Molecular evidence favouring step-wise evolution of Mozambique *Vibrio cholerae* O1 El Tor hybrid strain

Kalpataru Halder,¹ Bhabatosh Das,^{1†} G. Balakrish Nair² and Rupak K. Bhadra¹

Correspondence

Rupak K. Bhadra
rupakbhadra@iicb.res.in

¹Indian Institute of Chemical Biology (CSIR), Kolkata 700 032, India

²National Institute of Cholera and Enteric Diseases (ICMR), Kolkata 700 010, India

The *ctxAB* operon, encoding cholera toxin (CT) in *Vibrio cholerae*, is carried by the genome of a filamentous phage, CTX Φ . Usually, specific CTX Φ infect each of the two important biotypes, classical and El Tor, of epidemic *V. cholerae* strains belonging to serogroup O1, and are called CTX^{class} Φ and CTX^{ET} Φ , respectively. However, an unusual hybrid El Tor strain carrying CTX^{class} Φ caused the cholera epidemic in Mozambique in 2004. To understand the evolution of that strain, we have further analysed some representative hybrid El Tor strains isolated in Kolkata, India, in 1992, and the results indicate that both the Mozambique and the Indian strains are infected with a unique CTX^{class} Φ having only four copies of the tandem heptamer repeat sequence 5'-TTTGGAT-3' present in the *ctxAB* promoter (P_{ctxAB}) region, like in CTX^{ET} Φ . Usually, the P_{ctxAB} of the classical biotype contains seven to eight copies of such sequences. However, sequence analyses of the P_{ctxAB} regions of several classical strains indicated that the copy number of heptamer repeat sequences might vary from four to eight copies, which was previously unknown. Since the hybrid strains analysed in this study carry four copies of the heptamer sequences, it may thus serve as a marker to trace the strain in future. Interestingly, while the Mozambique strain is devoid of an El Tor-specific free RS1 element or pre-CTX prophage, the Indian hybrid strains carry such elements. The free RS1 has been mapped, cloned and sequenced. As in pre-CTX and CTX prophages, multiple copies of free RS1 elements were found to be integrated in tandem in the large chromosomal *dif* site. Since Indian hybrid El Tor strains carry either free RS1 or pre-CTX prophage in their large chromosomes, it is possible that the Mozambique hybrid El Tor strain has evolved from these progenitor strains by step-wise deletion of CTX genetic elements from their large chromosomes.

Received 7 July 2009
Revised 23 September 2009
Accepted 30 September 2009

†Present address: Centre de Génétique Moléculaire, Bâtiment 26, CNRS UPR2167, Avenue de la Terrasse, 91198 Gif Sur Yvette, France.

Abbreviations: CT, cholera toxin; TLC element, toxin-linked cryptic element.

The GenBank/EMBL/DDBJ accession number for the nucleotide sequence of the tandem RS1 copies present in the large chromosome of VC44 is EU860439. The GenBank accession numbers for the P_{ctxAB} regions of the following strains: V154, B33, VC44, NIH41, GP8, NIH35A3 and H218 are FJ641042–FJ641048, respectively, and that of VC106 is FJ660421.

A supplementary figure showing PCR analysis of the large chromosomal CTX element integration sites of hybrid El Tor strains VC44 and VC106 is available with the online version of this paper.

Part of this work was presented at the 13th International Conference on Emerging Infectious Diseases in the Pacific Rim under the US–Japan Co-operative Medical Science Program, Kolkata, India, 6–9 April 2009.

INTRODUCTION

Cholera is an acute diarrhoeal disease caused by the Gram-negative bacterium *Vibrio cholerae*. *V. cholerae* strains belonging to serogroups O1 and O139 are responsible for the disease and they are distinguished from other serogroups by their capacity to produce cholera toxin (CT), which is the principal virulence factor responsible for profuse watery diarrhoea (Kaper *et al.*, 1995). *V. cholerae* O1 strains are divided into two important biotypes, classical and El Tor. The classical biotype was responsible for cholera pandemics until 1961. After that, the El Tor biotype evolved as a seventh pandemic strain, which is still continuing. *V. cholerae* O139 is a recently evolved strain, and several lines of evidence indicate that it is genetically closely related to the El Tor biotype (Sack *et al.*, 2004).

The *ctxAB* operon, encoding CT in *V. cholerae*, is carried by a filamentous lysogenic bacteriophage, called CTX Φ (Waldor & Mekalanos, 1996). CTX Φ specifically integrates in the dimer resolution site *dif* present in both chromosomes of *V. cholerae* (Huber & Waldor, 2002; Val *et al.*, 2005). The genome of CTX Φ comprises RS2 (2.4 kb) and a core region (4.5 kb). The RS2 element contains three genes, 5'-*rstR-rstA-rstB*-3', while the core region harbours several genes, including the *ctxAB* operon (Waldor & Mekalanos, 1996; Waldor *et al.*, 1997). Interestingly, several distinct CTX Φ genomes have been described, namely CTX^{ET} Φ , CTX^{class} Φ and CTX^{calc} Φ (Davis *et al.*, 1999). The best-characterized classical and El Tor variants of CTX Φ carry distinct *rstR* and *ctxB* alleles, *rstR*^{class} *ctxB*^{class} and *rstR*^{ET} *ctxB*^{ET}, respectively (Davis *et al.*, 1999; Olsvik *et al.*, 1993). Only *V. cholerae* O1 El Tor and O139 strains carry an RS2-like element, called RS1, which usually flanks the 5' and/or 3' region of the integrated CTX prophage (Davis *et al.*, 2000; Davis & Waldor, 2000). It has been reported that RS1 may integrate as a 'free' genetic element in El Tor strains (Sharma *et al.*, 1997). Along with *rstRAB* genes, the RS1 element carries the unique *rstC* gene, which encodes the CTX phage antirepressor protein RstC (Davis *et al.*, 2002). No information is currently available about the chromosomal location of free RS1, its copy number or arrangement in *V. cholerae*. Although the RS1 element has been demonstrated to be a satellite phage (Faruque *et al.*, 2002), the exact role of CTX prophage-associated RS1 or free RS1 in the phage biology or virulence of *V. cholerae* is unknown.

