Chapter 3
The Genomics of Cholera

Bhabatosh Das and G. Balakrish Nair

Introduction

Vibrio cholerae, a bacterium autochthonous to the aquatic environment and introduced into the human intestine through contaminated water or food, is the etiological agent of the acute secretory diarrheal disease described as cholera. The pathogen has a free-living planktonic existence in aquatic bodies and has the ability to transmit into humans and cause disease. The process of completing an annual cycle in the environment and the transmission from contaminated water or food to humans is described as the ecology and epidemiology of the pathogen. This species contains a wide variety of both pathogenic and nonpathogenic strains. At the subspecies level, the organism is classified into more than 200 serogroups (Li et al. 2002). The differentiation of *V. cholerae* into serogroups is based on the differences in the sugar composition and therefore antigenicity of the heat-stable surface somatic "O" antigen. Only strains of serogroups O1 and O139 that produce cholera toxin defined as toxigenic strains have been recognized as agents of sporadic, endemic, epidemic, and pandemic cholera (Fig. 3.1) (Kaper et al. 1995; Sack et al. 2004). Most other serogroups of *V. cholerae* are not pathogenic or rarely cause local outbreaks, or mild gastroenteritis. *V. cholerae* strains belonging to serogroup O1 are further differentiated into two biotypes, classical and El Tor. The differentiation into biotypes is based on a combination of phenotypic, biochemical, and genetic traits, that include susceptibility to polymixin B, hemagglutination of chicken erythrocytes, hemolysis of sheep erythrocytes, the Voges–Proskauer test, susceptibility to phages, and nucleotide

B. Das
Centre National de la Recherche Scientifique, Centre de Genetique Moleculaire,
91198 Gif-sur-Yvette, France

G.B. Nair (✉)
National Institute of Cholera and Enteric Diseases, P-33, CIT Road Scheme XM,
Beliaghata, Kolkata 700 010, India
e-mail: nairgb@icmr.org.in

K.E. Nelson and B. Jones-Nelson (eds.), *Genomics Applications for the Developing World*, Advances in Microbial Ecology,
DOI 10.1007/978-1-4614-2182-5_3,

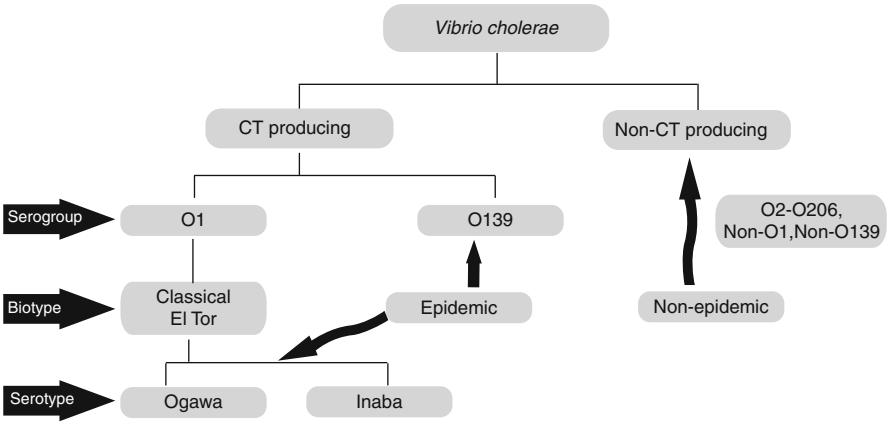

Fig. 3.1 Schematic representation of current scheme of classification of epidemic and nonepidemic *V. cholerae* isolates

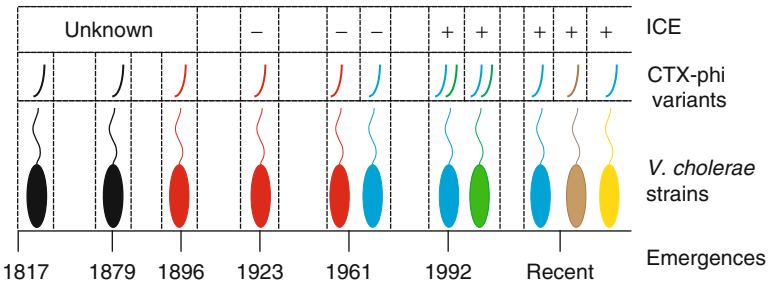

Fig. 3.2 Emergence of epidemic *V. cholerae* strains carrying variant of CTX prophages and ICEs recorded in the human history. Color code for bacteria and phages: *black*, unknown; *red*, classical; *blue*, El tor; *green*, O139; *gray*, altered El Tor; *yellow*, O1 hybrid

sequences of specific genes (Kaper et al. 1995). The other serogroups of *V. cholerae*, collectively called non-O1, non-O139, are not associated with epidemics and are ubiquitously distributed in the aquatic environment (Faruque et al. 1998).

