique strain of *V. cholerae* O1 with diverse *V. cholerae* O1 and non-O1 strains. The identity of strains shown in different lanes are A–E: WBDV-101E, WBDH697, WBDH712, FY2G, and NWBD40–1F1; F–K: WBDH552, WBDH336, AM15622, AM19226, MZO2, and MZO3; L–N: 569B, NIH41, and O395; O–R: MAK757, C6709, HK1, and N16961; S: Mozambique strain B-33; T and U: 2470–80 and NCTC; and V: MO10. The serogroups of different strains are shown above the figure.

genes that were found to be absent from the Mozambique strain is the TLC gene cluster (8), which was also absent from most of the non-O1 non-O139 strains analyzed in the study. These non-O1 non-O139 strains, which included both toxigenic and nontoxigenic strains, were from different sources and were found to be clonally diverse in a previous investigation (13). Whereas the non-O1 non-O139 strains carried between 92% and 95% of the N16961 genes of both the chromosomes, the O1 strains appear to be more clonal and conserve ≈98–99% of these genes. In agreement with previous studies, we found that occasionally the test strains were negative for clusters of genes that correspond to multigene insertions or chromosomal "islands" in the N16961 genome. However, the microarray data did not always indicate that every gene of an island was absent. In some cases, PCR and/or Southern blot analysis was conducted to confirm the presence or absence of genes within a suspected island insertion. This finding was particularly true for environmental isolates and supported our previous assumption that the gene clusters might have been derived from genes that are dispersed as individual genes among environmental strains.

Sequence of the CTX Prophage and Phage Integration Sites. Determination of the nucleotide sequence of the excised CTXΦ genome in the Mozambique strain B-33 revealed that the phage genome comprised 7,007 nucleotides and contained all of the ORFs and intergenic sequences of a typical CTXΦ genome. The orientation of the ORFs were also identical to that of previously reported CTX phage sequences. The sequence of *rstR* gene was identical to that of the CTX^{Class} prophage in strain 569B. There

Table 1. Occupancy of two chromosomal attachment sites by CTX or related phages and the TLC element in various *V. cholerae* strains

Strain	Chr I	TLC	Chr II	TLC
N16961	CTX^{ET}	Yes	Empty	No
AM19226	f237	No	VSK	No
O395	CTX^{Class}	Yes	CTX^{Class}	No
B33	Empty	No	CTX^{Class}	No

B33 is the 2004 Mozambique isolate; N16961 seventh pandemic El Tor; O395 classical biotype; and AM19226 is a non-O1 non-O139 *V. cholerae*.

were minor differences among the sequences of other genes of CTX prophage in B-33 and those of $CTX^{Class}\Phi$ carried by strain 569B and $CTX^{ET}\Phi$ of strain N16961. However, none of these nucleotide replacements or deletions appear to have resulted in frame-shift mutations in any of the genes.

Analysis of the sequence of the phage integration sites (*att*) on both chromosomes of the Mozambique strain B-33, revealed interesting differences between typical El Tor strains and strain B-33. The seventh pandemic El Tor strain represented by N16961 carries the CTX^{ET} prophage located at the *att* region between dif core1 and dif core2 of chromosome I, with adjacent elements known as TLC and RS1 (Table 1 and Figs. 2 and 3). Dif sites are located opposite the origin of replication and are binding sites for recombinases XerC and XerD to resolve dimeric chromosomes before cell division (14, 15). Previous studies have shown that CTXΦ integration requires XerC and XerD, and CTXΦ integrates into the normally overlapping dif core1 and dif core2 sites (16, 17). $CTX^{ET}\Phi$ is integrated into the attRS1 site within dif core1, thus separating dif 1 and dif 2 in N16961 (Fig. 3). Surprisingly, we found that B33 carried only the dif core2 site in chromosome I. Although B33 does not carry a CTX prophage in chromosome I, two copies of the CTX^{Class} prophage in tandem are located on chromosome II. These copies apparently inserted at a second filamentous phage attachment site in B33 that is functional for $CTX^{Class}\Phi$, such as those carried by the typical classical strain O395. These and other differences in the B33 *att* and related sequences are summarized in Table 1 and Fig. 3. Interestingly, the CTX-negative strain AM19226 carries two other filamentous phages (VSK and f237), each inserted at the dif core sites on the two chromosomes. The absence of TLC and dif core1 site in chromosome I of B33 could be explained by a recombination event between flanking attRS1-related end repeat (ER) sites (Fig. 4). Based on sequence across the filled and empty dif sites in different strains, it appears that during this recombination event, the precursor to B33 might have lost the TLC element, CTXΦ, and the dif core1 (Table 1 and Figs. 2 and 3). Without dif core1, which contains an attRS1 site required for integration of $CTX^{ET}\Phi$ into chromosome I, the precursor strain was no longer able to acquire $CTX^{ET}\Phi$. Thus, if cholera toxin is required for virulence, this hypothetical precursor strain was "forced" to pick up $CTX^{Class}\Phi$, which can use the chromosome II attachment site.