Since *V. cholerae* possesses two circular chromosomes (Heidelberg *et al.*, 2000; Trucksis *et al.*, 1998), one large (~3.0 Mb) and one small (~1.0 Mb), CTX Φ can integrate in either or both of the chromosomes (Heidelberg *et al.*, 2000; Nandi *et al.*, 2003; Trucksis *et al.*, 1998). In El Tor/O139 genomes, the CTX prophage may be present either as a single copy or as multiple copies arranged in tandem (Bhadra *et al.*, 1995; Mekalanos, 1983; Nandi *et al.*, 2003). In sharp contrast, in the classical biotype, the prophage is usually present in two copies, one copy in each chromosome (Trucksis *et al.*, 1998). However, one of these copies carries a truncated CTX prophage (Davis *et al.*, 2000). Whole-genome sequencing of the El Tor strain N16961 has revealed the integration of only one copy of the CTX prophage flanked by the RS1 element in its large chromosome (Heidelberg *et al.*, 2000). We previously reported the presence of tandem copies of the CTX prophage in the small chromosome of the pre-O139 El Tor strain (Nandi *et al.*, 2003). Recently, it has been shown that the Mozambique epidemic El Tor strain isolated in 2004 also carries tandemly arranged multiple copies of the CTX prophage in its small chromosome (Das *et al.*, 2007; Faruque *et al.*, 2007). Although Mozambique isolates were genotypically and phenotypically El Tor; nevertheless, the strains harboured classical CTX prophage (Lee *et al.*, 2006), indicating that they are hybrid in nature. Interestingly, our laboratory has shown that El Tor strains isolated in Kolkata, India, in 1992 also carried tandem copies of

classical CTX prophage in their small chromosomes (Das *et al.*, 2007). Since the sixth pandemic strains of classical biotype are believed to be extinct (Nair *et al.*, 2002), the evolution of the classical phage genomes in Mozambique or Kolkata El Tor isolates remains unknown.

In this study we have shown that although the hybrid El Tor strains isolated in Mozambique and Kolkata, India, do carry a classical-type CTX prophage, genetically it is different from the well-studied CTX^{class} Φ genome. It has been reported earlier that the *ctxAB* promoter (*P*_{*ctxAB*}) of the classical strain contains seven to eight copies of a tandemly arranged heptamer repeat sequence, 5'-TTTGGAT-3', but in the El Tor biotype, only three to four copies of such repeat sequences are present. This distinctive difference between the two biotypes is believed to be the critical determinant of CT production; classical being a stronger producer of toxin than El Tor due to differences in the regulation of expression. It was found that the *P*_{*ctxAB*} of classical prophages present in the Mozambique/Kolkata hybrid strains contains only four copies of the 5'-TTTGGAT-3' tandem repeats usually found in El Tor strains. Interestingly, sequencing of previously uncharacterized *P*_{*ctxAB*} regions of several classical strains revealed that the heptamer repeat copy number does vary in this biotype. Furthermore, unlike the Mozambique hybrid strain, the Kolkata strains carry an intact toxin-linked cryptic (TLC) element along with either pre-CTX prophage (CTX prophage with naturally deleted *ctxAB* genes; Boyd *et al.*, 2000) or free RS1 elements arranged in tandem in the *dif* sites of their large chromosomes. The entire free RS1 region of a Kolkata hybrid strain was sequenced, and this indicated the presence of the *rstR*^{ET} allele as well as the *rstC* gene. Altogether, it appears that diverse CTX Φ along with free RS1 may play a critical role in the microevolution of *V. cholerae*, leading to the genesis of distinct genetic variants, which in turn helps the pathogen to evolve further and maintain its pathogenic and epidemic potential in endemic zones. This type of genetic rearrangement leading to the evolution of novel strains may thus hamper the development of immunoprophylactic measures against pathogenic *V. cholerae*.

METHODS

Bacterial strains, oligonucleotides and growth conditions.

Bacterial strains and plasmids used in this study are listed in Table 1. Details of oligonucleotides used for various purposes are given in Table 2. Both *Escherichia coli* (DH5 α strain) and *V. cholerae* cells were routinely grown in Luria broth (LB) at 37 °C with shaking as described previously (Das *et al.*, 2007). Antibiotics were used at the following concentrations unless otherwise indicated: ampicillin, 100 μ g ml⁻¹; streptomycin, 100 μ g ml⁻¹; kanamycin, 50 μ g ml⁻¹. Bacterial cells were maintained at -70 °C in LB containing 20% (v/v) glycerol as described previously (Nandi *et al.*, 2003).

