Microbiology and Immunology

Microbiol Immunol 2018; 62: 635–650
doi: 10.1111/1348-0421.12648

ORIGINAL ARTICLE

Comparative analyses of CTX prophage region of *Vibrio cholerae* seventh pandemic wave 1 strains isolated in Asia

Tho Duc Pham[1], Tuan Hai Nguyen[1], Hanako Iwashita[2], Taichiro Takemura[1,3], Kouichi Morita[1,4] and Tetsu Yamashiro[2]

[1]Leading Program, Graduate School of Biomedical Sciences, Nagasaki University, Nagasaki, [2]Department of Bacteriology, Graduate School of Medicine, University of the Ryukyus, Nishihara, Okinawa and [3]Vietnam Research Station, [4]Department of Virology, Institute of Tropical Medicine, Nagasaki University, Nagasaki, Japan

ABSTRACT

***Vibrio cholerae* O1 causes cholera, and cholera toxin, the principal mediator of massive diarrhea, is encoded by *ctxAB* in the cholera toxin (CTX) prophage. In this study, the structures of the CTX prophage region of *V. cholerae* strains isolated during the seventh pandemic wave 1 in Asian countries were determined and compared. Eighteen strains were categorized into eight groups by CTX prophage region-specific restriction fragment length polymorphism and PCR profiles and the structure of the region of a representative strain from each group was determined by DNA sequencing. Eight representative strains revealed eight distinct CTX prophage regions with various combinations of CTX-1, RS1 and a novel genomic island on chromosome I. CTX prophage regions carried by the wave 1 strains were diverse in structure. *V. cholerae* strains with an area specific CTX prophage region are believed to circulate in South-East Asian countries; additionally, multiple strains with distinct types of CTX prophage region are co-circulating in the area. Analysis of a phylogenetic tree generated by single nucleotide polymorphism differences across 2483 core genes revealed that *V. cholerae* strains categorized in the same group based on CTX prophage region structure were segregated in closer clusters. CTX prophage region-specific recombination events or gain and loss of genomic elements within the region may have occurred at much higher frequencies and contributed to producing a panel of CTX prophage regions with distinct structures among *V. cholerae* pathogenic strains in lineages with close genetic backgrounds in the early wave 1 period of the seventh cholera pandemic.**

Key words cholera toxin prophage, phylogenetic tree, seventh pandemic wave 1, *Vibrio cholerae* O1.

Vibrio cholerae, gram-negative rod shaped bacteria, are widely distributed in estuary environments where fresh river water and seawater mix (1). More than 200 distinct serogroups have been identified as bacterial somatic antigens. Of these, *V. cholerae* belonging to serogroups O1 and O139 are recognized as the causative agents of cholera, which is characterized by massive diarrhea and subsequent dehydration (2). Serogroup O1 is classified into two biotypes: classical and El Tor, on the basis of results of phenotype tests such as hemolysis of sheep erythrocytes, agglutination of chicken erythrocytes, the Voges-Proskauer test, sensitivity to polymyxin B and specific phages (3). Recently, genotypes of pathogenic genes including cholera toxin subunit B (*ctxB*) (4), toxin co-regulated pilus (*tcpA*) (5) and repeat sequence transcriptional regulator (*rstR*) (6) have also been used to differentiate biotypes.

Correspondence
Tetsu Yamashiro, Department of Bacteriology, Graduate School of Medicine, University of the Ryukyus, 207 Uehara, Nishihara, Okinawa 903-0215 Japan. Tel.: +81 98 8951125; fax: +81 98 8951408; email: tyamashi@med.u-ryukyu.ac.jp

Received 7 June 2018; revised 2 September 2018; accepted 4 September 2018.

List of Abbreviations: CTX, cholera toxin; *ctx A* and *B*, cholera toxin A and B (gene); ORF, open reading frame; RFLP, restriction fragment length polymorphism; *rst*, repeat sequence transcriptional regulator (gene); RTX, repeats in toxin; SNP, single nucleotide polymorphism; TCP, toxin-coregulated pilus; TLC, toxin-linked cryptic.

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

V. cholerae are vulnerable to taking up exogenous genetic elements and hence are reportedly in dynamic population change (7–9), represented by a replacement of predominant strains from biotype classical to El Tor strains. Seven cholera pandemics have been recorded since the early 19th century. The first six pandemics were attributed to biotype classical, whereas the most recent and ongoing seventh cholera pandemic has been caused by strains belonging to biotype El Tor (3, 10). These two biotypes differ in their pathogenicity and ability to adapt to surrounding environments (3). For example, biotype classical strains induce more severe diarrhea, whereas, biotype El Tor strains are reported to survive better in the environment and to be transmitted more efficiently from human to human (11). There has been an ongoing modest population change among strains belonging to the seventh cholera pandemic El Tor; this has resulted in the emergence of wave 2 and wave 3 strains out of the wave 1 strains (7, 8).

Cholera toxin and TCP play important roles in the pathogenesis of *V. cholerae* and are closely associated with induction of severe watery diarrhea (12). Cholera toxin is encoded by *ctxAB*, which is harbored in a 6.9 kb lysogenic filamentous bacteriophage, CTX phage, composed principally of the two elements of RS2 and a core region. The core region contains ORFs for *cep*, *orfU*, *ace*, *zot*, *ctxA* and *ctxB*, which are essential for the morphogenesis of CTX phage particles and bacterial toxicity, whereas the RS2 region contains ORFs for *rstR*, *rstA* and *rstB*, whose products are required for the replication and regulatory functions of CTX phage (13–15). CTX phage was originally an exogenous genetic element and infects with *V. cholerae* that expresses TCP, the principle intestinal colonization factor of pathogenic *V. cholerae,* which serves as the receptor for CTX phage (13). CTX phage integrates into bacterial chromosomes via a specific *attP* site with a compatible bacterial *dif* site by XerC catalyzed site-specific recombination (16, 17). RS1 is a satellite phage carried principally by the seventh pandemic strains, consisting of *rstR*, *rstA*, *rstB* and *rstC* (18). RS1 integrates into the bacterial genome close to the CTX prophage, exploiting a Xer recombination mechanism (16).

The CTX prophage region, which spans from TLC gene clusters (19) to RTX gene clusters (20), is one of the most important regions for pathogenesis of *V. cholerae*, because it potentially carries a sole or a combination of CTX prophage, RS1 and other genomic elements. The structure of the region is reportedly diverse among the seventh pandemic *V. cholerae* wave 2 and wave 3 strains; however, there is insufficient information on those of wave 1 strains (21).

To expand our understanding about the structure of the CTX prophage region of the seventh cholera pandemic wave 1 strains, 18*V. cholerae* strains isolated from patients in the 1956–1962 seasons in South-East Asian countries and Japan were analyzed and compared.

MATERIALS AND METHODS

Vibrio cholerae strains used in the study

Eighteen *V. cholerae* clinical isolates from Thailand, Malaysia, Indonesia, the Philippines, Taiwan and Japan obtained from 1956 to 1962 that had been stored with glycerol with occasional passages were provided by Dr. Masahiko Ehara (Institute of Tropical Medicine, Nagasaki University, Japan). Information on serogroup, biotype, year and the place of isolation for each strain are presented under anonymized strain identification numbers in Table 1. Before use, all the strains were confirmed by biochemical and serological methods as being *V. cholerae* serogroup O1; additionally, their *ctxB*, *rstR* and *tcpA* genes were sequenced to determine their genotypes (22–24) (Table 1).

Genomic DNA preparation

Genomic DNA was prepared from a 5–10 mL overnight culture of each *V. cholerae* strain in heart infusion broth at 37°C with gently shaking using the appropriate DNeasy Blood & Kits (Qiagen, Tokyo, Japan). Harvested genome DNA was used for PCR, RFLP analysis with Southern blotting, and for sequencing analyses with the Sanger or next generation sequencing methods with Illumina HiSeq. Genomic DNA was also prepared by a standard phenol/chloroform/alcohol procedure for PacBio RS II system sequencing.

RFLP and PCR for determination of the structure of the CTX prophage region of *V. cholerae* strains

RFLP with Southern blotting was performed as described elsewhere (25, 26) to determine the structure of the CTX prophage region spanning from the TLC gene clusters (19) to the RTX gene clusters (20) of the 18 *V. cholerae* strains. Briefly, genomic DNA prepared from each strain was digested with *Bgl*I and *Sac*I, separated by 0.8% agarose gel, and then transferred to positively charged nylon membranes (GE Healthcare, Tokyo, Japan). A DIG DNA Labelling and Detection Kit (Roche, Tokyo, Japan) was used to produce digoxigenin-labeled *ctxA*, *zot* and *rstC* gene specific probes. After prehybridization with a hybridization buffer and blocking reagent (Roche, Tokyo, Japan), the membranes were hybridized overnight with each denatured gene probe, then washed intensively. DNA fragments that hybridized with the probes were detected using anti-digoxigenin antibody in accordance with the manufacturer's

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

Table 1. *Vibrio cholerae* clinical strains used to determine the sequence of CTX prophage s in this study

Strain ID.	Serogroup; Biotype; Serotype	Genotypes (*ctxB; rstR; tcpA*)	Year, place of isolation
A1M	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1956, Bangkok Thailand
J6	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Sarawak Malaysia
J9	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Sarawak Malaysia
C1	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Sulawesi Indonesia
C2	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Sulawesi Indonesia
C7	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Sulawesi Indonesia
P2	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Philippines
P3	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Philippines
P4	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Philippines
P7	O1; El Tor; Inaba	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Philippines
P16	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Philippines
P18	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Philippines
P31	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1961, Philippines
193	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1962, Taiwan
341	O1; El Tor; Ogawa	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1962, Taiwan
T10	O1; El Tor; Inaba	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1962, Taiwan
T100	O1; El Tor; Inaba	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1962, Taiwan
M25	O1; El Tor; Inaba	ctxB3; $rstR^{ET}$; $tcpA^{ET}$	1962, Moji Japan

instructions. Additionally, all the 18 *V. cholerae* strains were examined by PCR with a series of primers (27) obtain more information about the presence and arrangement of different types of CTX prophage regions. Each strain was categorized into a group on the basis of its RFLP and PCR profiles and a representative strain was chosen from each group for sequencing.