Seven cholera pandemics have been recorded throughout human history. The sixth pandemic and presumably the earlier pandemics were caused by *V. cholerae* O1 of the classical biotype. The current seventh pandemic, which originated in Celebes islands in Indonesia in 1961, is caused by the *V. cholerae* O1 of the El Tor biotype (Kaper et al. 1995). In 1992, for the first time in the recorded history of cholera, a new serogroup, namely, O139 of *V. cholerae*, was detected as the causative agent of epidemic cholera in India and Bangladesh (Ramamurthy et al. 1993; group Cw 1993; Nair et al. 1994). Although *V. cholerae* O139 is currently restricted only to the Asian continent, several new epidemic lineages of *V. cholerae* O1 El Tor strain have been isolated from different parts of the world in the past 20 years (Fig. 3.2).

One of the original motivations to study bacterial genomics was to understand what constitutes a bacterial species, but now, bacterial genomics is used as an essential tool to understand adaptation and evolution in the prokaryotes. With more than 1,491 complete microbial genome sequences now available (NCBI genome database), it is clear that most bacterial chromosomes are mosaics, composed of DNA obtained via horizontal gene transfer as well as vertical inheritance. In this chapter, we discuss current knowledge of *V. cholerae* genomics and contribution of several integrative mobile genetic elements (IMGEs) to pathogenicity, antibiotic resistance, and bacterial evolution. We provide specific examples of how pathogenicity is derived and how different integrative mobile genetic elements have shaped the genomes and contributed to phenotypes for the survival and adaptation of vibrios in the natural environment as well as becoming an efficient pathogen, modulating its genetics to adapt to the hostility of host factors.

Genome Features

An important milestone in the field of cholera research was achieved when the whole genome sequence of a seventh pandemic cholera strain N16961 of O1 El Tor biotype was reported (Heidelberg et al. 2000). The strain N16961 was isolated in 1971 from a cholera patient in Bangladesh. Currently, more than 27 complete or draft genome sequences of pathogenic and environmental *V. cholerae* isolates (Table 3.1) are available in the public domain (www.ncbi.nlm.nih.gov/sites/genome). Sequencing was done either by traditional Sanger sequencing (Heidelberg et al. 2000; Chun et al. 2009) or by more advanced third-generation single-molecule real-time DNA sequencing methods (Chin et al. 2011). *V. cholerae* harbors two distinct, nonhomologous circular chromosomes (chromosomes I and II) (Heidelberg et al. 2000; Trucksis et al. 1998). The chromosomes I and II of N16961 consist of 2,961,146 and 1,072,315 base pairs (bp), with an average G+C content of 46.9% and 47.7%, respectively (Table 3.2). In silico analysis predicted a total of 4,009 open reading frames (ORFs) with average length of 1,000 bp (Table 3.2). There is pronounced asymmetry in the distribution of genes known to be essential for growth and virulence between the two chromosomes. Significantly, more genes encoding functions required for DNA replication and repair, transcription, translation, cell-wall biosynthesis, and a variety of central catabolic and biosynthetic pathways are located in chromosome I (Heidelberg et al. 2000). All 24 rRNA genes transcribed from eight different rRNA operons are also located in chromosome I. Many features of chromosome II, like its partitioning system and a large fractions of genes encoding unknown functions, are plasmid like. The preferential transcription of genes from the smaller chromosome, during colon colonization and different environmental stress conditions, suggests that this genomic organization is important for pathogenicity and environmental adaptation, although several genes important for normal cell function, e.g., *dsdA*, *thrS*, and the genes encoding ribosomal proteins L20 and L35, also are encoded on chromosome II (Heidelberg et al. 2000; Xu et al. 2003).

Table 3.1 Complete or draft genome of *V. cholerae* strains isolated over the past 100 years from different parts of the world