Discussion

Studying the population structure and molecular evolution of *V. cholerae* as a species and the temporal genetic changes that lead to the emergence of new epidemic strains are challenging. Several studies using a variety of molecular typing methods have concluded that seventh pandemic strains represent a globally distributed clone that is closely related to the more recently emerged O139 clone (18, 19). The emergence of the O139 serogroup provided an unique opportunity to witness major genetic changes in an existing seventh pandemic El Tor strain and the epidemiological changes associated with the displacement of an existing serogroup by a new emerging one (20, 21). On the other hand, the recently emerged strain in Mozambique shows most of the chracteristics of an O1 El Tor strain although unexpectedly carrying a classical type CTX prophage (6). These characteristics include resistance to polymyxin B, agglutination of chicken erythrocytes, production of the El Tor hemolysin, sensitivity to group IV El Tor phage, and resistance to the

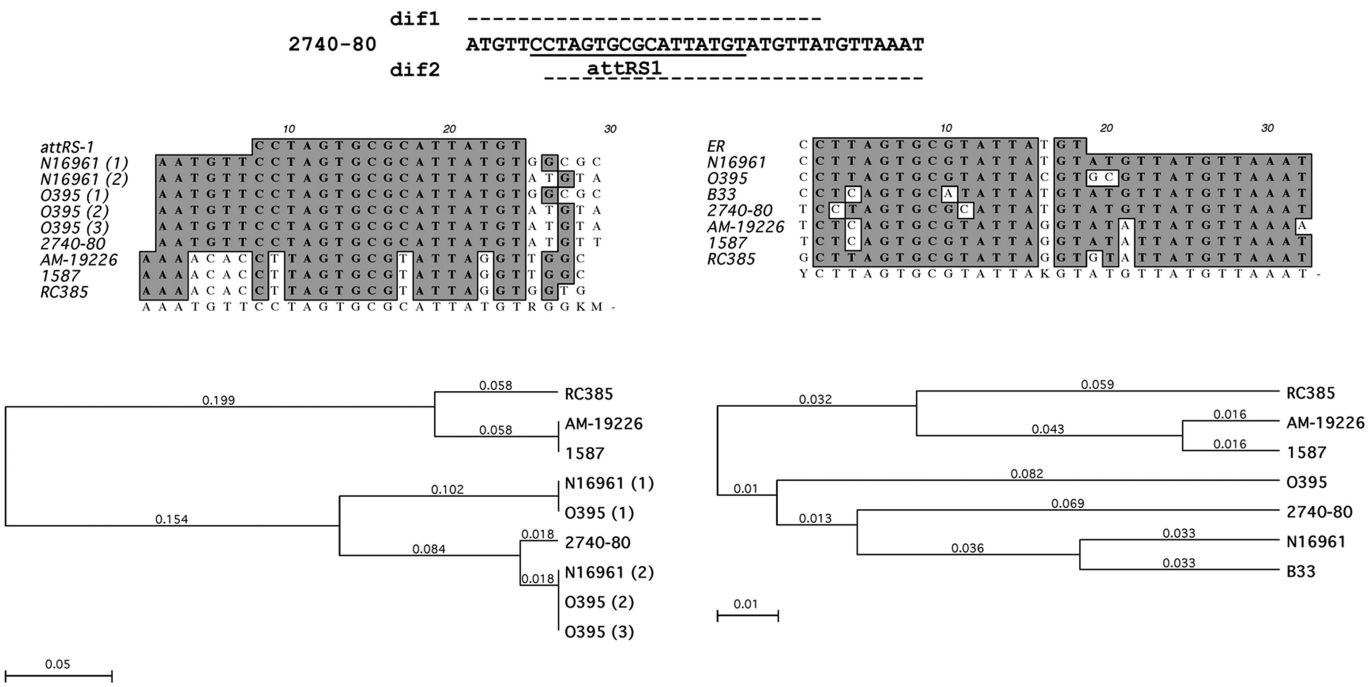

Fig. 2. Sequence comparison and alignment of chromosome I dif sites in various *V. cholerae* strains. The attRS1 site is underlined; B33 is the 2004 Mozambique isolate; N16961 seventh pandemic El Tor; O395 classical biotype; 2740–80 is an environmental isolate that is clonal with $TCP^{+}CTX^{+}$ isolates but lacks CTX; AM19226 and 1587 are non-O1, non-O139 clinical isolates of *V. cholerae*; and RC385 is a non-O1, non-O139 environmental isolate of *V. cholerae*.