Amplification of CTX prophage region

The CTX prophage region, which spans from the 3′-end portion of the TLC gene cluster corresponding to a primer TLC3F sequence to the 5 ′-end portion of the RTX gene cluster corresponding to a primer RTX5R sequence (Table S1) (27), of each representative *V. cholerae* strain was amplified from the bacterial genome by long range PCR using Prime STAR Max DNA polymerase (Takara, Tokyo, Japan) in accordance with the manufacturer's instructions. A set of primer pairs (Table S1) was chosen for these reactions on the basis the structure of the CTX prophage region. If necessary, a nested PCR procedure was applied to prepare fragments that carry a unique gene constellation for sequencing. El Tor strains are considered to not carry CTX prophage on chromosome II; however, a region corresponding to CIIF and CIIR primer sequence (Table S1) (27) on chromosome II was amplified and analyzed.

DNA sequencing to determine CTX prophage region

Genomic libraries were prepared using Nextera XT DNA kits (Illumina, San Diego, CA, USA). Paired-end sequencing was performed using a HiSeq Reagent Kit v3 (600 cycles) on the Illumina HiSeq platform. Quality trimming and filtering of the obtained sequence reads were performed using CLC Genomics Workbench v8.5.1 (CLC bio, Aarhus, Denmark) with the following parameters: quality limit = 0.001; adapters trimming = Yes; remove 5′ terminal nucleotides = No; remove 3′ terminal nucleotides = No; and discard reads below length = 25. If necessary, a PacBio RS II was used to estimate the best model for a CTX prophage region. After DNA isolation, purified genomic DNA was randomly fragmented using G-Tube (Zageno, Boston, MA, USA) to a nominal size of 10–12 kb and used for single molecule real time sequencing bell ligation, which was carried out on a PacBio RS II using P5-C3 chemistry, MagBead loading and stage start instrument operation protocols. CLC Genome Finishing Module software (CLC bio) was used for analyses.

Phylogenetic analysis

A maximum-likelihood phylogenetic tree of 32 *V. cholerae* strains was constructed based on SNP differences across 2483 core genes, excluding likely recombination events, with RAxML 8.2.4 with 1000 bootstrap replicates. Of the 18 strains analyzed in the study, 14 were composed of two strains of wave 1 (N16961, RC9), four were wave 2 (MO10, VC51, B33, MJ1236) and four wave 3 (4675, CIRS101, 2010EL1786, 4679) of the seventh cholera pandemic, three were pre-seventh pandemic (M66-2, MAK757, C5), and MS6, a reported "US Gulf Coast" strain (28) was also included

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

(Table S2). The whole genome core genes were identified using conserved gene neighborhood information implemented in Roary. The best fit DNA substitution model considering rate heterogeneity (GTR + GAMMA) was recommended out of 24 candidates by jModelTest v2.1.10 using both Akaike Information Criterion and Bayesian Information Criterion values. MS6 was used as an outgroup to root the tree. The phylogenetic trees were visualized and managed using a python Environment for Tree Exploration ETE3.

RESULTS

The 18 investigated strains of serogroup O1 biotype El Tor were categorized as early wave 1 of the seventh pandemic *V. cholerae* strains, because all the strains contained CTX-1 that harbored $rstR^{El\ Tor}$ and *ctxB3* (Table 1) on chromosome I. The 18 *V. cholerae* strains were categorized into eight groups on the basis of RFLP with Southern blotting and PCR profile as described in the Materials and Methods. One strain was chosen from each groups as a representative of the other strains categorized in that group, and sequenced.

Determination of CTX prophage region of *V. cholerae* strain 341, a representative of group ET-1

Six strains had a similar RFLP profile and were categorized into a group designated as ET-1 (Table 2, Fig. S1). Database information on a reference strain indicated that *Bgl* I specific digestion sites are estimated to be located near the 3′ end of the TLC gene cluster, at a site in *zot* gene, and at a site near the 3′ end of the RTX gene cluster in the CTX prophage region. When the genome of each strain in the group was digested with *Bgl* I, two fragments (20 and 8.5 kb in size) and one fragment (20 kb) were visualized when *zot*- and *ctxA*-specific probes, respectively, were used (Table 2). *Sac* I specific digestion sites are reportedly located at a site upstream of TLC and *rstC*, and a site close to the 3′ end of RTX in the CTX prophage. When the genome of each strain was digested with *Sac* I, an 11 kb fragment was visualized when either *zot*- or *ctxA*-specific probes were used, whereas three fragments (11, 4.8, and 3.5 kb in size) were visualized when an *rstC*-specific probe was used (Table 2). These data indicate that there were at least one copy of CTX prophage and multiple copies of *rstC*-like ORFs in the genome. Th ePCR profile of each strain predicted that a satellite phage RS1, consisting of $rstR^{El\ Tor}$–*rstA*–*rstB*–*rstC*, was flanking the TLC and that *ctxA* and *ctxB* were present but apart from the RTX on chromosome I (Table 2). To determine the structure and sequence of the CTX prophage region of strains in the group, strain 341 was chosen as a representative and subjected to DNA sequencing. Long range PCR with primer pairs TLC3F and rstC2, rstC1 and ctxBR, and ctxBF and RTX5R was performed on the genome to produce 2.8 kb, 7 kb and 8 kb products, respectively. The 8 kb-product was further fragmented into three fragments; a 4 kb-fragment with ctxBF and Frp16_CIR1 primers; a 2.8 kb-fragment with rstC1 and FrP16_CIR2 primers; and a 2 kb-fragment with FrP16_CIF1 and RTX5R primers. These primers were synthesized in reference to a preceding study or by gene walking in this study. Five fragments, estimated to carry a unique-gene constellation on chromosome I, were sequenced and it was found that strain 341 contains a single copy of CTX prophage harboring $rstR^{El\ Tor}$ and *ctxB3* genotypes, categorized as CTX-1, which is genetically identical to that carried by strain N16961, a reference strain of *V. cholerae* biotype El Tor (Fig. 1). A single copy of RS1 was present on chromosome I, but found to contain five nucleotide substitutions in *rstA* when compared with that of N16961 (Table 3) and designated as *RS1. These five SNPs were not associated with amino acid substitutions. In addition, two copies of a 4 kb genomic island, designated as VCET1_GI, were identified flanking the RTX (Fig. 1). Taken together, the best estimated model for the CTX prophage region of strain 341 is "TLC–*RS1–CTX-1–VCET1_GI–VCET1_GI–RTX" (approximately 15.8 kb in size) on chromosome I (Fig. 1a). Long range PCR with a pair of chromosome II-specific primer set CIIF and CIIR failed to amplify a typical CTX prophage-like cluster from the genome; however, an approximately 4 kb product was amplified, and identified as the "VCET1_GI" by sequencing (Fig. 1a).

The VCET1_GI, either on chromosome I or II, was determined to carry seven ORFs, designated as ET1_207, ET1_192, ET1_528, ET1_1140, ET1_279, ET1_351 and ET1_222 (Fig. 1b). Blast search revealed that two ORFs at the 5′-end, ET1_207 and ET1_192; and three ORFs at the 3′-end, ET1_279, ET1_351, and ET1_222, are homologous to the corresponding ORFs of MS6CTXAGI prophage (28) with 99%–100% identities, and also to those of "novel hybrid RS1" (29) with 97%–99% identities. In addition, the ORF ET1_222 is a homologue of *rstC* of strain N16961 with amino acid substitutions from Asp to Ser at position 54, and from Gly to Lys at 67. Of the two center ORFs, ET_1140 is homologous to *rstA1*, a corresponding gene of the "novel hybrid RS1" with 97% identity; however, ORF ET1_528 is unique, no homologue sequences having been reported. In the Illumina HiSeq analysis on strain 341, map reads to strain N16961 predicted the presence of a solitary CTX-1 with an elevated coverage rate at *rstC* gene-like sequence (data not shown) that is probably attributable to the presence of a single copy of genuine *rstC* of CTX-1 on chromosome I, and three copies of ET1_222, the *rstC* homologue in VCET1_GI, on both chromosomes I and II.

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

Determination of CTX prophage region of *V. cholerae* strain J6, a representative of group ET-2

Two strains showed a similar RFLP profile and were categorized into the same group, designated as ET-2 (Table 2, Fig. S2). The profile indicated that the strains carry multiple copies of CTX prophage and at least one copy of *rstC*-like ORF in the genome (Table 2). The PCR profile of each strain indicated that RS1 is present, but apart from TLC and RTX. El Tor type RS2, *ctxA* and *B* genes were estimated to lie flanking the TLC and RTX, respectively, on chromosome I. When CIIF and CIIR primers were applied on the genome, a non-specific amplicon was produced (Table 2); however, no genes associated with CTX prophage were present on chromosome II. Strain J6 was chosen as a representative of the group and sequenced to determine the structure and sequence of the CTX prophage region. Primer pairs TLC3F and rstC2, rstC1 and ctxAR, and ctxAF and RTX5R were used on the genome to produce 10 kb, 7 kb, and 1.4 kb products, respectively. Further, the 10 kb-product was fragmented into two fragments; a 4 kb-fragment with TLC3F and NCorfU1R primers; and a 7 kb-fragment with NCorfU1F and rstC2 primers. Four fragments were sequenced and it was found that strain J6 contains two copies of CTX-1 on chromosome I. A single copy of *RS1 identical to that of strain 341 and containing five nucleotide substitutions (Table 3), lay between two CTX-1. In HiSeq analysis on strain J6, the map reads to strain N16961 predicted multiple copies of CTX-1 with a sole RS1 in the genome. Taken together, the best estimated model for the CTX prophage region of strain J6 is "TLC–CTX-1–*RS1–CTX-1–RTX" (approximately 16 kb) on chromosome I and no CTX prophage-associated genes on chromosome II (Fig. 2).

Determination of CTX prophage region of strain C1, a representative of group ET-3

The RFLP profile of strain C1 was unique so it was categorized as an independent group, ET-3 (Table 2, Fig. S3). The analysis predicted that there are multiple copies of CTX prophage and *rstC*-like ORFs on the genome. Th ePCR profile indicated that RS1 is present but apart from TLC and RTX. In the analyses, El Tor type RS2 and *ctxA* and *B* were estimated to be present close to TLC, and RTX, respectively. No CTX prophage-associated genes were amplified when a pair of CIIF and CIIR primers was used (Table 2). HiSeq analysis and subsequent map reads to strain N16961 revealed that C1 carries multiple copies of *RS1, which is identical to that of 341, and CTX-1 on chromosome I. When C1 was analyzed by PacBio RS II system with CLC Genome Finishing Module software, several numbers of corrected reads containing [a portion of CTX-1–*RS1–CTX-1–*RS1–a portion of CTX-1] structure were harvested out of 45,973 full genome corrected reads, indicating that three copies of CTX-1 and two copies of *RS1 are present in the CTX prophage region. Taken together, "TLC–CTX-1–*RS1–CTX-1–*RS1–CTX-1–RTX" (approximately 25.6 kb) is estimated as the best model with 84× to 34× sequence coverage of the CTX prophage region on chromosome I. No CTX prophage-associated genes are present on chromosome II (Fig. 3).