Strains	Serogroups/biotypes	Year of isolation	Geographical origin	Genome size (nt)	Source	Reference
NCTC8457	O1 El Tor	1910	Saudi Arabia	4,063,388	AAWD01000000	(Chun et al. 2009)
MAK757	O1 El Tor	1937	Celebes Islands	3,919,418	AAUS00000000	(Chun et al. 2009)
M66-2	O1 El Tor	1937	Indonesia	3,938,905	NC012578/NC012580	(Feng et al. 2008)
TM11079-80	O1 El Tor	1980	Brazil	4,055,140	ACHW00000000	(Chun et al. 2009)
2740-80	O1 El Tor	1980	US Gulf Coast	3,945,478	AAUT01000000	(Chun et al. 2009)
RC9	O1 El Tor	1985	Kenya	4,211,011	ACHX00000000	(Chun et al. 2009)
12,129(1)	O1 El Tor	1985	Australia	3,969,506	ACFQ00000000	(Chun et al. 2009)
BX330286	O1 El Tor	1986	Australia	4,000,672	ACIA00000000	(Chun et al. 2009)
RC27	O1 classical	1991	Indonesia	4,011,779	NZADAI00000000	LANL, NMPDR
INDRE91/1	O1 El Tor	1991	Mexico	3,947,707	NZADAK00000000	LANL, NMPDR
MJ-1236	O1 hybrid	1994	Bangladesh	4,236,368	CP001485/CP001486	(Chun et al. 2009)
CIRS 101	O1 El Tor altered	2002	Bangladesh	4,059,686	ACVW00000000	(Chun et al. 2009)
B33	O1 El Tor altered	2004	Mozambique	4,154,698	ACHZ00000000	(Chun et al. 2009)
V52	O37	1968	Sudan	3,974,495	AAKJ02000000	(Chun et al. 2009)
TMA21	non-O1/O139	1982	Brazil	4,023,772	ACHY00000000	(Chun et al. 2009)
V51	O141	1987	USA	3,782,275	AAKI02000000	(Chun et al. 2009)
1,587	O12	1994	Peru	4,137,501	AAUR01000000	(Chun et al. 2009)
RC385	O135	1998	Chesapeake Bay	4,120,133	AAKH02000000	(Chun et al. 2009)
MZO-2	O14	2001	Bangladesh	3,862,985	AAWF01000000	(Chun et al. 2009)
MZO-3	O37	2001	Bangladesh	4,146,039	AAUU01000000	(Chun et al. 2009)
AM-19226	O39	2001	Bangladesh	4,056,157	AATY01000000	(Chun et al. 2009)
623-39	non-O1/O139	2002	Bangladesh	3,975,259	AAWG00000000	(Chun et al. 2009)
VL426	Albensis	Unknown	United Kingdom	3,987,383	ACHV00000000	(Chun et al. 2009)
CT 5369-93	non-O1/O139	1993	Brazil	3,557,309	NZADAL00000000	LANL, NMPDR

LANL, Los Alamos National Laboratory; *NMPDR*, National Microbial Pathogen Data Resource

Table 3.2 Comparative genomics of representative *V. cholerae* pathogenic isolates belonging to O1 classical, O1 El tor, and O139 serogroups

Feature	O395 (O1 classical)	N16961 (O1 El Tor)	MO10 (O139)
Genome size (bp)	4,132,319	4,033,464	4,034,412
Number of chromosomes	2	2	2
G+C content (%)	47.53	47.48	47
Total genes	3,998	4,009	3,942
Protein coding genes	3,878	3,887	3,870
tRNA genes	96	98	72
rRNA genes	24	24	24
No of complete CTX prophage	2 (1in Ch1 and 1 in Ch2)	1 (Ch1)	Multiple (Ch1)
No of SXT element	0	0	1

Several intermediaries of metabolic pathways are encoded only on chromosome II. Both chromosomes contain distinct replication origins and dimmer resolution systems (Val et al. 2008).

Comparative Genomics: Genetic Keys to Unravel Bacterial Evolution to Species Differentiation

With the availability of highly efficient, fast, and cost-effective third-generation DNA sequencing technology, comparative genomics is now emerging as feasible genetic tool to unravel what constitutes a bacterial species and to identify genetic factors driving adaptation and evolution in the prokaryotes. Several studies have looked at pathogenic species and their nonpathogenic relatives in an effort to discover which genes might be unique to the pathogen and therefore drive pathogenesis (Casjens 1998). Similarly, genomic comparisons among multiple strains of the same species have revealed that the total gene content of any two strains is not identical, that is, the genomic information of a bacterial species is not represented by the genome sequence of a single strain. Recently, Colwell and coworkers (Chun et al. 2009) reported complete genome sequences of 23 *V. cholerae* strains of clinical and environmental origins isolated from a variety of sources over the past 100 years (Table 3.1). They showed that the number of unique genes in the 23 sequenced *V. cholerae* stains far exceeded the number of genes found in any one strain. Analysis of the whole genome information of 23 *V. cholerae* strains revealed 12 distinct *V. cholerae* lineages. All known pandemic strains of both the sixth and seventh pandemics were assigned to one lineage, called the *V. cholerae* phylocore genome (PG) clade. However, sequence analysis of core genome revealed *V. cholerae* O1 classical biotype strains substantially differ from O1 El Tor strains, while seventh pandemic strains, including O139 and new variant O1 El Tor strains, showed highly related genome sequences. From the published sequenced genomes, the main genomic features of three representative strains belonging to O1 El tor (N16961), O1 classical (O395), and O139 (MO10) serogroups are summarized in the Table 3.2.