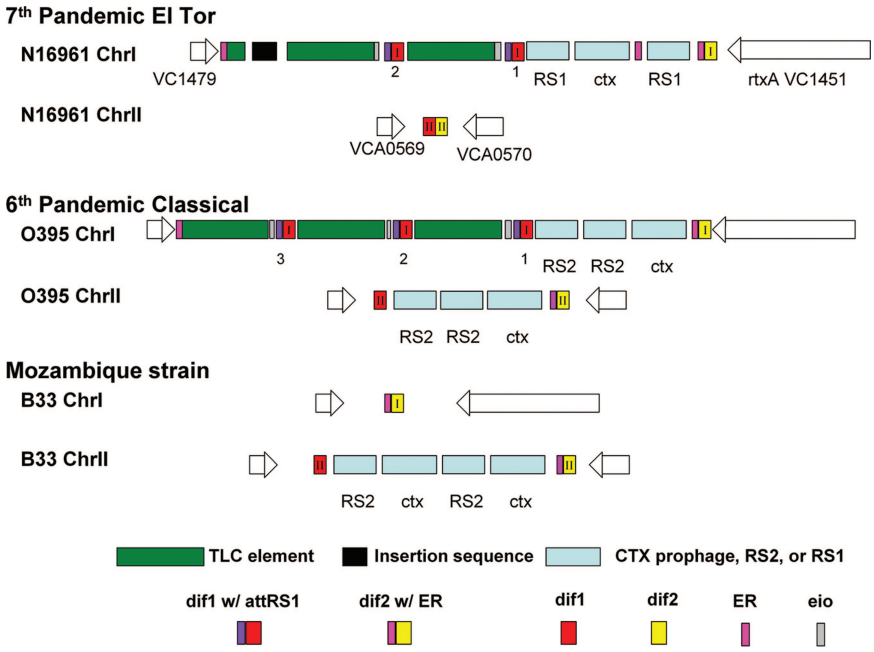

Fig. 3. Arrangement of the CTX prophage, RS1, TLC, and their adjacent genes as well as various attachment sites and related sequences on both chromosomes of a typical El Tor strain N16961, a classical strain O395, and the recently emerged Mozambique strain B-33. N16961 Chromosome I, dif core1 and dif core2 sites are separated by CTX^{ET}. Both dif core sites are deviated from consensus, and they contain either an attRS1 or end repeat (ER, similar 17-bp sequence, instead of 18 bp). N16961 chromosome II, no phage insertion, therefore dif core1 and dif core2 are intact (overlapping). O395 chromosome I, similar to N16961 dif core1 and dif core2, they have either overlapping attRS1 or ER site, but dif cor1 and dif core2 are separated only by CTX^{Class}. O395 Chromosome II, CTX^{Class} is inserted into dif core1 and dif core2 sites with the exception that dif core1 does not contain attRS1 or ER site. B33 Chromosome I, contains a dif core2 site with ER but does not have a dif core1 site and lacks TLC. B33 Chromosome II, as in O395 chromosome II, CTX^{Class} is integrated into dif core1 and dif core2 sites and dif core1 does not include attRS1 or ER.

classical group V phage. Based on the array data, the Mozambique strain B-33 has some absent genes characteristic of both classical and non-O1 non-O139 strains, but carries most of the genes (>99%) of N16961 and other typical seventh pandemic El Tor strains analyzed. These include the VSP-1 and VSP-2 islands (13), as well as the three genes (VCA0728-VCA730), which are