Determination of CTX prophage region of *V. cholerae* strain C2, a representative of group ET-4

The RFLP profile of strain C2 differs from the others so it was categorized as an independent group, ET-4 (Table 2, Fig. S4). The analysis predicted the presence of multiple copies of CTX prophage and *rstC*-like ORFs on the genome (Table 2). The PCR profile indicated that RS1 is present but apart from TLC. El Tor type RS2 is present close to the TLC and *ctxA* and *ctxB* are present but apart from the RTX (Table 2). To determine the structure and sequence of CTX prophage region of C2, primer pairs TLC3F and rstC2, rstC1 and ctxBR, and ctxBF and RTX5R were used on the genome to produce 10 kb, 7 kb, and 8 kb products, respectively. The 10 kb-product was further fragmented into two fragments, a 7 kb-fragment with a pair of TLC3F and ctxBR primers and a 3 kb-fragment with ctxBF and rstC2 primers. The 8 kb product was further fragmented into a 4 kb fragment with primer pairs ctxBF and Frp16_CIR1, a 2.8 kb-fragment with rstC1 and FrP16_CIR2 primers, and a 2 kb-fragment with FrP16_CIF1 and RTX5R primers. Finally, six fragments were sequenced and it was found that strain C2 contains two copies of CTX-1 and a single copy of RS1 that lie between two copies of CTX-1 on chromosome I. In addition, two copies of VCET1_GI found in strain 341 were identified in proximity to the RTX. When a CIIF and CIIR primer set was applied on the C2 genome, an approximately 4 kb product was amplified and identified as VCET1_GI by sequencing. In the HiSeq analysis on strain C2, map reads to strain N16961 predicted multiple copies of CTX-1 and of *rstC*-like sequences. Taken together, the best estimated model for CTX prophage region of strain C2 is "TLC–CTX-1–RS1–CTX-1–VCET1_GI–VCET1_GI–RTX" (approximately 22.8 kb) on chromosome I and a single "VCET1_GI" on chromosome II (Fig. 4).

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

Table 2. Grouping of *V. cholerae* strains by profiles of RFLP and PCR in CTX prophage regions

		Size of fragments produced by RFLP, kbp					PCR profile														
		BglI digestion		*SacI* digestion																	
Strain ID.	Group	*zot* probe	*ctxA* probe	*zot* probe	*ctxA* probe	*rstC* probe	TLC3F/ rstAR	TLC3F/ rstC2	TLC3F/ rstRclaR	TLC3F/ rstRETR	rstC1/ rstC2	ctxAF/ rtx5R	ctxBF/ ctxBR	rstRclaF/ rstRclaR	rstRETF/ rstRETR	CIIF/ CIIR	CIIF/ rstAR	CIIF/ rstC2	CIIF/	CIIF/ rstRETR	ctxAF/ CIIR
A1M	ET-1	(20); 8.5	(20)	11	11	(11); 4.8; [3.5]	P	P	N	P	P	N	P	N	P	N	N	N	N	N	N
P7	ET-1	(20); 8.5	(20)	11	11	11; 4.8; 3.5	P	P	N	P	P	N	P	N	P	N	N	N	N	N	N
193	ET-1	(20); 8.5	(20)	11	11	11; 4.8; 3.5	P	P	N	P	P	N	P	N	P	N	N	N	N	N	N
*341	ET-1	(20); 8.5	(20)	11	11	11; 4.8; 3.5	P	P	N	P	P	N	P	N	P	N	N	N	N	N	N
T10	ET-1	(20); 8.5	(20)	11	11	11; 4.8; 3.5	P	P	N	P	P	N	P	N	P	N	N	N	N	N	N
T100	ET-1	(20); 8.5	(20)	11	11	(11); 4.8; 3.5	P	P	N	P	P	N	P	N	P	N	N	N	N	N	N
*J6	ET-2	11.2; 10; 5.4	11.2; 10	22; 12	22; 12	12	P	N	N	P	P	P	P	N	P	P	N	N	N	N	N
J9	ET-2	11.2; 10; 5.4	11.2; 10	22; 12	22; 12	12	P	N	N	P	P	P	P	N	P	P	N	N	N	N	N
*C1	ET-3	11.2; [10]; 5.4	11.2; [10]	(22); 12; 10	(22); 12; 10	12; 10	P	N	N	P	P	P	P	N	P	P	N	N	N	N	N
*C2	ET-4	20; 10; 5.4	20; 10	22; 11	22; 11	11; 4.8; 3.5	P	N	N	P	P	N	P	N	P	N	N	N	N	N	N
*P2	ET-5	11.2; 7; 5.4	11.2; 7	no bands	no bands	no bands	P	N	N	P	N	P	P	N	P	P	N	N	N	N	N
P3	ET-5	(11.2); 7; 5.4	11.2; 7	no bands	no bands	no bands	P	N	N	P	N	P	P	N	P	P	N	N	N	N	N
P4	ET-5	(11.2); 7; 5.4	11.2; 7	no bands	no bands	no bands	P	N	N	P	N	P	P	N	P	P	N	N	N	N	N
P31	ET-5	(11.2); 7; 5.4	11.2; 7	no bands	no bands	no bands	P	N	N	P	N	P	P	N	P	P	N	N	N	N	N
*C7	ET-6	11.2; [7]; 5.4	11.2; [7]	no bands	no bands	no bands	P	N	N	P	N	P	P	N	P	P	N	N	N	N	N
*P16	ET-7	(20); 7; 5.4	(20); 7	no bands	no bands	4.8; [3.5]	P	N	N	P	N	N	P	N	P	N	N	N	N	N	N
P18	ET-7	(20); 7; 5.4	(20); 7	no bands	no bands	4.8; [3.5]	P	N	N	P	N	N	P	N	P	N	N	N	N	N	N
*M25	ET-8	13.6; 8.5	13.6	10	10	10; 4.8	P	P	N	P	P	N	P	N	P	P	N	N	N	N	N

(): Size of fragments with weak intensities.

[]: Size of fragments with augmented intensities.

*: Representative strains and sequenced to determine the structure and DNA sequence of CTX prophage array.

P, positive in producng an amplicon; N, negative in producing an amplicon.

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

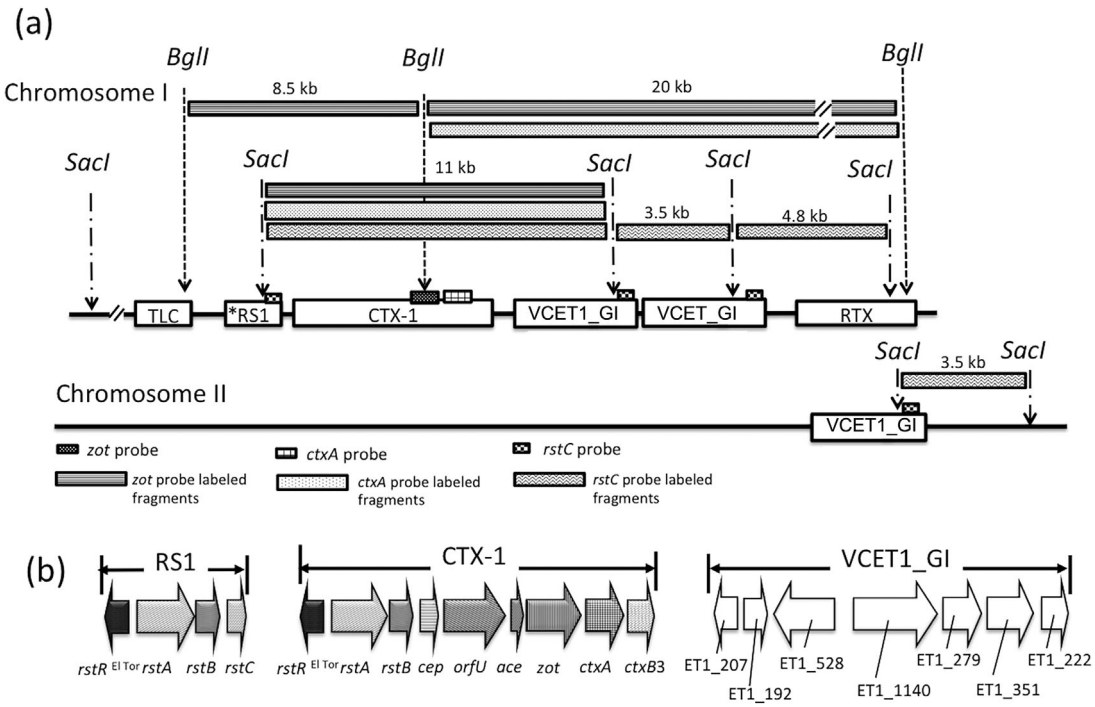


Fig. 1. Estimated CTX prophage region structure of *Vibrio cholerae* strain 341, the representative of group ET-1. Six *V. cholerae* wave 1 strains revealed similar profiles of CTX prophage region-specific RFLP and PCR, and were categorized into a group designated as ET-1. (a) The strain 341 was chosen as a representative and sequenced. The best estimated model, "TLC–*RS1–CTX-1–VCET1_GI–VCET1_GI–RTX" for chromosome I and "VCET1_GI" for chromosome II, is shown. The sites to which each of *zot*, *ctxA* and *rstC* probes bind and produced fragments in the RFLP analysis are shown in the figure. (b) Estimated structure of a novel genomic island VCET1_GI. Structures of satellite phage RS1 and CTX-1 are shown in references. TLC, toxin-linked cryptic gene cluster; CTX-1, CTX prophage harboring $rstR^{El\ Tor}$ and *ctxB3*; *RS1, RS1 which harbors five nucleotide substitutions in *rstA* compared with that of N16961.