Table 3.3 Recently described gene cassettes in *V. cholerae* conferring resistance to most commonly used antimicrobials

Gene	Resistant to	Carrier	References
sul2	Sulfonamide	ICE	(Waldor et al. 1996)
dfrA18	Trimethoprim	ICE	(Waldor et al. 1996)
floR	Florfenicol	ICE	(Waldor et al. 1996)
tetG	Tetracycline	Plasmid	(Ceccarelli et al. 2006)
aphA1	Kanamycin	Plasmid	(Ceccarelli et al. 2006)
aadA1	Spectinomycin	Transposon	(Goldstein et al. 1986)
arr2	Rifampicin	Integron	(Fluit and Schmitz 2004)
blaP1	Ampicillin	Plasmid	(Ceccarelli et al. 2006)
cat1	Chloramphenicol	Plasmid	(Ceccarelli et al. 2006)
qnrVC3	Ciprofloxacin	Integron	(Thungapathra et al. 2002)
aac-Ib	Amikacin	Integron	(Thungapathra et al. 2002)
strAB	Streptomycin	ICE	(Hochhut et al. 2001)
ereA2	Erythromycin	Integron	(Thungapathra et al. 2002)

Comparative Genomics: Genetic Keys to Unravel Emergence of Multidrug-Resistant Bacterial Strains

Acquisition and dissemination of antibiotic resistance traits among pathogenic bacterial isolates is the most striking example of evolution that has been observed in bacteria over the past six decades. Six years after the launch of several commercial antimicrobials like streptomycin (Sm), spectinomycin (Sp), tetracycline (Tc), and chloramphenicol (Cm), isolates of *Shigella dysenteriae* that were resistant to each of these antibiotics were identified (Mitsuhashi et al. 1961). However, *V. cholerae* continued to remain sensitive to all commonly used antibiotics till 1970s. Ampicillin (Ap), kanamycin (Kn), Sm, Sp, and Tc resistance in *V. cholerae* pathogenic isolates were reported during 1977 and 1980 in Tanzania and Bangladesh, respectively (Threlfall et al. 1980). The resistance traits were attributed to either chromosomal mutation or the acquisition of resistance genes from mobile genetic elements (Threlfall et al. 1980). *V. cholerae* strains isolated during the past three decades carry resistance traits to almost all commonly used antibiotics (Table 3.3) (Mukhopadhyay et al. 1998; Ghosh and Ramamurthy 2011). Now, more than 27 complete or draft sequenced genomes of antibiotic-sensitive and resistant variants available in the public domain provide an excellent opportunity to explain emergence and reemergence of multidrug-resistant pathogenic variants.

Genome comparison reveals that *V. cholerae* strains isolated before 1977 have native antibiotic target and no resistance gene cassettes. Strains isolated during or after 1977 acquired antibiotic resistance either by modifying the antibiotic's target due to spontaneous mutation or by acquiring resistance gene cassettes from different mobile genetic elements like transposons, integrating conjugative elements (ICEs), and mobile integrons (Table 3.3) (Goldstein et al. 1986; Waldor et al. 1996). IMGE encoding resistance to sulfamethoxazole (Su), and trimethoprim (Tm) was reported for the first time in *V. cholerae* O139 strain MO10 isolated in the Madras in South

India in 1992 and described as SXT (Waldor et al. 1996). SXT is a 100-kb mobile genetic element that integrates site specifically into the 5-end of a gene (*prfC*) which encodes peptide chain release factor 3 (RF3) found on the ChI of *V. cholerae*. SXT or similar type of ICEs conferring resistance to several useful antibiotics are not restricted only in the O139 isolates; instead, they are prevalent in most epidemic *V. cholerae* strains including serogroup O1 isolated from different parts of the world (Wozniak et al. 2009). Acquisition and rapid dissemination of SXT and related ICEs in the both clinical and environmental strains is one of the key drivers for emergence of multidrug resistance among bacterial pathogens.

Genesis of the Hybrid *V. cholerae* Strains

Over the last two decades, several new epidemic lineages of *V. cholerae* O1 El Tor strain have been recorded (Fig. 3.2). In 1992, for the first time in the history of cholera, a new serogroup, namely, O139 of *V. cholerae*, was detected as the causative agent of epidemic cholera in the Indian subcontinents (Ramamurthy et al. 1993; group Cw 1993). Although the initial concern was that a new pandemic was beginning, the geographic range of *V. cholerae* O139 is currently restricted to only the Asian continent. Additionally, *V. cholerae* O1 hybrids and altered El Tor variants carrying CTXΦ in either or both chromosomes (Das et al. 2010) have been isolated repeatedly in Bangladesh and Mozambique (Nair et al. 2002; Faruque et al. 2007; Das et al. 2007). Altered *V. cholerae* O1 El Tor isolates produce cholera toxin of the classical biotype but can be biotyped as El Tor by conventional phenotypic assays, whereas *V. cholerae* O1 hybrid variants cannot be biotyped using standard phenotypic tests and can produce cholera toxin of either biotype. These new variants have subsequently replaced the prototype seventh pandemic *V. cholerae* O1 El Tor strains in Asia and Africa, with respect to frequency of isolation from clinical cases of cholera (Safa et al. 2010). Sequence analyses of the recent clinical isolates predict presumably these new O1 variants originated from the common O1 El Tor ancestor. The unique combination of genomic characteristics achieved by the acquisition of mobile genetic elements make current O1 clinical isolates superior compared to prototype O1 El Tor clones. In this context, Siddique et al. (Siddique et al. 2010) have shown that cholera caused by the hybrid variants is more severe than that caused by prototype El Tor cholera. Further, it has also been noticed that recent cholera epidemics of cholera caused by the hybrids tend to be more protracted over time (recent examples are the cholera epidemics in Zimbabwe and Haiti) as compared to earlier outbreaks of cholera caused by prototype El Tor biotype strains. Studies have also shown that the hybrid variants of the El Tor biotype produce larger amounts of cholera toxin almost equivalent to that of the classical biotype and much more than the prototype El Tor biotype (Ghosh-Banerjee et al. 2010). The improved environmental fitness and enhanced virulence in terms of higher production of cholera toxin of the hybrid variants presumably was responsible for the replacement of the prototype O1 El Tor clones in the cholera endemic zones.