GM₁-ELISA of CT. For detection of CT production by *V. cholerae* strains, cells were grown in two different conditions: (i) in LB medium with 66 mM NaCl at 30 °C with shaking (Miller &

Table 1. *V. cholerae* strains used in this study

Strain*	Serotype (biotype)	Origin	Year of isolation/country	Reference
O395	O1 (classical)	Clinical	1964/India	Nandi <i>et al.</i> (2003)
569B	O1 (classical)	Clinical	1948/India	Nandi <i>et al.</i> (2003)
NIH35A3	O1 (classical)	Clinical	1941/India	Mekalanos (1983)
H218	O1 (classical)	Clinical	Unknown	Choudhury <i>et al.</i> (1994)
V154	O1 (classical)	Clinical	Unknown	Choudhury <i>et al.</i> (1994)
NIH41	O1 (classical)	Clinical	1940/India	Mekalanos (1983)
GP8	O1 (classical)	Clinical	1970/India	Choudhury <i>et al.</i> (1994)
N16961	O1 (El Tor)	Clinical	1975/Bangladesh	Nandi <i>et al.</i> (2003)
VC44	O1 (El Tor)	Clinical	1991/India	Nandi <i>et al.</i> (2003)
VC106	O1 (El Tor)	Clinical	1991/India	Sharma <i>et al.</i> (1997)
B33	O1 (El Tor)	Clinical	2004/Mozambique	Das <i>et al.</i> (2007)
VCE232	O4	Environmental	1988/India	Maiti <i>et al.</i> (2006)

*All strains possessed the CT gene.

Mekalanos, 1988); and (ii) in AKI medium (Iwanaga *et al.*, 1986; Waldor & Mekalanos, 1994). CT present in each culture supernatant was assayed by GM₁-ELISA (Nandi *et al.*, 2003) using pure CT (Sigma-Aldrich) and PBS (10 mM, pH 7.2) as positive and negative controls, respectively.

PCR and other molecular biological techniques. Standard molecular biological methods (Sambrook *et al.*, 1989) were followed for chromosomal and plasmid DNA preparations, electroelution of DNA fragments, restriction enzyme digestion, DNA ligation, bacterial transformation and agarose gel electrophoresis, unless stated

otherwise. All restriction enzymes and nucleic acid-modifying enzymes were purchased from New England Biolabs (NEB), and were used essentially as directed by the manufacturer. Primers used in this study were either reported earlier or were designed using the whole-genome sequence information for *V. cholerae* strain N16961 (Heidelberg *et al.* 2000). PCR assays were carried out essentially as described previously (Maiti *et al.*, 2006).

Cloning and DNA sequencing. For cloning of the free RS1 element present in the genome of *V. cholerae* strain VC44 (Table 1), the desired large chromosomal region of about 6.0 kb in size was PCR-

Table 2. Sequences of primers used in this study

Primer	Sequence (5'-3')	Use	Source or reference
CtxAF2	ATATCGGGCAGATTCTAGAC	PCR	This study
Ig1F	GAGCCTGTGACACTCACCTTGAT	Sequencing	Maiti <i>et al.</i> (2006)
RstAR	GCATATGGAACCGACCAAGCAAGAT	Sequencing	Maiti <i>et al.</i> (2006)
RSR	GCTCAGTCAATGCCTTGAGTTG	PCR, sequencing	Maiti <i>et al.</i> (2006)
RstBF	GTGGTGTGTTCCCTGTTTGAC	Sequencing	Maiti <i>et al.</i> (2006)
RstR ^{ET} F	GATCTCCAGAGCATCTGCG	Sequencing	This study
ErstR-R	TCGAGTTGTAATTCATCAAGAGTG	PCR	This study
RstCF	GGATGTTTACGATAGCCTAG	PCR, sequencing	This study
ChrIdF	GCACAAATCGAAGATCAAC	PCR,	This study
ChrIdR	CATACTTAACATTGTCAAATG	PCR	This study
RstBR	CGAGATTTTCATATTAATCACC	PCR, sequencing	This study
TlcF	GATTGTGCGTCTTGCAATTTAGG	PCR	Maiti <i>et al.</i> (2006)
TlcR	GTGAATAAATCAGGTGTAATGTGCG	PCR	Maiti <i>et al.</i> (2006)
RstAF2	CACTCACGCATTAAGTGG	Sequencing	This study
CepR2	CGCCTCGTTCCAGACATAAT	PCR	Maiti <i>et al.</i> (2006)
Ig1R	GCCTTAATGCGTATTATACG	Sequencing	This study
XBS1F	TGTGGCGCGTCATTATGTTG	Sequencing	This study
ZtPF	AGCGTTATCGCCTCGTAGAC	PCR, sequencing	This study
CtPR	GCCTAACAAATCCCGTCTGA	PCR, sequencing	This study
RstAF1	GATTACCGTATGGCTCGC	Sequencing	This study
CIR	CAGACGTGGTTTCACAAAAAAGG	PCR	This study
1479F	GTTTCCACATTAGGTGACAAC	PCR	This study
RtxAR	GCTGATAAGCTTCGTCAAGCCAA	PCR	Maiti <i>et al.</i> (2006)