Determination of CTX prophage region of *V. cholerae* strain P2, a representative of group ET-5

Four strains showed a similar RFLP profile and were categorized into the same group, designated as ET-5 (Table 2, Fig. S5). Strains in the group were estimated to carry multiple copies of CTX prophage and no *rstC*-like ORFs in the genome (Table 2). The PCR profile predicted that El Tor type RS2, *ctxA* and *ctxB* were close to TLC and RTX, respectively, and that there was no *rstC* in the genome. There were no CTX prophage-associated genes on chromosome II (Table 2). Of the four strains, P2 was chosen as a representative and sequenced. Primer pairs TLC3F and ctxBR, ctxBF and ctxAR, and ctxAF and RTX5R were used on the genome to produce 7 kb, 6.5 kb, and 1.4 kb products, respectively, and sequenced. It was found that P2 contains two copies of CTX-1 on chromosome I. HiSeq analysis and subsequent map reads to N16961 predicted multiple copies of CTX-1 and no *rstC* in the genome. The best estimated model for CTX prophage region of strain P2 is "TLC–CTX-1–CTX-1–RTX" (approximately 13.3 kb) on chromosome I with no CTX prophage-associated genes on chromosome II (Fig. 5).

Determination of CTX prophage region of *V. cholerae* strain C7, a representative of group ET-6

Th eRFLP profile of strain C7 is similar to that of strain P2, the representative strain of group ET-5; however, more 7 kb fragments were produced by *Bgl* I digestion with *zot* and *ctxA* probes than in group ET-5 (Table 2, Fig. S6). The strain was therefore categorized as an independent group, ET-6. The analysis predicted that multiple copies of CTX prophage are present and that there are no *rstC*-like ORFs in the genome (Table 2). The PCR profile indicated that El Tor type RS2, *ctxA* and *ctxB* lie close to TLC and RTX, respectively (Table 2). No CTX prophage-associated genes were amplified when a pair of CIIF and CIIR primers was used (Table 2). HiSeq analysis and subsequent map read to strain N16961 revealed that C7 carries multiple copies of CTX-1 and has no *rstC* in the genome. When C7 was analyzed by PacBio RS II system with CLC Genome Finishing Module, two joined contigs, 2.54 Mb and 1.07

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

Table 3. SNPs in CTX prophage regions

		CTX prophage									RS1		
Strain ID.	Group	*rstR*	*rstA*	*rstB*	*cep*	*orfU*	*ace*	*zot*	*ctxA*	*ctxB*	*rstR*	*rstA*	*rstB*	*rstC*
341	ET-1	$rstR^{El\ Tor}$	–	–	–	–	–	–	–	ctxB3	$rstR^{El\ Tor}$	T579C, T609C, C654T, C639T, T894C	-	-
J6	ET-2	$rstR^{El\ Tor}$	–	–	–	–	–	–	–	ctxB3	$rstR^{El\ Tor}$	T579C, T609C, C654T, C639T, T894C	-	-
C1	ET-3	$rstR^{El\ Tor}$	–	–	–	–	–	–	–	ctxB3	$rstR^{El\ Tor}$	T579C, T609C, C654T, C639T, T894C	-	-
C2	ET-4	$rstR^{El\ Tor}$	–	–	–	–	–	–	–	ctxB3	$rstR^{El\ Tor}$	-	-	-
P2	ET-5	$rstR^{El\ Tor}$	–	–	–	–	–	–	–	ctxB3	NA	NA	NA	NA
C7	ET-6	$rstR^{El\ Tor}$	*G301A	T84C	–	–	–	–	G622A	ctxB3	NA	NA	NA	NA
P16	ET-7	$rstR^{El\ Tor}$	–	–	–	–	–	–	–	ctxB3	NA	NA	NA	NA
M25	ET-8	$rstR^{El\ Tor}$	–	–	–	–	–	–	–	ctxB3	$rstR^{El\ Tor}$	-	-	-

Nucleotide sequences of CTX prophage array of each strain were compared with that of N16961, a reference strain of *V. cholerae* biotype El Tor.

–: no SNPs compared with the corresponding gene of N16961 are present

NA: not applicable, satellite phage RS1 was not present in the genome

rstREl Tor: ref. (23); B3: ref. (21)

*: strain C7 contains four copies of CTX phage arrays in the genome. SNPs were found in three of four CTX prophage arrays of the strain (see text for more detail).

Mb in size, were eventually harvested. The 2.54 Mb joined contig was found to carry a four copy tandem repeat of CTX-1, designated as CTX-1_1, CTX-1_2, CTX-1_3 and CTX-1_4, spanning between TLC and RTX (Fig. 6). SNPs in *rstB* (T84C) and in *ctxA* (G622A) are carried by all four CTX-1; however, SNP in *rstA* (G301A) is carried by CTX-1_1, CTX-1_3 and CTX-1_4 (Table 3). The SNPs in *rstA* and *ctxA* are associated with amino acid substitutions from Asp to Asn, and from Ala to Thr, respectively. Taken together, "TLC–$^{+}$CTX-1–$^{\#}$CTX-1–$^{+}$CTX-1–$^{+}$CTX-1–RTX" (approximately 27 kb) is estimated to be the best model for the CTX prophage region on chromosome I with 80× to 50× sequence coverage. $^{+}$CTX-1 represents CTX-1, carrying three SNPs, namely *rstA* (G301A), *rstB* (T84C) and *ctxA* (G622A), and $^{\#}$CTX-1 represents CTX-1, carrying two SNPs, namely *rstB* (T84C) and *ctxA* (G622A). No CTX prophage-associated genes are present on chromosome II (Fig. 6).

Determination of CTX prophage region of *V. cholerae* strain P16, a representative of group ET-7

Two strains revealed a similar RFLP profile and were categorized into the same group, ET-7 (Table 2, Fig. S7). Strains in the group were estimated to carry multiple copies of CTX prophage and of *rstC*-like ORFs in the genome (Table 2). The PCR profile predicted that El Tor type RS2 is present close to TLC and *ctxA* and *ctxB* are present but apart from RTX (Table 2). Typical *rstC* was not amplified. Strain P16 was chosen as a representative and sequenced. To determine the structure and sequence of the CTX prophage region, primer pairs TLC3F and ctxBR, ctxBF and ctxAR, and ctxAF and RTX5R were used on the genome to produce 7 kb, 6.5 kb and 9 kb products, respectively. The 9 kb-product was further fragmented into a 5 kb-fragment with ctxAF and FrP16_CIR1 primers, a 2.8 kb-fragment with rstC1 and FrP16_CIR2 primers, and a 2 kb-fragment with FrP16_CIF1 and RTX5R primers, and sequenced. It was found that strain P16 contains two copies of CTX-1 on chromosome I. In addition, two copies of VCET1_GI were identified close to RTX. When a CIIF and CIIR primer set was applied on the P16 genome, an approximately 8 kb product was amplified. The product was fragmented into three fragments and sequenced to identify two copies of VCET1_GI lying in tandem. HiSeq analysis and subsequent map read to N16961 predicted multiple copies of CTX-1 and of *rstC*-like sequences, as for strains 341 and C2. Taken together, the best estimated model for the CTX prophage region of strain P16 is "TLC–CTX-1–CTX-1–VCET1_GI–VCET1_GI–RTX" (approximately 20 kb) on chromosome I and "VCET1_GI–VCET1_GI" on chromosome II (Fig. 7).

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

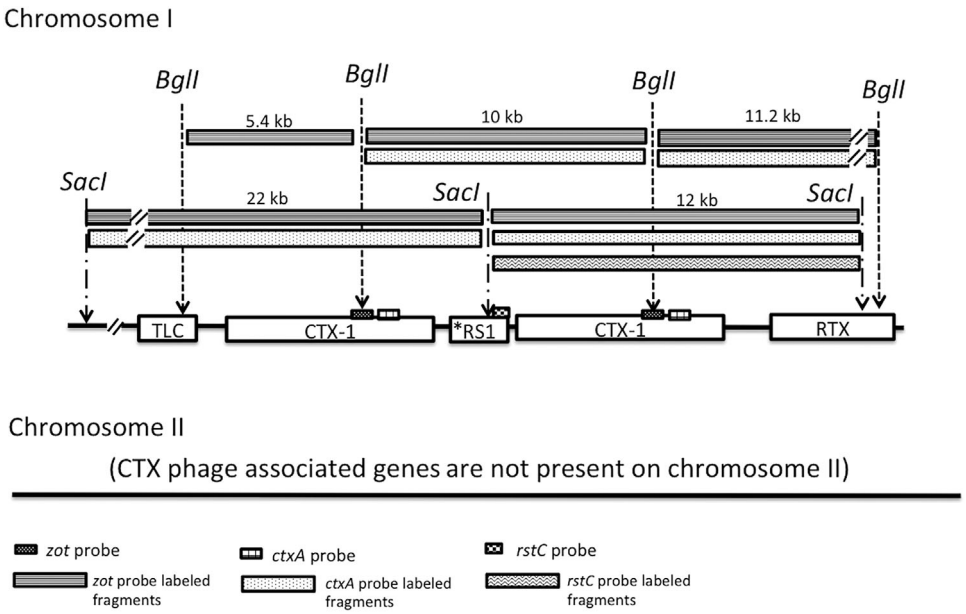


Fig. 2. Estimated CTX prophage region structure of *V. cholerae* strain J6, the representative of group ET-2. Two *V. cholerae* wave 1 strains revealed similar profiles of CTX prophage region-specific RFLP and PCR and were categorized into a group designated as ET-2. The strain J6 was chosen as a representative and sequenced. The best estimated model of CTX prophage region is "TLC–CTX-1–*RS1–CTX-1–RTX" for chromosome I; no CTX prophage-associated genes are present on chromosome II.