Plasticity of the *V. cholerae* Genome

Microbial genomes are highly mosaic and dynamic (Casjens 1998). The composite structure of microbial genomes is developed by the acquisition of IMGEs either from closely or distantly related organisms by conjugation, transformation, or transduction. The fluidity of genome components and their relative arrangements allows bacteria to respond rapidly to changing environmental conditions and to new environmental challenges. The dynamic nature of microbial genomes can be attributed to the ability of bacteria to share genetic information by horizontal gene transfer systems and to the variety of homologous and nonhomologous recombinatorial processes that integrate, excise, and translocate genes into their genome. Many IMGEs directly contribute to the evolutionary fitness of their host because they harbor antibiotic resistance, xenobiotic degradation, and/or virulence factors (Frost et al. 2005; Waldor and Mekalanos 1996). In addition, some genetic elements, those that can be regarded as purely selfish, may not carry any genes of specific function but can indirectly participate in the adaptation of their host to new environmental conditions since they contribute to the plasticity of the host genome. Analysis of more than 27 complete or draft genomes of different pathogenic and environmental *V. cholerae* isolates revealed several IMGEs carrying pathogenic determinants or antimicrobial resistance traits or genes presumably associated to the host fitness exist in their chromosomes and contribute to the genome diversity and host evolution (Fig. 3.3). Here, we present some well-known IMGEs crucial for the pathogenicity, antimicrobial resistance, and environmental fitness of *V. cholerae*.

CTXΦ

The most important virulence factor of *V. cholerae*, cholera toxin (CT), is encoded in the genome of an integrated prophage CTXΦ. Although DNA sequences of CTXΦ isolated from different toxigenic *V. cholerae* strains are quite different, their genome is identical at the functional level. CTXΦ has a ~7-kb ss(+) DNA genome arranged in two modular structures, the "RS2" and "core" (Waldor et al. 1997). The functions needed for the phage morphogenesis are present in the core region and are encoded by the five genes (e.g., *psh*, *cep*, *gIIICTX*, *ace*, *zot*). The core region also harbors the *ctxAB* genes not required for the phage morphogenesis but responsible for the severe diarrhea associated with cholera. Three proteins, designated as RstR, RstA, and RstB, are encoded in RS2. RstA, initiator of rolling circle replication, is essential for phage replication (Waldor et al. 1997). RstB, a single-stranded DNA binding protein, plays a crucial role in phage integration (Falero et al. 2009). RstR acts as a transcriptional repressor by inhibiting the activity of P_{rstA}, the only phage promoter required for CTXΦ replication and integration (Kimsey and Waldor 2004).

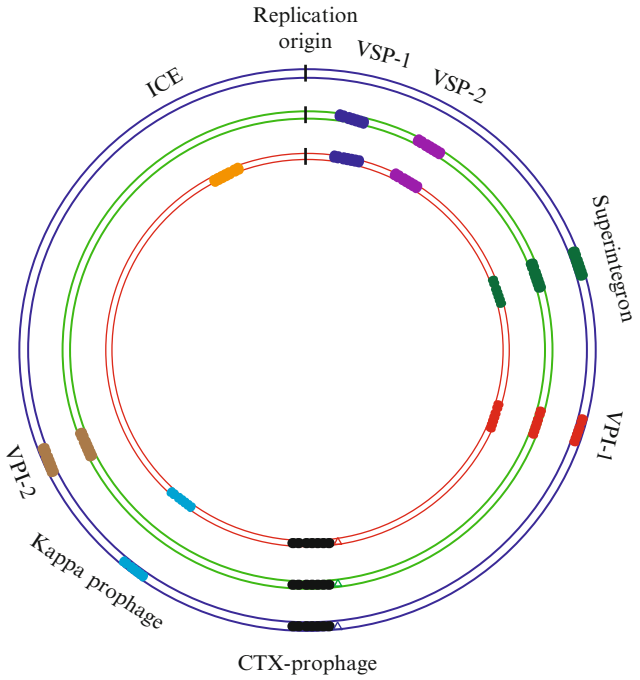

Fig. 3.3 Schematic representation of the genetic elements that influence the plasticity of the *V. cholerae* genome. From the outermost ring: genome of O1 classical O395, genome of O1 El Tor N16961, and genome of O139 MO10 (Data obtained from (Grim et al. 2010))