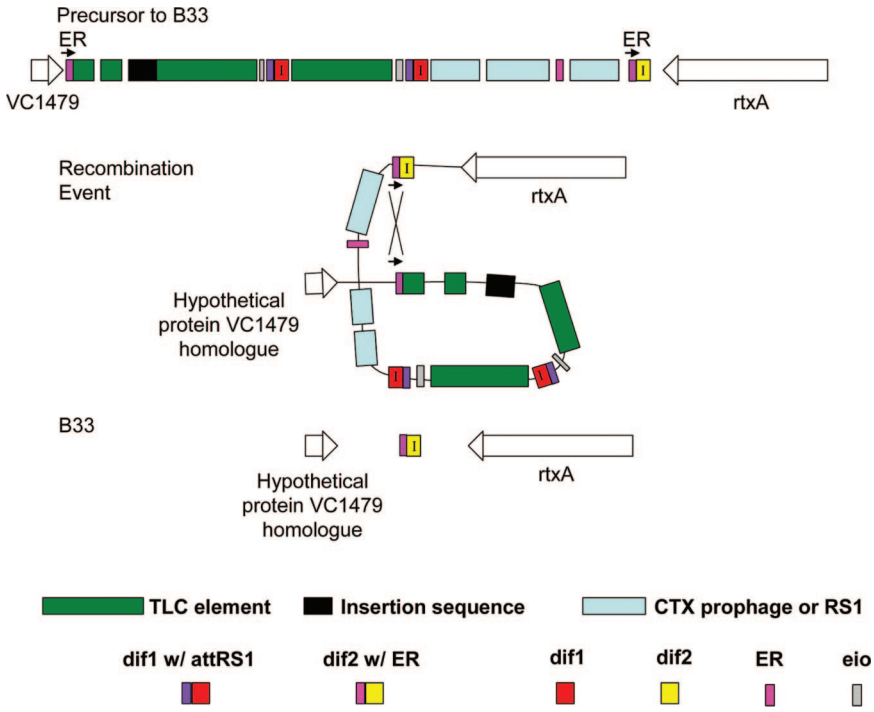

Fig. 4. Proposed recombination event in the deletion of the entire region containing CTX prophage, RS1, and TLC from the large chromosome of a putative precursor strain in the pathway to the emergence of the Mozambique strain B-33.

carried by all El Tor strains but are absent from classical and non-O1 non-O139 strains. These data are consistent with strain B33's having evolved from the seventh pandemic strain.

We propose a pathway in which multiple recombination events were involved in the origination of the Mozambique strain from a precursor seventh pandemic strain similar to N16961. In this pathway, the Mozambique strain was derived from an El Tor-like strain that lost the El Tor CTX phage, TLC, and RS1, based on a loop-out recombination between related *att* sites, preventing future El Tor phage acquisition (Fig. 4). Because the remaining attachment site on chromosome II can be used only by a $CTX^{Class}\Phi$, the strain acquired a $CTX^{Class}\Phi$ to generate B33. The acquisition of CTX phage might have been driven by strong selective pressure to colonize and amplify in the host intestinal environment.

It is not clear whether the progenitor strain of B33 received the classical CTX phage genome through an infectious $CTX^{Class}\Phi$, particularly because production of CTXΦ particles by classical *V. cholerae* strains has not been demonstrated. However, the possibility that infectious $CTX^{Class}\Phi$ particles existed in nature or the phage genome might have been packaged into another phage cannot be ruled out. It is also not clear whether the circular but apparently nonreplicating form of the classical CTXΦ genome demonstrated in this study is a required intermediate for its transfer between *V. cholerae* strains. Interestingly, a recent study has suggested that *V. cholerae* may become naturally competent to uptake DNA from its environment when induced by growing on a chitin substrate (22). Thus, the precursor strain of B33 might have accepted a free CTX^{Class} phage genome or a chromosomal segment of a classical *V. cholerae* strain carrying the CTX^{Class} prophage from the environment.