Determination of CTX prophage region of *V. cholerae* strain M25, a representative of group ET-8

The RFLP profile of strain M25 is unique so it was categorized as an independent group, ET-8 (Table 2, Fig. S8). The analysis predicted that there was at least one copy of CTX prophage and multiple copies of *rstC*-like ORFs in the genome (Table 2). Th ePCR profile indicated that RS1 is present close to TLC and *ctxA* and *ctxB* are present but apart from RTX. No CTX prophage-associated genes were amplified when a pair of

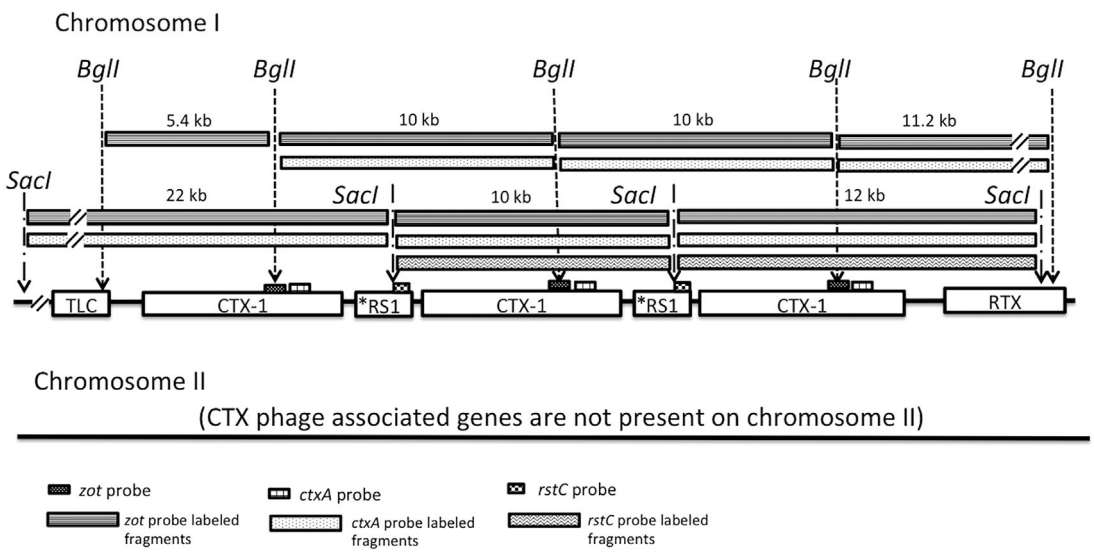


Fig. 3. Estimated CTX prophage region structure of *V. cholerae* strain C1, the representative of group ET-3. Profiles of CTX prophage region-specific RFLP and PCR of strain C1 are unique and it was categorized as an independent group, ET-3. The best estimated model for CTX prophage region is "TLC–CTX-1–*RS1–CTX-1–*RS1–CTX-1–RTX" for chromosome I; no CTX prophage-associated genes are present on chromosome II.

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

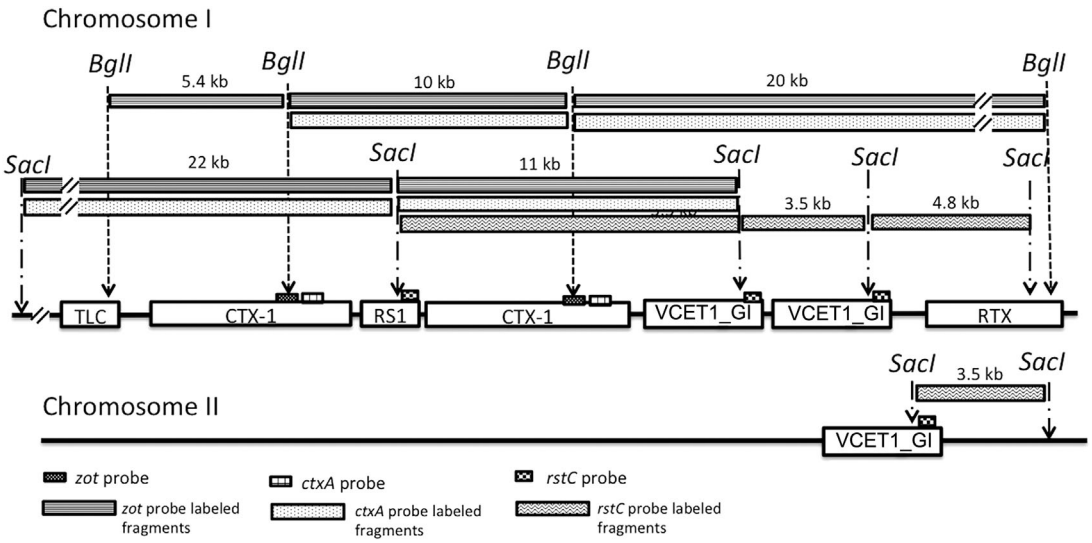


Fig. 4. Estimated CTX prophage region structure of *V. cholerae* strain C2, the representative of group ET-4. Profiles of CTX prophage region-specific RFLP and PCR of strain C2 are unique and it was categorized as an independent group, ET-4. The best estimated model for CTX prophage region of strain C2 is "TLC–CTX-1–RS1–CTX-1–VCET1_GI–VCET1_GI–RTX" on chromosome I and a single "VCET1_GI" on chromosome II.

CIIF and CIIR primers was used (Table 2). To determine the structure and sequence of the CTX prophage region of M25, primer pairs TLC3F and rstC2, rstC1 and ctxBR, and ctxBF and RTX5R were used on the genome to produce 2.8 kb, 7 kb and 3.5 kb fragments, respectively, and sequenced. It was found that strain M25 contains a single copy of CTX-1 and two copies of RS1, one flanking the upstream and the other the downstream of CTX-1 on chromosome I. HiSeq analysis and subsequent map read to N16961 predicted a single copy of CTX-1 and multiple copies of RS1 like regions are present in the genome. Taken together, the best estimated model for CTX prophage region of strain M25 is "TLC–RS1–CTX-1–RS1–RTX" (approximately 11.8 kb) on chromosome I and no CTX prophage-associated genes on chromosome II (Fig. 8).

Phylogenetic analysis of *V. cholerae* strains

A phylogenetic tree using 32 *V. cholerae* strains comprising 18 strains analyzed in the study and 14 reference strains is shown in Figure 9. After branching from M66-2 and MAK757 at [*], and C5 at [1], all 18 strains analyzed in the study and 10 strains representing waves 1, 2, and 3 of the seventh pandemic converged into relatively closer clusters under a node [1-2]. Strain C7 was solely placed under a node [1-2-1], whereas the remaining 17 strains were placed in clusters under a node [1-2-2] composed of a node [1-2-2-1] including strains J9, J6 and C1 and a node [1-2-2-2-1] including strains P18, P16, A1M, P7, T100, C2, 341, 193, T10, P3, P4, P31 and P2 (Fig. 9). Strain M25 was placed under a node [1-2-2-2-2] together with 10 reference waves 1–3 representative strains (Fig. 9).

V. cholerae strains were categorized into eight groups on the basis of the structure of the CTX prophage region (Table 2). Six strains belonging to the group ET-1 were segregated into a cluster at the node [1-2-2-2-1-2], consisting of a node [1-2-2-2-1-2-a] including strains A1M and P7, a node [1-2-2-2-1-2-b] including strains T-100 and 341 and a node [1-2-2-2-1-2-c] including strains 193 and T10. Strain C2, the sole strain belonging to the group ET-4 was placed at [1-2-2-2-1-2-b] together with two ET-1 strains. Strains J9 and J6, and strain C1, belonging to groups ET-2 and ET-3, respectively, were categorized into clusters at the node [1-2-2-1]. All four strains belonging to group ET-5 were segregated at a node [1-2-2-2-1-2-d]. Two strains (P18, P16) belonging to the group ET-7 were segregated in a cluster at the node [1-2-2-2-1-1]. Strain C7 and M25 sole strain belonging to ET-6 and ET-8, respectively, was placed into a cluster at node [1-2-1] and node [1-2-2-2-2], respectively.

DISCUSSION

Vibrio cholerae, an important gastrointestinal pathogen, causes cholera, which is characterized by severe diarrhea and subsequent dehydration. Cholera causes epidemics, principally in countries with poor sanitation and accounts for several million cases annually globally

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

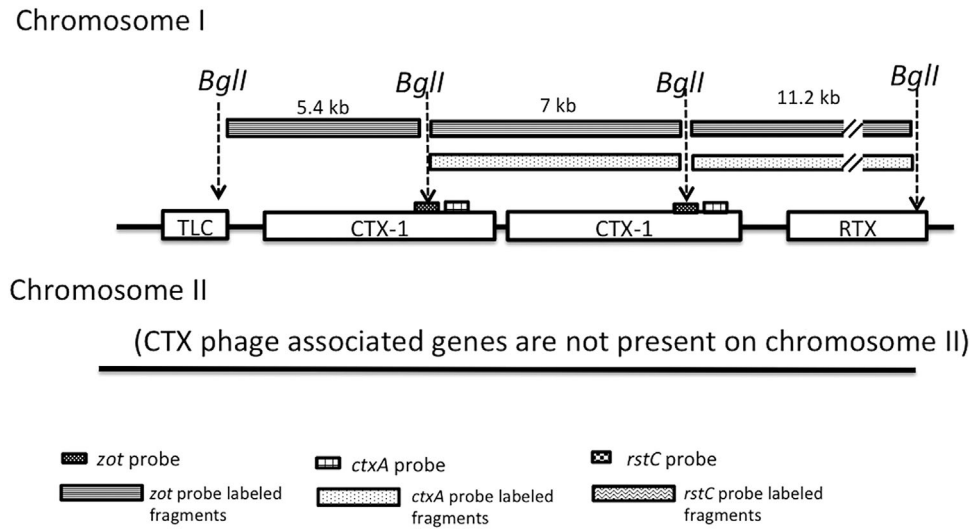


Fig. 5. Estimated CTX prophage region structure of *V. cholerae* strain P2, the representative of group ET-5. Four *V. cholerae* wave 1 strains revealed similar profiles of CTX prophage region-specific RFLP and PCR, and were categorized into a group designated as ET-5. The strain P2 was chosen as a representative and sequenced. The best estimated model of CTX prophage region is "TLC–CTX-1–CTX-1–RTX" for chromosome I; no CTX prophage-associated genes are present on chromosome II.