Conventionally, CTXΦs can be classified into four families based on the sequence of their *rstR* gene (Davis et al. 1999). These categories are designated as $CTX\Phi^{ET}$, $CTX\Phi^{Cl}$, $CTX\Phi^{Cla}$, and $CTX\Phi^{Env}$ according to the host cells in which they were most frequently isolated. It is surprising to find that several *V. cholerae* strains isolated in the current pandemic from different geographical origins harbor CTX prophages with mosaic genome compare to prototype $CTX\Phi^{ET}$ and $CTX\Phi^{Cl}$ phages (Fig. 3.4). For example, CTX prophages present in the genome of the hybrid variants MJ-1236, B33, and CIRS101 strains (Table 3.1) carry *ctxB* allele of $CTX\Phi^{Cl}$ origin where as all other genes in the core region are identical to those of $CTX\Phi^{ET}$ phage (Chun et al. 2009; Halder et al. 2010). It seems that the CTX core region of current epidemic strains is hybrid in nature carrying the "pre-CTX" core region of the El Tor type, with *ctxB* of the classical type. However, recent Haitian *V. cholerae* isolates carry the *ctxB* allele with three nonsynonymous substitutions compared to canonical $CTX\Phi^{ET}$ encoded CtxB (Chin et al. 2011). Two of these changes (Ile68Thr and Tyr39His) are characteristic of CtxB of prototypical $CTX\Phi^{Cl}$ phage, and the third substitution mutation (His20Asn) has previously been observed only in El Tor variant strains from South Asia (Goel et al. 2008) and in very recent isolates from West Africa (Quilici et al. 2010).

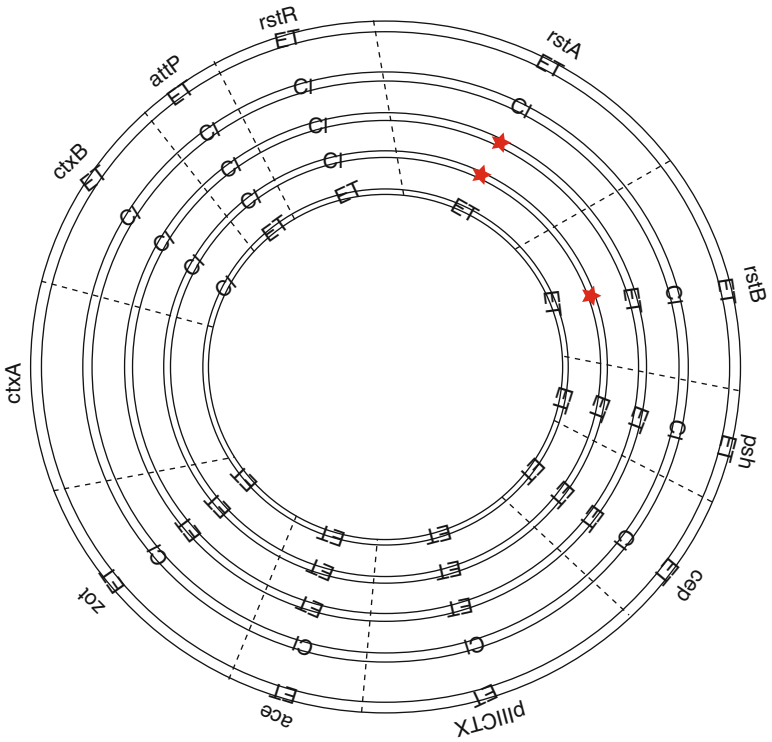

Fig. 3.4 Schematic presentation of genome of CTXΦ integrated in the chromosomes of different epidemic *V. cholerae* isolates (from the outermost ring): CTX prophage of N16961, CTX prophage of O395, CTX prophage of Mj-1236, CTX prophage of B33, and CTX prophage of CRIS101. ET and Cl indicate alleles are identical to prototypical CTXΦ^{ET} or CTXΦ^{Cl} phages, respectively. An *asterisk* represents allele, which is not identical to that of prototypical CTXΦ^{ET} or CTXΦ^{Cl} phages. *ctxA* is identical in both CTXΦ^{ET} and CTXΦ^{Cl} phages