It is now generally agreed that there were at least two critical gene transfer events in the evolution of pathogenic *V. cholerae* from a nonpathogenic precursor strain (9). First, the progenitor strain acquired the TCP pathogencity island (via a mechanism that is not fully understood) and thereby became colonization proficient while simultaneously acquiring the receptor for CTXΦ (3). Then, perhaps in the intestine, this TCP-positive strain was infected by CTXΦ. Previously, the finding that all TCP-positive *V. cholerae* O1 strains are also positive for TLC, irrespective of their CTX status, suggested the possibility that acquisition of the TLC element by TCP-positive *V. cholerae* strains is an intermediate step preceding the acquisition of CTXΦ in the evolution of fully pathogenic *V. cholerae* (9). We examined the sequences surrounding chromosome I dif sites of several non-O1 non-O139 strains that are negative for TCP, TLC, and CTX as well as strain 2740-80, an environmental U.S. Gulf Coast isolate that is clonal with TCP^+CTX^+ isolates but lacks CTX. From the sequence alignment and phylogenetic analysis (Fig. 2), dif1 sites of strains containing TLC are closely related and have intact attRS1 sites, whereas the three non-O1 non-O139 strains have a different dif 1 site and do not contain an attRS1 site, which is required for CTXΦ integration. Furthermore, the TLC element is linked to CTXΦ by an element-intervening ORF (*eio*) (9); the TLC and *eio* sequences are duplicated once in N16961 and twice in O395. The number of TLC duplications corresponds to the number of dif1 and attRS1 sites present in the region; explicitly, there are one TLC element and one dif 1/attRS1 in 2740–80, two in N16961, and three in O395. Thus, acquisition of TLC copies is correlated with duplications of dif 1 sites. Interestingly, the *eio* genes between the duplicate TLC and dif1/attRS1 sequences are truncated, suggesting that different versions of TLC carrying different *eio* sequences were acquired through time. Based on the chromosomal location and DNA sequence analysis, TLC may have initially brought in the attRS1 site that allowed the seventh pandemic strain to emerge through lysogenic conversion by the El Tor CTX phage. We predict that the precursor of the seventh pandemic clone needed to pick up the TLC element currently in seventh pandemic strains because this element can potentially generate an attRS1 site upon insertion into the chromosomal I site. TLC encodes a filamentous phage replicase and is likely a satellite phage but has not been studied in this context. The concept that a satellite phage (TLC) can carry only a short sequence that, upon integration into the chromosome, could generate a new functional *att* site for another phage ($CTX^{ET}\Phi$) that carries a selectable trait (i.e., cholera toxin) is fascinating and deserves further experimental investigation.

The reconstruction of the events that led to the origination of strain B33 can also shed light on the emergence of the seventh pandemic clone. This also provides evidence that the seventh pandemic El Tor genomic backbone remains more fit than other TCP^+CTX^+ strains, because the Mozambique strain remains a seventh pandemic El Tor strain, despite carrying the classical CTX prophage. Based on molecular markers, the sixth-pandemic, seventh-pandemic, and U.S. Gulf coast isolates have been classified as three distinct clones, apparently evolving independently from environmental, nontoxigenic, non-O1 strains (23). The emergence of the Mozambique strain and its association with incidences of cholera may represent yet another pathway for the evolution of pathogenic strains from environmental *Vibrio*s that have their origin in pathogenic strains that have lost virulence factors. Whether the loss and gain of virulence factors provide strains with increased fitness in the environment or in the host remains to be determined. It will also be interesting to see whether the Mozambique clone of *V. cholerae* El Tor O1 will eventually replace the typical seventh pandemic strain as a causative agent of cholera in the sub-Saharan African region and elsewhere in the world. If this occurs, we propose that the minor differences in the sequence or regulation of the *ctxAB* genes on the two copies of the CTX^{Class} phage carried by B33 are the most likely explanation for this improved evolutionary fitness.

Materials and Methods

***V. cholerae* Strains.** A total of 18 *V. cholerae* isolates from different cholera patients in Beira, Mozambique (6) between January and March 2004, were included in the study. *V. cholerae* strains used for comparison were obtained from the culture collection of the International Center for Diarrhoeal Disease Research in Dhaka, Bangladesh (ICDDRD,B). Strains were stored in sealed deep nutrient agar at room temperature. Before use, all strains were rechecked by biochemical serological methods (24) to confirm their identity.

Probes and PCR Assays. All strains were analyzed for the presence of different virulence genes, as well as distinctive genes carried by the two biotypes of *V. cholerae* O1. Presence of *tcpA* genes specific for the classical and El Tor biotypes was determined by using a multiplex PCR assay as described previously (25). Presence of classical, El Tor, and Calcutta type *rstR* genes of the CTX prophage and the *rstC* gene of the RS1 element were determined by PCR as well as specific DNA probes as described by us (25–27). All probes were labeled by random priming (28) using a random primers DNA labeling kit (Invitrogen, Carlsbad, CA) and [α-^{32}P]deoxycytidine triphosphate [3,000 Ci/mmol (1 Ci = 37 GBq); Amersham Biosciences, Uppsala, Sweden]. Southern blots were hybridized with the labeled probes and washed under stringent conditions, and autoradiographs were developed by using standard methods (29).