(30, 31). *V. cholerae* have undergone several dynamic population changes (7–9). Genome changes, acquisition of integrative mobile elements, and alterations in CTX phage are considered to have been the key mechanisms associated with the population changes (7, 8, 21). Mutreja *et al.* have reported identifying 50–250 SNPs related to the reference N16961 within the core genome of 123 *V. cholerae* El Tor strains isolated in the 1957–2010 seasons, making the rate of SNP accumulation 3.3 SNPs per year (7). In addition to CTX phage, El Tor strains have acquired at least three mobile elements; namely RS1, which facilitates CTX phage to reproduce, and VSP-I and VSP-II, whose functions in pathogenesis remain unknown. These changes eventually resulted in global dissemination of the strains and the seventh cholera pandemic since year 1961 (32, 33). Alterations in CTX prophage region have become evident in the evolution of *V. cholerae* (7, 34). Eleven distinct types of CTX prophage has been reported so far (21), categorized principally on the basis of genotypes of *rstR* and *ctxB* and

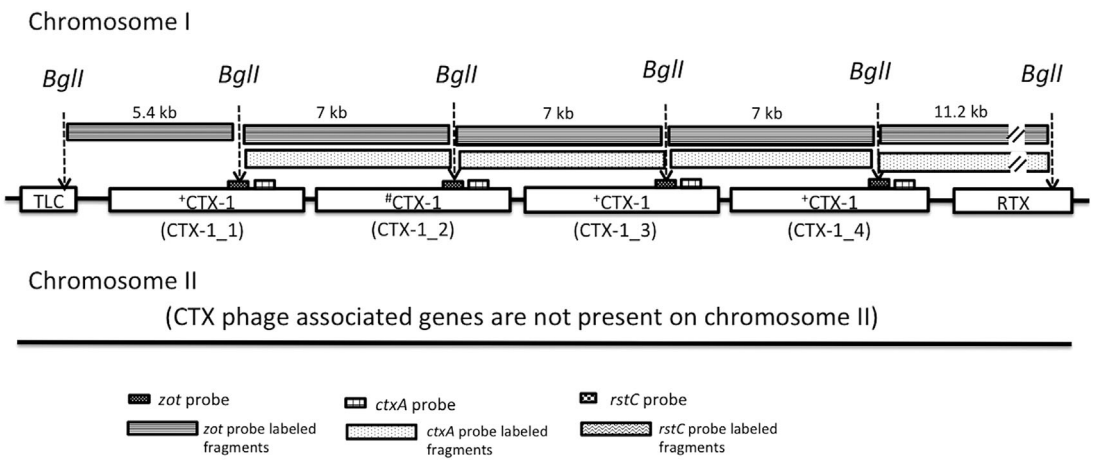


Fig. 6. Estimated CTX prophage region structure of *V. cholerae* strain C7, the representative of group ET-6. Profiles of CTX prophage region-specific RFLP and PCR of strain C7 are unique and it was categorized as an independent group, ET-6. The best estimated model for CTX prophage region of strain C7 is "TLC–$^{+}$CTX-1–$^{\#}$CTX-1–$^{+}$CTX-1–$^{+}$CTX-1–RTX" on chromosome I; no CTX prophage-associated genes are present on chromosome II. $^{+}$CTX-1, CTX-1 harboring SNPs in *rstA* (G301A), *rstB* (T84C), and in *ctxA* (G622A); $^{\#}$CTX-1, CTX-1 harboring SNPs in *rstB* (T84C) and in *ctxA* (G622A).

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

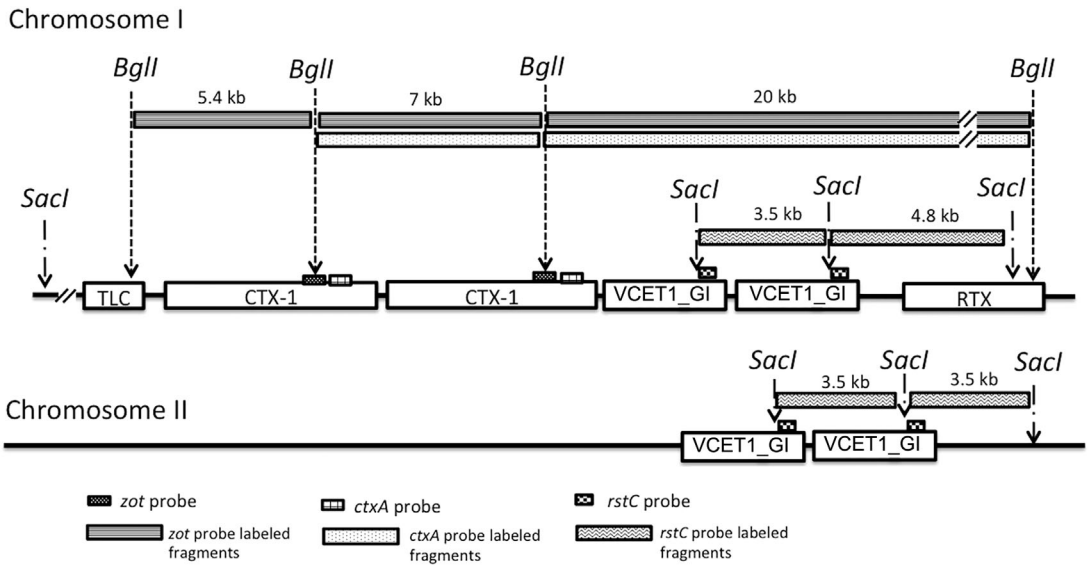


Fig. 7. Estimated CTX prophage region structure of *V. cholerae* strain P16, the representative of group ET-7. Two *V. cholerae* wave 1 strains revealed similar profiles of CTX prophage region-specific RFLP and PCR, and were categorized into a group designated as ET-7. The strain P16 was chosen as a representative and sequenced. The best estimated model for CTX prophage region of strain P16 is "TLC–CTX-1–CTX-1–VCET1_GI–VCET1_GI–RTX" on chromosome I and "VCET1_GI–VCET1_GI" on chromosome II.

on SNPs throughout the CTX prophage genome (22, 35). The structure of the CTX prophage region, which contains CTX prophage, RS1 and a possible mobile genomic island, is reportedly diverse among the seventh pandemic waves 2 and 3 strains of *V. cholerae;* however, little is known about the structure of the region in wave 1 strains (21). Hence, 18 *V. cholerae* strains isolated from patients in the 1956–1962 seasons, categorized as early wave 1 period, in South-East Asian countries and Japan, were analyzed and estimated structures of the region compared in thisstudy.

All 18 wave 1 *V. cholerae* strains were categorized into eight groups on the basis of their CTX prophage region specific RFLP with Southern blotting and PCR profiles (Table 2). One strain from each of the eight groups was chosen as a representative and sequenced to determine the structure of the CTX prophage region. All the representative strains were found to carry at least one copy of CTX-1, which harbors $rstR^{El\ Tor}$ and *ctxB3*, specific to El Tor strains (22), on chromosome I, but not on chromosome II. The eight representative strains were found to have eight distinct CTX prophage region structures with various combinations of CTX-1, RS1 and a novel genomic island (Figs. 1–8).

V. cholerae strains with ET-6 or ET-7 type CTX prophage regions have been isolated exclusively in the Philippines, whereas strains with ET-2 type have been isolated exclusively in Malaysia. In addition, four of six *V. cholerae* strains of ET-1 type were isolated in Taiwan in this study. Taken together, these findings indicate that *V. cholerae* strains with an area specific CTX prophage region were circulating in South-East Asian countries at an early stage of the seventh pandemic wave 1 period. During the same period, *V. cholerae* strains with three different types of CTX prophage region, namely ET-3, ET-4 and ET-5; and strains with ET-1, ET-6 and ET-7, were isolated in Indonesia and the Philippines, respectively, in 1961, indicating that *V. cholerae* strains with different types of CTX prophage regions were co-circulating in South-East Asian countries at an early stage of the wave 1 period (Figs. 1–8).

Phylogenetic tree analysis generated by SNP differences across 2483 core genes of the 32 *V. cholerae* strains revealed that strains categorized in the same CTX prophage-group were placed in relatively closer clusters (Fig. 9), indicating that the CTX prophage structure based-grouping is valid. Notably, strain C2, which belongs to group ET-4, was placed at node [1-2-2-2-1-2-b] together with two ET-1 strains and strain C1, which belongs to group ET-3, was placed at node [1-2-2-1] together with two ET-2 strains (Fig. 9). It is interesting to note that the structure of the CTX prophage regions of groups ET-1 and ET-4 have a similar combination of molecular elements (Figs., 4 1). For instance, ET-1 and ET-4 were each found to carry one copy of CTX-1 and tandem repeats of VCET1_GI on chromosome I and to carry sole VCET1_GI on chromosome II. Considering the evolutionary directions inferred from

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

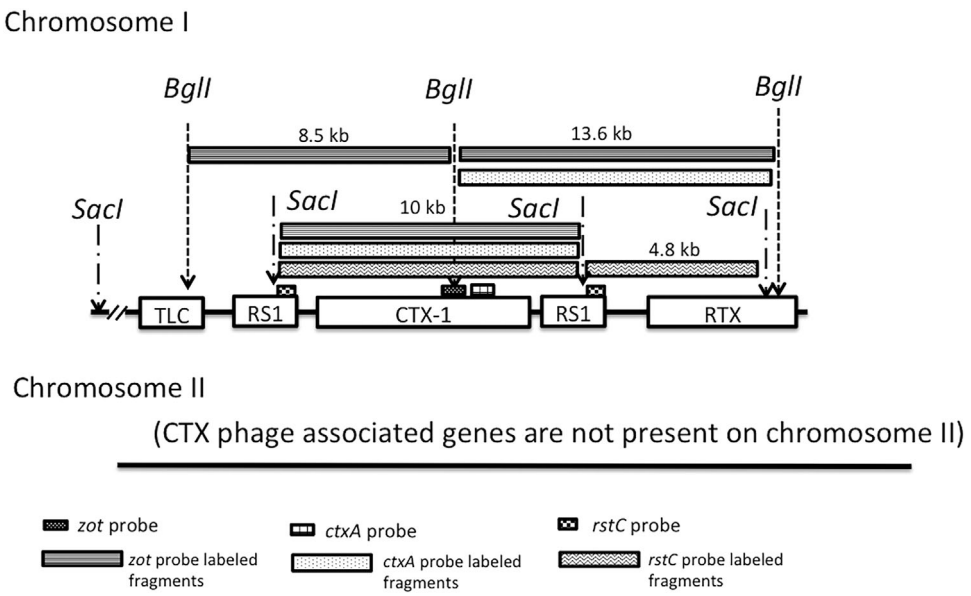


Fig. 8. Estimated CTX prophage region structure of *V. cholerae* strain M25, the representative of group ET-8. Profiles of CTX prophage region-specific RFLP and PCR of strain M25 are unique and it was categorized as an independent group, ET-8. The best estimated model for CTX prophage region of strain M25 is "TLC–RS1–CTX-1–RS1–RTX" on chromosome I and there are no CTX prophage-associated genes on chromosome II.

the phylogenetic tree, we propose that ET-4, represented by strain C2, was produced from ET-1, or vice versa, by the mechanisms of CTX prophage region-specific recombination events or multiple unusual replication events that occurred specifically in the CTX prophage region. A similar scenario could explain the placement of ET-2 and ET-3 strains, which are in the same lineage, at [1-2-2-1].