amplified using the primers ChrIdF/ChrIdR (Table 2) and cloned in the vector pDrive (Amp^r Kan^r; Qiagen), generating the recombinant plasmid pKH8G. Since two copies of RS1 sequences were tandemly arranged in the recombinant pKH8G, we faced difficulties in direct automated sequencing of the insert of the clone by primer walking, and the primer sequences used for this purpose are given in Table 2. Therefore, we took the step-wise sequencing of the entire 6.0 kb insert DNA in the following manner. There is a conserved *Bgl*II site in the *rstR* gene of RS1 (Bhadra *et al.*, 1995; Pearson *et al.*, 1993). Since RS1 is 2.7 kb in size, the 6.0 kb insert of pKH8G should carry two copies of the element in tandem. Therefore, digestion of pKH8G with *Bgl*II should definitely generate a 2.7 kb fragment (since there is no *Bgl*II site in the vector pDrive) carrying the entire RS1 element except truncated *rstR* genes from each RS1 copy flanking both the ends of the fragment. This 2.7 kb *Bgl*II fragment was gel-purified and used as a template for sequencing using the following primers: Ig1F, Ig1R, RstAR, RSR, RstBF, RstR^{ET}F, RstCF, RstBR, RstAF2, XBS1F and RstAF1 (Table 2). *Bgl*II-digested pKH8G liner plasmid DNA was also gel-purified and religated, followed by transformation to obtain the clone pKH9A. Sequencing of the insert DNA present in plasmid pKH9A was done by using the universal M13 forward and reverse primers (NEB) and other primers (Table 2). The *P_{ctxAB}* regions of different *V. cholerae* strains were PCR-amplified by using the primers ZtPF/CtPR (Table 2). In each case the desired ~0.6 kb PCR fragment was amplified and gel-purified, followed by sequencing of both the strands using the same pair of primers, ZtPF/CtPR, that were used for PCR amplification of the fragment. Sequencing reactions were performed using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems) essentially as recommended by the manufacturer. The samples were run on an automated DNA sequencer (Genetic Analyzer 3130, Applied Biosystems). In each case, complete sequencing of both strands was done, and to avoid errors in sequencing, each sample was sequenced at least twice.

Nucleotide sequence and other computer-based analysis. DNA sequence data were compiled and analysed by using the DNASIS program (Hitachi Corporation). The National Center for Biotechnology Information (NCBI) BLASTN program version 2.2.19 was used to search homologous sequences in the database ([http://](http://blast.ncbi.nlm.nih.gov/Blast.cgi)

blast.ncbi.nlm.nih.gov/Blast.cgi). The NCBI ORF Finder tool (<http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi>) was used to identify potential ORFs within the sequenced fragment. For designing PCR and other primers, Primer3 software was used (<http://frodo.wi.mit.edu/>). Genomatix software (<http://www.genomatix.de/cgi-bin/dialign/dialign.pl>) was used to align DNA sequences.

RESULTS AND DISCUSSION

Free RS1 and pre-CTX prophage in the large chromosomes of El Tor hybrid strains

Faruque *et al.* (2007) recently reported that the large chromosomal CTX phage integration site is 'empty' in the Mozambique hybrid El Tor strain B33. They have shown that the TLC element (Rubin *et al.*, 1998), which is usually present in two copies in tandem just upstream of integrated CTX prophage, is also deleted in B33. Das *et al.* (2007) have reported that the El Tor strains which were isolated in Kolkata, India, considerably before the emergence of the Mozambique strain, are also hybrid in nature, and like the Mozambique strain carry tandem copies of the classical CTX prophage in their small chromosomes. Since the status of the large chromosomal CTX phage integration site (*dif* site) of Indian hybrid strains is not known, we wished to check the locus using *V. cholerae* strains B33, N16961, O395 and VCE232 (Table 1) as controls. When the TLC-specific primers TlcF/TlcR (Table 2) were used, unlike Mozambique strain B33, both VC44 and VC106 Kolkata isolates (Table 1) gave the desired 2.0 kb PCR fragment, suggesting that they carry a TLC element (data not shown; Fig. 1). Similar PCR analysis using the primers ChrIdF (just upstream of the large chromosomal CTX integration site) and RSR (*rstC* gene) (Fig. 1; Table 2) gave

Fig. 1. Genetic organization (not drawn to scale) of the pre-CTX prophage and RS1 element in the large chromosomes of Kolkata hybrid strains VC106 and VC44. Various genetic regions of the CTX element are as indicated. For simplicity, the *rtxA* gene is shown here as truncated. The large chromosome of Mozambique hybrid strain B33 is free of any CTX element, including *tlc* (Faruque *et al.*, 2007). The number 1479 indicates ORF VC1479 of *V. cholerae* (Heidelberg *et al.*, 2000). Small arrows indicate region-specific PCR primers (Table 2) used in various combinations for determining the arrays of the pre-CTX prophage and RS1 in *V. cholerae* hybrid strains. The RS1 region (~6.0 kb) of VC44, which was PCR-amplified using the forward and reverse primers ChrIdF and ChrIdR (Table 2), respectively, is shown. Broken lines indicate regions deleted in the genomes of the VC44 and B33 strains.