Although RS2 carries only three genes and is smaller than the core region, allelic variations in recent isolates are more prominent (Fig. 3.4). Some of lineages carry identical genes like prototypical CTXΦ^{ET} (e.g., CIRS101) or classical type *rstR* but El Tor type *rstB* (MJ-1236) or *rstAB* alleles different from either of prototypical CTXΦ^{ET} or CTXΦ^{Cl} phages (e.g., B33). Current classification of CTXΦ is now not very informative. With the identification of an increasing number of phage variants, which harbor allele of several origins, this suggests that a new CTXΦ classification is required. We believe the element essential for phage integration would be simplest and most informative candidate for new CTXΦ classification. As CTXΦ does not encode any integrase, its attachment sites might be appropriate for its designation. To date, three categories of CTXΦ attachment sites were identified which target three different *dif* sites identified in the *V. cholerae* strains sequenced so far (*dif1, dif2, difG*) (Das et al. 2010). Based on the integration specificity, CTXΦs can be categoried as CTXΦ^{1} (specific for *dif1*) or CTXΦ^{1,2} (can integrates in both *dif1* and *dif2*) or CTXΦ^{G} (specific for *difG*) (Das et al. 2010; Das et al. 2011).

ICE

Integrating and conjugative elements (ICEs) are a class of self-transmissible mobile genetic elements that are key mediators of horizontal gene transfer among prokaryotes (Burrus et al. 2002). These mobile elements share features common with other well-known mobile genetic elements like plasmids and phages. Like conjugative plasmids, ICEs mediate transfer via conjugation. In contrast to plasmids, they do not replicate autonomously. To ensure vertical transmission, ICEs integrate into the host chromosomes, a feature similar to many temperate bacteriophages. Burrus et al. (2002) first introduced the term ICE meant to include conjugative transposons. The most well-characterized ICE in *V. cholerae* is SXT^{MO10} first reported in 1992 in *V. cholerae* O139 clinical isolates from India (Waldor et al. 1996). This is an ~100-kb ICE that encodes resistance to four antibiotics: Su, Tm, Cm, and Sm. SXT ^{MO10} and also other characterized *V. cholerae* ICEs integrate site specifically into the 5'-end of a gene (*prfC*) that encodes peptide chain release factor 3 (RF3) found on the ChI of *V. cholerae* (Fig. 3.4) (Hochhut and Waldor 1999). ICE integration is mediated by its own tyrosine recombinase (Int). Recombination occurs between 17 bp, nearly identical DNA sequences of ICE (*attP*) and *V. cholerae* chromosome (*attB*) (Hochhut and Waldor 1999).

Another well-characterized ICE, R319 (~89 kb), which mediate Kn and mercury (Hg) resistance, was derived from African *Providencia rettgeri* (Coetzee et al. 1972). All the SXT/R319 ICEs characterized share a common structure and have sizes ranging from 79 to 108 kb (Burrus et al. 2006b). They contain sets of 52 conserved genes of around 47-kb size and encode functions essential for their integration, excision, conjugative transfer, and regulation (Burrus et al. 2006a), as well as many genes of unknown functions. In addition to sharing 52 core genes, ICEs also carry diverse sequences that confer the elements specific function, like antibiotics resistance. ICEs are one of the major keys that influence *V. cholerae* genome plasticity and are extremely important for the pathogen to combat antimicrobial treatment.

Integron

Integrons are natural genetic elements able to incorporate and disseminate exogenous ORFs by site-specific recombination. Integrons have three basic components necessary for the capture, expression, and dissemination of circular exogenous gene cassettes: (Li et al. 2002) *intI*, encoding the tyrosine recombinase essential for integration and excision of gene cassettes; (1) *attI*, integration site of integrons; and (2) Pc, an outward-orientated promoter that ensure transcription of the captured promoter-less genes (Cambray et al. 2010). Depending on their location, all integrons can be divided into two distinct subsets: the mobile integrons (MIs), which are linked to mobile DNA elements, and the chromosomal superintegrons (CIs)

(Mazel 2006). MIs are always physically linked to mobile genetic elements, like insertion sequences (ISs), transposons, or conjugative plasmids, which can mediate their transmission between bacterial species. A pool of more than 130 cassettes encoding resistance against different antibiotics has been identified in the MIs (Cambray et al. 2010). In the late 1990s, CI was first identified in *V. cholerae* (Mazel et al. 1998). The CI of *V. cholerae* serotype O1 biotype El Tor strain N16961 located in its small chromosome contains at least 215 ORFs and occupies ~3% of the genome (Heidelberg et al. 2000). Most of the superintegron gene cassettes examined so far seem to be unique to the host species and have no known functions. However, some of them have significant sequence homology with known antibiotic resistance gene cassettes indicating that they might confer the resistance potential if exposed to proper substrates. In *V. cholerae*, several genes encoding virulence factors are also located in the superintegron (Ogawa and Takeda 1993; Barker and Manning 1997).