It is well recognized that recombination events have rarely occurred in the *V. cholerae* seventh pandemic strains since 1961, when they emerged (33, 36); however, recombination seems to have occurred frequently in pre-seventh pandemic strains (33). In this study, however, we found that a panel of CTX prophage regions with distinct structures is carried by a variety of pathogenic *V. cholerae* strains isolated from geographically diverse area, possibly indicating that CTX prophage region specific recombination events occurred at much higher frequencies than expected among pathogenic *V. cholerae* strains in the early wave 1 period. Levade *et al.* recently reported that *V. cholerae* have measurably evolved by accumulation of single nucleotide variants and a mechanism of gain and loss of genomic elements and as a consequence show variations in gene content (37). They concluded that these variations were induced within patients during infection and are unlikely to be culture or sequencing artifacts (37). Consistent with their observations, approximately 50 times continuous passages with LB media of several *V. cholerae* clinical isolates, including strain P16, did not induce any CTX prophage region changes according to specific RFLP and PCR findings (data not shown). We hypothesize that CTX prophage region specific recombination events may have occurred principally within patients during infection, driven probably by continuous pressure of specific immunity.

The structure of the new genomic island VCET1_GI is similar to that of RS1, possibly as a result of using a *dif* sequence exploiting Xer recombination mechanisms for integration, the evidence for this being that all VCET1_GI examined were found to carry attachment site (*attP*) (data not shown), which is responsible for integration into the genome, and are present close to CTX-1, which uses the same mechanism (17, 35, 38). VCET1_GI was detected in the CTX prophage region carried by *V. cholerae* strains isolated in geographically diverse countries (Table 1, Table 2, Fig. 1). Okada *et al.* have reported that a *V. cholerae* El Tor strain MS6 isolated from a patient in a Thailand–Myanmar border area in the 2008–2012 period carries a genomic island, named MS6CTXAGI prophage (28), which has a similar structure to VCET1_GI. Strain MS6 reportedly carries CTX prophage region with CTX-1 and the MS6CTXAGI island, which is similar to groups ET-1, ET-4 and ET-7 in structure. It is intriguing to postulate that a *V. cholerae* strain carrying the CTX prophage region with VCET1_GI like element was once prevalent in South-East Asian countries, specifically at an early stage of the seventh pandemic wave 1 period.

Typical El Tor wave 1 strains are believed to carry both CTX-1 and RS1 in CTX prophage region on

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

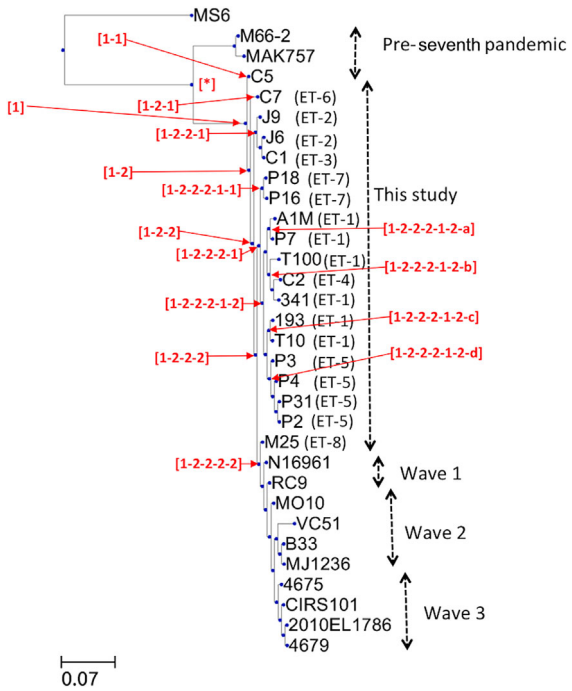


Fig. 9. Phylogenetic analysis of *Vibrio cholerae* strains. Maximum-likelihood phylogenetic tree of 32 *V. cholerae* strains including the 18 strains analyzed in the study, 14 referral strains composed of two strains of wave 1 (N16961, RC9), four strains of wave 2 (MO10, VC51, B33, MJ1236) and four strains of wave 3 (4675, CIRS101, 2010EL1786, 4679) of the seventh cholera pandemic, three strains of the pre-seventh pandemic (M66-2, MAK757 and C5), and MS6, a reported "US Gulf Coast" strain, was constructed based on SNP differences across 2483 core genes excluding recombination events. All 18 strains analyzed in the study and 10 strains representing waves 1, 2 and 3 of the seventh pandemic converged in relatively closer clusters under a node [1-2]. Note that *V. cholerae* strains categorized in the same CTX prophage-group were placed in relatively closer clusters.

chromosome I (39). However, *V. cholerae* strains carrying group ET-5, ET-6 or ET-7, which have been circulating in Indonesia and the Philippines, are devoid of RS1. One report found one *V. cholerae* El Tor wave 1 strain isolated in Indonesia in 1961 to lack RS1 (33), but this phenomenon is rare. Hassan *et al.* have reported that RS1 uses a *dif* sequence exploiting Xer recombination mechanism for integration or excision from genome (21, 40). It is hypothesized that the status of CTX prophage regions lacking RS1 is: (i) solely CTX-1 integrated next to TLC to form a CTX prophage region; or (ii) both CTX-1 and RS1 were involved in forming the CTX prophage region, after which RS1 was removed by homologous recombination leaving only CTX-1 (35, 41).

There are *V. cholerae* strains carrying CTX prophage region with multiple copy numbers of CTX-1. We found that strain C7, isolated in Indonesia in 1961, carries four tandem repeats of CTX-1 (Fig. 6); however, its ability to produce cholera toxin remains unknown. *ctxA* gene is reportedly stable and SNPs rarely occur in this area (22). Of note, we identified G622A mutations in *ctxA* in all four copies of CTX-1 carried by strain C7 (Table 3, Fig. 6), this mutation being associated with an Ala to Thr substitution at amino acid position 208. Effects on toxin productivity need to be clarified.

We identified five substitutions of T579C, T609C, C654T, C639T and T894C in *rstA* of satellite phage RS1 of strains of 341, J6 and C1, these strains representing groups ET-1, ET-2 and ET-3, respectively (Table 3). *V. cholerae* strains carrying RS1 with unique mutations are believed to have once been spread across South-East Asian countries.

To expand knowledge on the structure of the CTX prophage region of the seventh cholera pandemic wave 1 strains, 18*V. cholerae* strains isolated from patients in 1961–1962 in South-East Asian countries and Japan, were analyzed and compared. Similar to those of the wave 2 and wave 3 strains, CTX prophage region carried by wave 1 strains were diverse in structure. Some strains were estimated to carry four tandem repeats of CTX-1 and three copies of CTX-1 each separated by RS1 and to carry a novel mobile genomic island, VCET1_GI on chromosome I or II. A plausible hypothesis is that CTX prophage region-specific recombination events, or gain and loss of genomic elements within CTX prophage region, have occurred at much higher frequencies and consequently contributed to produce a panel of CTX prophage regions with distinct structures among *V. cholerae* pathogenic strains in lineages with close genetic backgrounds during the early wave 1 period of the seventh cholera pandemic.

ACKNOWLEDGMENTS

This research was supported by the Japan Agency for Medical Research and Development under Grant Number JP18fm0108001. The authors are grateful to the Program for Nurturing Global Leaders in Tropical and Emerging Communicable Diseases, Graduate School of Biomedical Sciences of Nagasaki University. We presented part of this study as an oral presentation at the 52nd Joint Panel Conference on Cholera and other Bacterial Enteric Infections (22 February 2018) and as a poster at the 91st Annual Meeting of the Japanese Society for Bacteriology (27–28 March 2018).

DISCLOSURES

The authors have no conflicts of interests to declare.