the expected amplicon of ~3.0 kb (Supplementary Fig. S1a), which confirmed that unlike those of B33, the large chromosomal *dif* sites of the Indian hybrid strains were not empty. Rather, they are occupied by at least one RS1 element, the characteristic feature of *V. cholerae* O1 El Tor strains (Davis *et al.*, 2000; Pearson *et al.*, 1993), and thus the strains are indeed hybrid in nature in this respect, since their small chromosomal *dif* sites are occupied by a classical type CTX prophage without an RS1 element (Das *et al.*, 2007; Faruque *et al.*, 2007). It is to be noted that Sharma *et al.* (1997) reported that *V. cholerae* O1 El Tor strains may carry free RS1, although its location is unknown. This result prompted us to map the large chromosomal *dif* sites of VC44 and VC106 in more detail, and for this, extensive PCR analysis (Maiti *et al.*, 2006) was done using various combinations of primers (Fig. 1; Table 2). The PCR results (see Supplementary Fig. S1) allowed us to conclude that while VC44 carries two free RS1 in tandem in the large chromosomal *dif* site, strain VC106 in the same locus contains two copies of RS1 followed by two copies of pre-CTX prophage in tandem (Fig. 1). It is to be noted that like VC44 and B33, strain VC106 also carries two tandem copies of the classical CTX prophage in its small chromosome (Supplementary Fig. S1; Das *et al.*, 2007). Thus, comparing all the PCR-based genetic mapping data it seems that hybrid El Tor strains are genotypically different in the large chromosomal CTX genetic element present in the *dif* loci, as shown in Fig. 1. Regarding the step-wise evolution of these hybrid strains, it is plausible that an early progenitor strain carried tandemly arranged multiple copies of the CTX prophage in each of the two chromosomes, after which the *ctxAB* operon was most likely deleted from the large chromosome creating a VC106-like strain (TLC⁺ RS1⁺ pre-CTX prophage⁺). Further deletion of pre-CTX prophages from this strain helped in the evolution of a VC44-like strain (TLC⁺ RS1⁺), which in turn served as a precursor for a B33-like strain with deletion of TLC and RS1 elements from the large chromosome. This conclusion may be further supported by various earlier studies in which it has clearly been indicated that except for the TLC element, all other CTX genetic elements are mobile in nature (Boyd *et al.*,

2000; Faruque *et al.*, 1998, 2002, Udden *et al.*, 2008; Waldor & Mekalanos, 1996). The genesis of the VC44/VC106-like strain differs from the model proposed by Faruque *et al.* (2007) regarding the evolution of the hybrid strain B33. According to those authors a large deletion of the CTX prophage including the TLC element from the *dif* site of a precursor *V. cholerae* strain made the locus defective for further integration of a CTX^{class}Φ and allowed the phage to integrate in the small chromosomal empty site, leading to the evolution of the Mozambique strain; however, this study indicates that the small chromosomal *dif* site was already occupied by the classical CTX prophage in hybrid strains.

Genetic localization of free RS1 in the genome of VC44 prompted us to further investigate its organization; therefore, we amplified the region (~6.0 kb) with the primers ChrIdF and CIR (Table 2), followed by cloning and generation of the recombinant plasmid pKH8G. Sequencing of both the strands of the insert DNA of pKH8G was done as described in Methods, and the sequence was submitted to GenBank (accession number EU860439). The nucleotide sequence of the tandem RS1 elements allowed us to interpret the upstream and downstream as well as RS1–RS1 junctions, as shown in Fig. 2. While there is no 18-base end-repeat sequence in the upstream junction, both RS1–RS1 and the downstream junctions showed the presence of one copy of such a repeat sequence, as shown in Fig. 2. To our knowledge, this is the first report in which the free RS1 element has been localized and its genetic organization, when present in multiple copies in tandem in the chromosome of *V. cholerae*, determined. It has been reported that the RS1 element can propagate horizontally as a filamentous phage, like CTXΦ, exploiting the morphogenesis genes of CTXΦ, and RS1 satellite phages may be produced in *V. cholerae* (Faruque *et al.*, 2002). Since *dif*-like sequences are present in the intergenic regions of the RS1 element, it is possible that upon entry of an RS1 satellite phage into *V. cholerae* cells, it can integrate into the host chromosomal *dif* site by homologous recombination, as proposed by Val *et al.* (2005).

Fig. 2. Schematic representation (not drawn to scale) of tandemly arranged free RS1 in the large chromosomal *dif* site of VC44. *attL* and *attR* represent 5' and 3' junctions, respectively. Nucleotides in bold and italic type indicate chromosomal and RS1-derived sequences, respectively. The XerC and XerD binding sites along with the *attL* and *attR* sequences are also shown (see text for details). Only *attR* and not *attL* contains a functional *dif* site with a 6 bp spacer present between the XerC and XerD binding sequences. The triangle depicts the 18 bp end repeat (ER) sequence (5'-CCTTAGTGGCGGCA-3') and an almost similar sequence is also present in the 5' region of the *attR* sequence marked as ER. The entire sequence of the region is available in GenBank under accession number EU860439.

The promoter sequence of the *ctxAB^{class}* operon of hybrid strains

Since CT is the most critical virulence factor of *V. cholerae* and the regulation of its expression is the subject of extensive research, we wished to analyse the P_{ctxAB} of the operon encoding this toxin. Several earlier studies have shown that among various genetic differences between the classical and El Tor biotypes, the difference in nucleotide sequence of P_{ctxAB} is distinctive. While in classical strains there are seven to eight copies of a heptamer (5'-TTTTGAT-3') direct repeat sequence present in tandem, in the El Tor biotype the same region usually contains three to four copies of this sequence. Experimental evidence indicates that the heptamer copy number in P_{ctxAB} is critical for expression of CT (Mekalanos *et al.*, 1983; Hung & Mekalanos, 2005). The presence of more copies of the heptamer sequence in classical strains leads to CT production in large quantities, as shown elsewhere (Mekalanos *et al.*, 1983; Hung & Mekalanos, 2005; Miller and Mekalanos, 1984). Furthermore, Hung & Mekalanos (2005) constructed an El Tor strain in which the P_{ctxAB} sequence carrying four copies of heptad repeats was replaced by the P_{ctxAB} of the classical strain 569B containing eight copies of the repeat sequences, and they found high expression of CT by the recombinant El Tor strain in LB medium at 37 °C in the presence of sodium cholate, which is a non-permissive condition for the El Tor biotype, and according to the authors this correlated with their earlier finding (Mekalanos *et al.*, 1983) that the difference in copy number of the heptad repeats in the classical strains