Microarray Analysis. The preparation of *V. cholerae* DNA microarrays and their use in comparative genomic studies have been described (13). Briefly, genomic DNA from each strain was extracted by using the EasyDNA kit (Invitrogen), and 3 μg of DNA from the test strain was used to generate Cy3- and Cy5-labeled samples (13). For each test strain, two independent

MICROBIOLOGY

gDNA preparations were used as template, and both Cy3 and Cy5-dCTP were used in independent labeling and hybridization experiments to account for any differences in DNA preparation or dye incorporation. Hybridized, washed slides were scanned for Cy5 and Cy3 fluorescence intensities by using a ScanArray 5000 (Packard Instruments). The resulting files were analyzed by using GENEPIX 3.0 software (Axon Instruments, Foster City, CA). In a single array experiment, the genomic content from one of the test strains was compared with that of N16961. For each of the test strains, data were compiled from at least two array experiments. Confirmation of absent genes included verification by Southern blot analysis, PCR analysis, or both methods. Standard 50-ml PCRs using TaqDNA polymerase were performed according to the manufacturer's instructions (Invitrogen). Southern blot analysis was performed by using standard protocols (29).

Sequencing the B-33 Genome. A B-33 genomic library was constructed according to standard methods. Sequencing was performed with pSMART primers (Lucigen, Middleton, WI) in 10-μl reactions in 384-well plates using Applied Biosystems BIGDYE V 3.1. Products were purified by using the CLEANSEQ system (Agencourt, Beverly, MA) on BIOMEK FX (Beckman Coulter, Fullerton, CA) and applied to an ABI 3730 sequencer. Chromatograms were transferred to the UNIX platform, and quality control–quality assessment scripts were run. Data were analyzed and assembled by using PHRED, PHRAP, and CONSED (30–32). Sequence data were submitted to the Annotation Engine at The Institute for Genomic Research (Rockville, MD). The GLIMMER system (33, 34) was used to identify putative ORFs.

In addition, the CTX prophage in B33 was also sequenced separately from the excised pCTX DNA. Initially, overlapping subclones of the excised pCTX derived from the CTX prophage of a representative Mozambique strain B-33 was constructed in pUC18 and were sequenced by using universal sequencing primers (Invitrogen). Nucleotide sequencing was performed with an Applied Biosystems automated DNA sequencing system (ABI Prism 310 PE; Applied Biosystems, Foster City, CA) using BigDye terminator cycle sequencing ready reaction kit (PE; Applied Biosystems). Nucleotide sequence of both strands of pCTX were further determined by primer walking with primers derived from the preliminary sequencing of pCTX subclones in pUC18. Sequences were processed by using the Sequencher alignment program, Version 4.0 (Gene Codes, Ann Arbor, MI). Nucleotide sequence of pCTX of strain B-33 was compared with sequences in the GenBank databases, and protein homology search was done by using the National Center for Biotechnology Information BLAST server program.

Sequence Analysis. Sequence analysis was performed by using Artemis (Sanger, Cambridge, U.K.), and Lasergene (DNASTar, Madison, WI). Sequence alignments were created by using the CLUSTALW alignment feature of MACVECTOR 7.2.3.

We thank Derek Sturtevant (Department of Microbiology and Molecular Genetics, Harvard Medical School) for assisting with microarrays. This work was supported, in part, by National Institutes of Health Research Grant GM068851 under a subagreement between the Harvard Medical School and the International Centre for Diarrhoeal Disease Research Bangladesh (ICDDRB) and by the Swedish International Development Agency (SIDA) under an agreement with ICDDRB. The ICDDRB is supported by countries and agencies that share its concern for the health problems of developing countries.