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

REFERENCES

1. Blake P.A., Allegra D.T., Snyder J.D., Barrett T.J., McFarland L., Caraway C.T., Feeley J.C., Craig J.P., Lee J.V., Puhr N.D., Feldman R.A. (1980) Cholera—a possible endemic focus in the United States. *N Engl J Med* **302**: 305–9.
2. Faruque S.M., Albert M.J., Mekalanos J.J. (1998) Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. *Microbiol Mol Biol Rev* **62**: 1301–14.
3. Kaper J.B., Morris J.G., Jr., Levine M.M. (1995) Cholera. *Clin Microbiol Rev* **8**: 48–86.
4. Olsvik O., Wahlberg J., Petterson B., Uhlen 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**: 22–5.
5. Keasler S.P., Hall R.H. (1993) Detecting and biotyping *Vibrio cholerae* O1 with multiplex polymerase chain reaction. *Lancet* **341**: 1661.
6. Kimsey H.H., Waldor M.K. (1998) CTXphi immunity: Application in the development of cholera vaccines. *Proc Natl Acad Sci USA* **95**: 7035–9.
7. Mutreja A., Kim D.W., Thomson N.R., Connor T.R., Lee J.H., Kariuki S., Croucher N.J., Choi S.Y., Harris S.R., Lebens M., Niyogi S.K., Kim E.J., Ramamurthy T., Chun J., Wood J.L., Clemens J.D., Czerkinsky C., Nair G.B., Holmgren J., Parkhill J., Dougan G. (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* **477**: 462–5.
8. Banerjee R., Das B., Balakrish Nair G., Basak S. (2014) Dynamics in genome evolution of Vibrio cholerae. *Infect Genet Evol* **23**: 32–41.
9. Naha A., Pazhani G.P., Ganguly M., Ghosh S., Ramamurthy T., Nandy R.K., Nair G.B., Takeda Y., Mukhopadhyay A.K. (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**: 1733–6.
10. Faruque S.M., Mekalanos J.J. (2012) Phage-bacterial interactions in the evolution of toxigenic *Vibrio cholerae*. *Virulence* **3**: 556–65.
11. Sack D.A., Sack R.B., Nair G.B., Siddique A.K. (2004) *Cholera*. *Lancet* **363**: 223–33.
12. Klose K.E. (2001) Regulation of virulence in *Vibrio cholerae*. *Int J Med Microbiol* **291**: 81–8.
13. Waldor M.K., Mekalanos J.J. (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**: 1910–4.
14. Davis B.M., Waldor M.K. (2003) Filamentous phages linked to virulence of *Vibrio cholerae*. *Curr Opin Microbiol* **6**: 35–42.
15. Mekalanos J.J. (2011) The Evolution of *Vibrio cholerae* as a Pathogen. In: Ramamurthy T., Bhattacharya S.K., eds. *Epidemiologidal and Molecular Aspects on Cholera*. New York: Springer, pp. 97–114.
16. Das B., Martinez E., Midonet C., Barre F.X. (2013) Integrative mobile elements exploiting Xer recombination. *Trends Microbiol* **21**: 23–30.
17. 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**: 559–66.
18. 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**: 4240–9.
19. Rubin E.J., Lin W., Mekalanos J.J., Waldor M.K. (1998) Replication and integration of a *Vibrio cholerae* cryptic plasmid linked to the CTX prophage. *Mol Microbiol* **28**: 1247–54.
20. Lin W., Fullner K.J., Clayton R., Sexton J.A., Rogers M.B., Calia K.E., Calderwood S.B., Fraser C., Mekalanos J.J. (1999) Identification of a *Vibrio cholerae* RTX toxin gene cluster that is tightly linked to the cholera toxin prophage. *Proc Natl Acad Sci USA* **96**: 1071–6.
21. Kim E.J., Lee C.H., Nair G.B., Kim D.W. (2015) Whole-genome sequence comparisons reveal the evolution of *Vibrio cholerae* O1. *Trends Microbiol* **23**: 479–89.
22. Kim E.J., Lee D., Moon S.H., Lee C.H., Kim D.W. (2014) CTX prophages in *Vibrio cholerae* O1 strains. *J Microbiol Biotechnol* **24**: 725–31.
23. Davis B.M., Kimsey H.H., Chang W., Waldor M.K. (1999) The *Vibrio cholerae* O139 Calcutta bacteriophage CTXphi is infectious and encodes a novel repressor. *J Bacteriol* **181**: 6779–87.
24. Iredell J.R., Manning P.A. (1994) Biotype-specific tcpA genes in *Vibrio cholerae*. *FEMS Microbiol Lett* **121**: 47–54.
25. Bakhshi B., Pourshafie M.R., Navabakbar F., Tavakoli A. (2008) Genomic organisation of the CTX element among toxigenic *Vibrio cholerae* isolates. *Clin Microbiol Infect* **14**: 562–8.
26. Basu A., Mukhopadhyay A.K., Garg P., Chakraborty S., Ramamurthy T., Yamasaki S., Takeda Y., Nair G.B. (2000) Diversity in the arrangement of the CTX prophages in classical strains of *Vibrio cholerae* O1. *FEMS Microbiol Lett* **182**: 35–40.
27. Mohapatra S.S., Mantri C.K., Turabe Fazil M.H., Singh D.V. (2011) *Vibrio cholerae* O1 biotype El Tor strains isolated in 1992 from Varanasi, India harboured El Tor CTXPhi and classical ctxB on the chromosome-I and classical CTXPhi and classical ctxB on the chromosome-II. *Environ Microbiol Rep* **3**: 783–90.
28. Okada K., Na-Ubol M., Natakuathung W., Roobthaisong A., Maruyama F., Nakagawa I., Chantaroj S., Hamada S. (2014) Comparative genomic characterization of a Thailand-Myanmar isolate, MS6, of *Vibrio cholerae* O1 El Tor, which is phylogenetically related to a "US Gulf Coast" clone. *PLoS ONE* **9**: e98120.
29. Wang H., Pang B., Xiong L., Wang D., Wang X., Zhang L., Kan B. (2015) The hybrid pre-CTXPhi-RS1 prophage genome and its regulatory function in environmental *Vibrio cholerae* O1 strains. *Appl Environ Microbiol* **81**: 7171–7.
30. Harris J.B., LaRocque R.C., Qadri F., Ryan E.T., Calderwood S.B. (2012) Cholera. *Lancet* **379**: 2466–76.
31. World Health Organization (2010) Cholera vaccines: WHO position paper. *Wkly Epidemiol Rec* **85**: 117–28.
32. Dziejman M., Balon E., Boyd D., Fraser C.M., Heidelberg J.F., Mekalanos J.J. (2002) Comparative genomic analysis of *Vibrio cholerae*: Genes that correlate with cholera endemic and pandemic disease. *Proc Natl Acad Sci USA* **99**: 1556–61.
33. Hu D., Liu B., Feng L., Ding P., Guo X., Wang M., Cao B., Reeves P.R., Wang L. (2016) Origins of the current seventh cholera pandemic. *Proc Natl Acad Sci USA* **113**: E7730–9.
34. Shah M.A., Mutreja A., Thomson N., Baker S., Parkhill J., Dougan G., Bokhari H., Wren B.W. (2014) Genomic epidemiology of *Vibrio cholerae* O1 associated with floods, Pakistan, 2010. *Emerg Infect Dis* **20**: 13–20.
35. Kim E.J., Lee D., Moon S.H., Lee C.H., Kim S.J., Lee J.H., Kim J.O., Song M., Das B., Clemens J.D., Pape J.W., Nair G.B., Kim D.W. (2014) Molecular insights into the evolutionary pathway of *Vibrio cholerae* O1 atypical El Tor variants. *PLoS Pathog* **10**: e1004384.
36. Mutreja A., Kim D.W., Thomson N.R., Connor T.R., Lee J.H., Kariuki S., Croucher N.J., Choi S.Y., Harris S.R., Lebens M., Niyogi S.K., Kim E.J., Ramamurthy T., Chun J., Wood J.L.N., Clemens

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd

J.D., Czerkinsky C., Nair G.B., Holmgren J., Parkhill J., Dougan G. (2011) Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* **477**: 462–5.
37. Levade I., Terrat Y., Leducq J.B., Weil A.A., Mayo-Smith L.M., Chowdhury F., Khan A.I., Boncy J., Buteau J., Ivers L.C., Ryan E.T., Charles R.C., Calderwood S.B., Qadri F., Harris J.B., LaRocque R.C., Shapiro B.J. (2017) *Vibrio cholerae* genomic diversity within and between patients. *Microb Genom* **3** doi: 10.1099/mgen.0.000142.
38. Huber K.E., Waldor M.K. (2002) Filamentous phage integration requires the host recombinases XerC and XerD. *Nature* **417**: 656–9.
39. 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., Umayam L., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P., McDonald L., Utterback T., Fleishmann R.D., Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C., Fraser C.M. (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**: 477–83.
40. Hassan F., Kamruzzaman M., Mekalanos J.J., Faruque S.M. (2010) Satellite phage TLCphi enables toxigenic conversion by CTX phage through dif site alteration. *Nature* **467**: 982–5.
41. Kamruzzaman M., Robins W.P., Bari S.M., Nahar S., Mekalanos J.J., Faruque S.M. (2014) RS1 satellite phage promotes diversity of toxigenic *Vibrio cholerae* by driving CTX prophage loss and elimination of lysogenic immunity. *Infect Immun* **82**: 3636–43.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Fig. S1. **Profile of RFLP with Southern blotting of genomic DNA prepared from *V. cholerae* 341 representing group ET-1.** (a) DNA digested with *BglI*. M, molecular marker; Lane 1, labeled with *zot* probe; Lane 2, labeled with *ctxA* probe. (b) DNA digested with *SacI*. Lane 3, labeled with zot probe; Llane 4, labeled with ctxA probe; Lane 5, labeled with *rstC* probe.

Fig. S2. **Profile of RFLP with Southern blotting of genomic DNA prepared from *V. cholerae* J6 representing group ET-2.** (a) DNA digested with *BglI*. M, molecular marker; Lane 1, labeled with *zot* probe; Lane 2, labeled with *ctxA* probe. (b) DNA digested with *SacI*. Lane 3, labeled with zot probe; Lane 4, labeled with ctxA probe; Lane 5, labeled with *rstC* probe.

Fig. S3. **Profile of RFLP with Southern blotting of genomic DNA prepared from *V. cholerae* C1 representing group ET-3.** (a) DNA digested with *BglI*. M, molecular marker; Lane 1, labeled with *zot* probe; Lane 2, labeled with *ctxA* probe. (b) DNA digested with *SacI*. Lane 3, labeled with *zot* probe; Lane 4, labeled with *ctxA* probe; Lane 5, labeled with *rstC* probe.

Fig. S4. **Profile of RFLP with Southern blotting of genomic DNA prepared from *V. cholerae* C2 representing group ET-4.** (a) DNA digested with *BglI*. M, molecular marker; Lane 1, labeled with *zot* probe; Lane 2, labeled with *ctxA* probe. (b) DNA digested with *SacI*. Lane 3, labeled with *zot* probe; Lane 4, labeled with *ctxA* probe; Lane 5, labeled with *rstC* probe.

Fig. S5. **Profile of RFLP with Southern blotting of genomic DNA prepared from *V. cholerae* C7 representing group ET-5.** (a) DNA digested with *BglI*. M, molecular marker; Lane 1, labeled with *zot* probe; Lane 2, labeled with *ctxA* probe. (b) DNA digested with *SacI*. Lane 3, labeled with *zot* probe; Lane 4, labeled with *ctxA* probe; Lane 5, labeled with *rstC* probe.

Fig. S6. **Profile of RFLP with Southern blotting of genomic DNA prepared from *V. cholerae*** P2 representing group ET-6. (a) DNA digested with *BglI*. M, molecular marker; Lane 1, labeled with *zot* probe; Lane 2, labeled with *ctxA* probe. (b) DNA digested with *SacI*. Lane 3, labeled with *zot* probe; Lane 4, labeled with *ctxA* probe; Lane 5, labeled with *rstC* probe.

Fig. S7. **Profile of RFLP with Southern blotting of genomic DNA prepared from *V. cholerae*** P16 representing group ET-7. (a) DNA digested with *BglI*. M, molecular marker; Lane 1, labeled with *zot* probe; Lane 2, labeled with *ctxA* probe. (b) DNA digested with *SacI*. Lane 3, labeled with *zot* probe; Lane 4, labeled with *ctxA* probe; Lane 5, labeled with *rstC* probe.

Fig. S8. **Profile of RFLP with Southern blotting of genomic DNA prepared from *V. cholerae*** M25 representing group ET-8. (a) DNA digested with *BglI*. M, molecular marker; Lane 1, labeled with *zot* probe; Lane 2, labeled with *ctxA* probe. (b) DNA digested with *SacI*. Lane 3, labeled with *zot* probe; Lane 4, labeled with *ctxA* probe; Lane 5, labeled with *rstC* probe.

Table S1. List of the primers and their sequences used in this study.

Table S2. List of the *Vibrio cholerae* O1 strains used for constructing phylogenetic tree.

© 2018 The Authors. *Microbiology and Immunology* published by The Societies and John Wiley & Sons Australia, Ltd