leads them to be higher producers of CT than the El Tor biotype. It may be argued that since the Mozambique and Indian hybrid El Tor strains carry CTX^{class} prophage (Faruque *et al.*, 2007; Das *et al.*, 2007), their P_{ctxAB} should contain seven to eight copies of heptamer repeats, at least according to earlier reports. To examine this, 600 bp DNA fragments containing the entire P_{ctxAB} regions of several hybrid El Tor strains of Indian as well as Mozambique origin along with control strains (O395, 569B and N16961; Table 1) were PCR-amplified using *zot* gene (the gene preceding the *ctxAB* operon)-specific forward and *ctxA* gene-specific reverse primers, ZtPF and CtPR (Table 2), respectively, followed by direct sequencing of both strands of the amplified products. Analysis of these sequences revealed interesting results: the hybrid strains irrespective of their source of isolation showed only four copies of heptamer repeat sequences in the P_{ctxAB} regions, indicating that the CT regulatory regions of these unusual strains are like those of canonical El Tor strains and not similar to those of classical strains, which usually carry seven to eight such repeats. This result prompted us to compare the nucleotide sequences of the P_{ctxAB} regions of several authentic classical strains so far used in various studies (Table 1). In contrast to our previous belief, it was found that seven to eight copies of heptamer repeats are not the signature sequence of the P_{ctxAB} of classical biotype strains, and it may in fact vary from four to eight copies. The features of the P_{ctxAB} regions of various *V. cholerae* strains examined in this study, including the control strains, are schematically shown in Fig. 3. It is noteworthy that apart

Fig. 3. Organization of heptamer repeats in the P_{ctxAB} regions in classical, El Tor and hybrid strains. Tandemly arranged direct heptamer repeat sequences (5'-TTTGTAT-3'), represented by black arrows, are present upstream of the -35 promoter sequence. The classical strains (569B, O395, V154, H218, NIH35A3 and GP8) carry higher number of repeats, except for strain NIH41, which carries only four such repeats, like the El Tor strains and the hybrid strains B33, VC44 and VC106. Other characteristics of the P_{ctxAB} region are also shown, such as the -35 region, the -10 region, the transcription start site (+1) and the initiation codon ATG. The solitary heptamer repeat is present just upstream of the -10 region. The P_{ctxAB} sequences of the classical strains 569B and O395 were known previously, and their accession numbers are X58785 and CP000627, respectively. The sequences of the P_{ctxAB} regions of the classical strains V154, NIH41, GP8, NIH35A3 and H218 were determined in this study and their nucleotide accession numbers are FJ641042, FJ641045, FJ641046, FJ641047 and FJ641048, respectively. Since the nucleotide sequence (accession number FJ641043) of the P_{ctxAB} region of B33 is identical to those of VC44 (accession number FJ641044), VC106 (FJ660421) and N16961 (Heidelberg *et al.*, 2000), the latter strains are not shown.

Fig. 4. CT production by hybrid El Tor strains. *V. cholerae* cells were grown either in (a) AKI condition or in (b) LB medium containing 66 mM NaCl at pH 6.5 and 30 °C, as indicated. The hybrid strains VC44, VC106 and B33 produced maximum CT under the AKI condition only. However, VC106 also produced a substantial amount of CT in LB medium. The N16961 (El Tor) and O395 (classical) strains were used as controls for AKI and LB culture conditions. The results are the mean of at least three independent experiments.

from the tandem arrangement of multiple copies of the heptamer repeats, there is a solitary 5'-TTTTGAT-3' repeat present in between the -35 and -10 regions, irrespective of the strain analysed (Fig. 3). Currently no information is available about the role of this solitary heptad in *ctxAB* operon expression; however, since this repeat is absolutely conserved it appears to be crucial in the expression of CT in *V. cholerae*.

Condition for CT production by hybrid El Tor strains

To see whether changes in the location and variation in the copy number of the CTX^{class} prophage, or the presence of a pre-CTX or RS1 element in the genome of hybrid El Tor strains have any effect on CT production, we used two different culture conditions (see Methods): LB medium at pH 6.5 and incubation of the culture at 30 °C with shaking (a CT-producing condition for classical strains); and the AKI condition, where the bacterial culture is initially grown statically followed by shaking at 37 °C (a CT-producing condition for El Tor strains). When these experiments were done, as expected, the hybrid El Tor strains produced a large amount of CT ($1.96\text{--}2.6\ \mu\text{g ml}^{-1}$) under the AKI condition (Fig. 4a), whereas in LB medium CT production was very poor (Fig. 4b). However, among the hybrid El Tor strains tested, VC106 seemed to be a strong CT producer (Fig. 4a), and even in LB medium, a non-permissive condition, the strain produced a substantial amount of CT (Fig. 4b). This result is consistent with earlier reports (Iwanaga *et al.*, 1986; Waldor & Mekalanos, 1994). Although the hybrid strains carry a classical CTX prophage, it seems that regulation of CT production under *in vitro* conditions is highly similar to that of the 'pure' El Tor biotype.