1. Faruque SM, Albert MJ, Mekalanos JJ (1998) *Microbiol Mol Biol Rev* 62:1301–1314.
2. Kaper JB, Morris JG, Levine MM (1995) *Cholera Clin Microbiol Rev* 8:48–86.
3. Waldor MK, Mekalanos JJ (1996). *Science* 272:1910–1914. 37.
4. Davis BM, Moyer KE, Boyd EF, Waldor MK (2000) *J Bacteriol* 182:6992–6998.
5. Boyd EF, Heilpern AJ, Waldor MK (2000) *J Bacteriol* 182:5530–5538.
6. Ansaruzzaman M, Bhuiyan NA, Nair GB, Sack DA, Lucas M, Deen JL, Ampuero JC, Chaignat L (2004) *Emerg Infect Dis* 10:2057–2059.
7. Davis BM, Kimsey HH, Chang W, Waldor MK (1999) *J Bacteriol* 181:6779–6787.
8. Lee JH, Han KH, Choi SY, Lucas MES, Mondlane C, Ansaruzzaman M, Nair GB, Sack DA, Seidlein LV, Clemens JD, *et al.* (2006) *J Med Microbiol* 55:165–170.
9. Rubin EJ, Lin W, Mekalanos JJ, Waldor MK (1998) *Mol Microbiol* 28:1247–1254.
10. Faruque SM, Asadulghani, Kamruzzaman M, Nandi RK, Ghosh AN, Nair GB, Mekalanos JJ, Sack DA (2002) *Infect Immun* 70:163–170.
11. Davis BM, Kimsey HH, Kane AV, Waldor MK (2002) *EMBO J* 21:4240–4249.
12. Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, Haft DH, Hickey EK, Peterson JD, Umayam L, *et al.* (2000) *Nature* 406:477–483.
13. Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ (2002) *Proc Natl Acad Sci USA* 99:1556–1561.
14. Blakely G, May G, McCulloch R, Arciszewska LK, Burke M, Lovett ST, Sherratt DJ (1993) *Cell* 75:351–361.
15. Kuempel PL, Henson JM, Dircks L, Tecklenburg M, Lim DF (1991) *New Biol* 3:799–811.
16. Huber KE, Waldor MK (2002) *Nature* 417:656–659.
17. Iida T, Makino K, Nasu H, Yokoyama K, Tagomori K, Hattori A, Okuno T, Shinagawa H, Honda T (2002) *J Bacteriol* 184:4933–4935.
18. Bik EM, Bunschoten AE, Gouw RD, Mooi FR (1995) *EMBO J* 14:209–216.
19. Faruque SM, Saha MN, Asadulghani, Sack DA, Sack RB, Takeda Y, Nair GB (2000) *J Infect Dis* 182:1161–1168.
20. Cholera Working Group, International Centre of Diarrhoeal Disease Research, Bangladesh (1993) *Lancet* 342:387–390.
21. Ramamurthy T, Garg S, Sharma R, Bhattacharya SK, Nair GB, Shimada T, Takeda T, Karasawa T, Kurazano H, Pal A, Takeda Y (1993) *Lancet* 341:703–704.
22. Meibom KL, Blokesch M, Dolganov NA, Wu C, Schoolnik GK (2005) *Science* 310:1824–1827.
23. Wachsmuth IK, Olsvik Ø, Evins GM, Popovic T (1994) in *Vibrio cholerae and Cholera: Molecular to Global Perspectives*, eds Wachsmuth IK, Blake PA, Olsvik Ø (Am Soc Microbiol, Washington, DC), pp 357–370.
24. World Health Organization (1974) *World Health Organization Guidelines for the Laboratory Diagnosis of Cholera* (Bacterial Disease Unit, World Health Organ, Geneva).
25. Keasler SP, Hall RH (1993) *Lancet* 341:1661.
26. Faruque SM, Chowdhury N, Kamruzzaman M, Dziejman M, Rahman MH, Sack DA, Nair GB, Mekalanos JJ (2004) *Proc Natl Acad Sci USA* 101:2123–2128.
27. Faruque SM, Chowdhury N, Kamruzzaman M, Ahmad QS, Faruque AS, Salam MA, Ramamurthy T, Nair GB, Weintraub A, Sack DA (2003) *Emerg Infect Dis* 9:1116–1122.
28. Feinberg A, Volgelstein B (1984) *Anal Biochem* 137:266–267.
29. Maniatis T, Fritsch EF, Sambrook J (1982) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY)
30. Ewing B, Green P (1998) *Genome Res* 8:186–194.
31. Ewing B, Hillier L, Wendl MC, Green P (1998) *Genome Res* 8:175–185.
32. Gordon D, Abajian C, Green P (1998) *Genome Res* 8:195–202.
33. Salzberg SL, Delcher AL, Kasif S, White O (1998) *Nucleic Acids Res* 26:544–548.
34. Delcher AL, Harmon D, Kasif S, White O, Salzberg SL (1999) *Nucleic Acids Res* 27:4636–4641.