Conclusion

Analysis of representative hybrid El Tor strains isolated in Kolkata, India, in 1992 suggested that the evolved

strains were unique, since both of their chromosomes are occupied by CTX genetic elements. Since the Mozambique hybrid strain evolved in 2004, and the Kolkata hybrid strains were isolates of 1992, the latter strains may thus be the progenitors of the former strain. Recently Chatterjee *et al.* (2009) have also provided evidence in favour of this conclusion by using ribotyping and other molecular methods. The small chromosomal arrangement of the classical CTX prophage is identical in the Mozambique and Kolkata hybrid El Tor strains (Faruque *et al.*, 2007; Das *et al.*, 2007), which also supports their ancestry. While the large chromosomal CTX genetic element of the Mozambique strain is deleted, the Kolkata hybrid strains carry either a pre-CTX prophage or a free RS1 element in the same location, and thus step-wise deletion of the pre-CTX prophage and free RS1 along with the *tlc* element probably helped in the genesis of the Mozambique strain. The presence of four heptamer repeats in the P_{ctxAB} of classical CTX prophages, carried by the hybrid El Tor strains, indicates that a single and novel type of classical CTX Φ infection had occurred in these strains. Previously it was not known that like that of El Tor, the P_{ctxAB} region of classical strains may carry a low copy number of heptad repeats, which is critical for CT gene expression. However, the natural selection of this type of virulence gene regulatory element, its implications in clinical manifestations, and the epidemiology of cholera certainly demand further investigation.

ACKNOWLEDGEMENTS

We are thankful to Professor Siddhartha Roy for his constant support. The work was financially supported by research grants from the Council of Scientific and Industrial Research (CSIR), Government of India, to R.K.B. K.H. and B.D. are grateful to CSIR for Research Fellowships.

REFERENCES

- Bhadra, R. K., Roychoudhury, S., Banerjee, R. K., Kar, S., Majumdar, R., Sengupta, S., Chatterjee, S., Khetawat, G. & Das, J. (1995). Cholera toxin (CTX) genetic element in *Vibrio cholerae* O139. *Microbiology* **141**, 1977–1983.
- Boyd, E. F., Heilpern, A. J. & Waldor, M. K. (2000). Molecular analyses of a putative CTX Φ precursor and evidence for independent acquisition of distinct CTX Φ s by toxigenic *Vibrio cholerae*. *J Bacteriol* **182**, 5530–5538.
- Chatterjee, S., Patra, T., Ghosh, K., Raychoudhuri, A., Pazhani, G. P., Das, M., Sarkar, B., Bhadra, R. K., Mukhopadhyay, A. K. & other authors (2009). *Vibrio cholerae* O1 clinical strains isolated in 1992 in Kolkata with progenitor traits of the 2004 Mozambique variant. *J Med Microbiol* **58**, 239–247.
- Choudhury, S. R., Bhadra, R. K. & Das, J. (1994). Genomic size and restriction fragment length polymorphism analysis of *Vibrio cholerae* strains belonging to different serovars and biotypes. *FEMS Microbiol Lett* **115**, 329–334.
- Das, B., Halder, K., Pal, P. & Bhadra, R. K. (2007). Small chromosomal integration site of classical CTX prophage in Mozambique *Vibrio cholerae* O1 biotype El Tor strain. *Arch Microbiol* **188**, 677–683.
- Davis, B. M. & Waldor, M. K. (2000). CTX Φ contains a hybrid genome derived from tandemly integrated elements. *Proc Natl Acad Sci U S A* **97**, 8572–8577.
- Davis, B. M., Kimsey, H. H., Chang, W. & Waldor, M. K. (1999). The *Vibrio cholerae* O139 Calcutta bacteriophage CTX Φ is infectious and encodes a novel repressor. *J Bacteriol* **181**, 6779–6787.
- Davis, B. M., Moyer, K. E., Boyd, E. F. & Waldor, M. K. (2000). CTX prophages in classical biotype *Vibrio cholerae*: functional phage genes but dysfunctional phage genomes. *J Bacteriol* **182**, 6992–6998.
- Davis, B. M., Kimsey, H. H., Kane, A. V. & Waldor, M. K. (2002). A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. *EMBO J* **21**, 4240–4249.
- Faruque, S. M., Asadulghani, Abdul Alim, A. R. M., Albert, M. J., Nasirul Islam, K. M. & Mekalanos, J. J. (1998). Induction of the lysogenic phage encoding cholera toxin in naturally occurring strains of toxigenic *Vibrio cholerae* O1 and O139. *Infect Immun* **66**, 3752–3757.
- Faruque, S. M., Asadulghani, Kamruzzaman, M., Nandi, R. K., Ghosh, A. N., Nair, G. B., Mekalanos, J. J. & Sack, D. A. (2002). RS1 element of *V. cholerae* can propagate horizontally as a filamentous phage exploiting the morphogenesis genes of CTX Φ . *Infect Immun* **70**, 163–170.
- Faruque, S. M., Tam, V. C., Chowdhury, N., Diraphat, P., Dzieman, M., Heidelberg, J. F., Clemens, J. D., Mekalanos, J. J. & Nair, G. B. (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.
- Heidelberg, J. F., Eisen, J. A., Nelson, W. C., Clayton, R. A., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D. & other authors (2000). DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**, 477–484.
- Huber, K. E. & Waldor, M. K. (2002). Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* **417**, 656–659.
- Hung, D. T. & Mekalanos, J. J. (2005). Bile acids induce cholera toxin expression in *Vibrio cholerae* in a ToxT-independent manner. *Proc Natl Acad Sci U S A* **102**, 3028–3033.
- Iwanaga, M., Yamamoto, K., Higa, N., Ichinose, Y., Nakasone, N. & Tanabe, M. (1986). Culture conditions for stimulating cholera toxin production by *Vibrio cholerae* O1 El Tor. *Microbiol Immunol* **30**, 1075–1083.
- Kaper, J. B., Morris, J. G., Jr & Levine, M. M. (1995). Cholera. *Clin Microbiol Rev* **8**, 48–86.
- Lee, J. H., Han, K. H., Choi, S. Y., Lucas, M. E., Mondlane, C., Ansaruzzaman, M., Nair, G. B., Sack, D. A., von Seidlein, L. & other authors (2006). Multilocus sequence typing (MLST) analysis of *Vibrio cholerae* O1 El Tor isolates from Mozambique that harbour the classical CTX prophage. *J Med Microbiol* **55**, 165–170.
- Maiti, D., Das, B., Saha, A., Nandy, R. K., Nair, G. B. & Bhadra, R. K. (2006). Genetic organization of pre-CTX and CTX prophages in the genome of an environmental *Vibrio cholerae* non-O1, non-O139 strain. *Microbiology* **152**, 3633–3641.
- Mekalanos, J. J. (1983). Duplication and amplification of toxin genes in *Vibrio cholerae*. *Cell* **35**, 253–263.
- Mekalanos, J. J., Swartz, D. J., Pearson, G. D., Harford, N., Groyne, F. & de Wilde, M. (1983). Cholera toxin genes: nucleotide sequence, deletion analysis and vaccine development. *Nature* **306**, 551–557.
- Miller, V. L. & Mekalanos, J. J. (1984). Synthesis of cholera toxin is positively regulated at the transcription level by *toxR*. *Proc Natl Acad Sci U S A* **81**, 3471–3475.
- Miller, V. L. & Mekalanos, J. J. (1988). A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J Bacteriol* **170**, 2575–2583.
- Nair, G. B., Faruque, S. M., Bhuniyan, N. A., Kamaruzzaman, M., Siddique, A. K. & Sack, D. A. (2002). New variants of *Vibrio cholerae* O1 biotype El Tor with attributes of the classical biotype from hospitalized patients with acute diarrhoea in Bangladesh. *J Clin Microbiol* **40**, 3296–3299.
- Nandi, S., Maiti, D., Saha, A. & Bhadra, R. K. (2003). Genesis of variants of *Vibrio cholerae* O1 biotype El Tor: role of CTX Φ array and its position in the genome. *Microbiology* **149**, 89–97.
- Olsvik, O., Wahlberg, J., Petterson, B., Uhlén, M., Popovic, T., Wachsmuth, I. K. & Fields, P. I. (1993). Use of automated sequencing of polymerase chain reaction-generated amplicons to identify three types of cholera toxin subunit B in *Vibrio cholerae* O1 strains. *J Clin Microbiol* **31**, 22–25.
- Pearson, G. D., Woods, A., Chiang, S. L. & Mekalanos, J. J. (1993). CTX genetic element encodes a site-specific recombination system and an intestinal colonization factor. *Proc Natl Acad Sci U S A* **90**, 3750–3754.
- Rubin, E. J., Lin, W., Mekalanos, J. J. & Waldor, M. K. (1998). Replication and integration of *Vibrio cholerae* cryptic plasmid linked to the CTX prophage. *Mol Microbiol* **28**, 1247–1254.
- Sack, D. A., Sack, R. B., Nair, G. B. & Siddique, A. K. (2004). Cholera. *Lancet* **363**, 223–233.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Sharma, C., Nair, G. B., Mukhopadhyay, A. K., Bhattacharya, S. K., Ghosh, R. K. & Ghosh, A. (1997). Molecular characterization of *Vibrio cholerae* O1 biotype El Tor strains isolated between 1992 and 1995 in Calcutta, India: evidence for the emergence of a new clone of the El Tor biotype. *J Infect Dis* **175**, 1134–1141.
- Trucksis, M., Michalski, J., Deng, Y. K. & Kaper, J. B. (1998). The *Vibrio cholerae* genome contains two unique circular chromosomes. *Proc Natl Acad Sci U S A* **95**, 14464–14469.
- Udden, S. M. N., Zahid, M. S. H., Biswas, K., Ahmad, Q. S., Cravioto, A., Nair, G. B., Mekalanos, J. J. & Faruque, S. M. (2008). Acquisition of classical CTX prophage from *Vibrio cholerae* O141 by El Tor strains aided by lytic phages and chitin-induced competence. *Proc Natl Acad Sci U S A* **105**, 11951–11956.

Val, M. E., Bouvier, M., Campos, J., Sherratt, D., Cornet, F., Mazel, D. & Barre, F. X. (2005). The single-stranded genome of phage CTX is the form used for integration into the genome of *Vibrio cholerae*. *Mol Cell* **19**, 559–566.

Waldor, M. K. & Mekalanos, J. J. (1994). Emergence of a new cholera pandemic: molecular analysis of virulence determinants in *Vibrio cholerae* O139 and development of a live vaccine prototype. *J Infect Dis* **170**, 278–283.

Waldor, M. K. & Mekalanos, J. J. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**, 1910–1914.

Waldor, M. K., Rubin, E. J., Pearson, G. D., Kimsey, H. & Mekalanos, J. J. (1997). Regulation, replication, and integration functions of the *Vibrio cholerae* CTXΦ are encoded by region RS2. *Mol Microbiol* **24**, 917–926.

Edited by: P. H. Everest