Genomic Islands

Genomic islands (GIs) are discrete segments of DNA detected by comparative genome sequencing among closely related bacterial species, which carry set of genes having important functions for bacterial pathogenesis and/or host fitness. The mechanism of acquisition and dissemination of GI is not well established. Toxigenic *V. cholerae* O1 and O139 strains harbor four well-characterized GIs, Vibrio pathogenicity island-1 (VPI-1), Vibrio pathogenicity island-2 (VPI-2), Vibrio seventh pandemic island-1 (VSP-1), and Vibrio seventh pandemic island-2 (VSP-2), associated with the pathogenesis and fitness of *V. cholerae*. These genomic islands are significantly different among the sequenced *V. cholerae* genomes and thus participate in host genome diversity and evolution.

VPI-1

VPI-1 is a 41-kb DNA molecule that integrates at the tmRNA (*ssrA*) locus on the large chromosome of most O1 and O139 toxigenic *V. cholerae* isolates. VPI-1 encodes several important proteins necessary for the *V. cholerae* pathogenicity. This includes *ctxAB* transcription regulators ToxT and TcpPH, the accessory colonization factor (ACF), and most importantly the CTXΦ receptor toxin-coregulated pilus (TCP) (Karaolis et al. 1998). Excision of VPI-1 from its chromosomal integration site has been noticed, and a circular intermediate was detected by conventional genetic approaches like PCR assay. Excision of VPI-I does not essentially depend on its cognate tyrosine recombinase. VPI-1 can be mobilized between *V. cholerae* serogroup O1 strains, although the mechanism is still not clear (Murphy and Boyd 2008).

VPI-2

VPI-2 is a 57-kb DNA molecule consisting of 52 ORFs and integrates at the tRNA-serine locus in the large chromosome of *V. cholerae* toxigenic isolates. VPI-2 is absent in the genome of several *V. cholerae* epidemic isolates, like MO10, CRIS101, and MAK757 (Chun et al. 2009). VPI-2 encodes a type 1 restriction modification system (Jermyn and Boyd 2002), sialic acid transport and catabolism, and neuraminidase (Hacker and Kaper 2000; Stewart-Tull et al. 1986). VPI-2 of pathogenic non-O1/O139 *V. cholerae* isolates also encodes type three secretion systems (Chen et al. 2007). In contrast to VPI-1, VPI-2 encode a tyrosine recombinase (Int_{VPI-2}) that can excise the GI from its chromosomal location to form nonreplicative circular intermediates. VPI-1 and VPI-2 do not share any genes.

VSP-I

VSP-I, a 16-kb DNA segment, harbors 13 ORFs found in all sequenced seventh pandemic *V. cholerae* isolates. Transcriptome analysis of several O1 and O139 *V. cholerae* strains revealed the presence of VSP-I in the seventh pandemic isolates (Dziejman et al. 2002). Recently, Grim et al. (2010) analyzed genomes of more than 300 *Vibrio* strains and reported that some non-O1/non-O139 *V. cholerae* and *V. mimicus* also possess VSP-1-like elements, indicating the prevalence of this GI across both species, namely, *V. cholerae* and *V. mimicus*. Most of the ORFs harbored within VSP-1 encode hypothetical or conserved hypothetical proteins whose function remains unknown. A XerCD-like putative tyrosine recombinase (Int_{VSP-I}) might participate in the integration and dissemination of VSP-1 and was detected in this GI. Compared to other GI, VSP-1 is highly conserved among O1 El Tor isolates.

VSP-II

VSP-II is a 27-kb region integrated at a tRNA-methionine locus and encodes a putative tyrosine recombinase (Int_{VSP-II}), presumably essential for its integration and dissemination (Murphy and Boyd 2008). Initially, comparative genomics by microarray technology revealed VSP-II as a 7.5-kb GI (Dziejman et al. 2002). VSP-II is not ubiquitous among seventh pandemic isolates, and a significant deletion of the island has been reported in the current epidemic strains CIRS101 and MO10 (Chun et al. 2009). Additionally, the Latin American strains of *V. cholerae* carry a specific deletion in the VSP-II region, which has been used as a molecular signature (Nusrin et al. 2009). VSP-II-encoded proteins have several important cellular functions like DNA repair, transcriptional regulation, cell attachment, chemotaxis, and site-specific recombination (O''Shea et al. 2004). The VSP-II appears to be restricted to the seventh pandemic strains of *V. cholerae* and is believed to have caused the initiation of the seventh pandemic of cholera in 1961.

Application of the Knowledge of the Genome of *Vibrio cholerae*

In conclusion, the exponential knowledge on the genome of this pathogen has been very useful to select experimental approaches to further knowledge to combat the ancient scourge in multiple ways:

1. Sequence data is useful to identify highly conserved metabolically essential pathogenic proteins as candidates to find new antimicrobials against the multi-drug-resistant strains of the pathogen. In addition, it will also be helpful to avoid spreading of antibiotic resistance within the pathogens by preventing use of capricious drugs, which stimulate bacterial SOS response and stimulating dissemination of antibiotic resistance genes harbored by the pathogens.
2. Comparative genomics can serve to identify the most reliable pathogenic isolates for selecting best candidates for effective vaccine development and help to prevent new epidemics in the cholera endemic zone of developing countries.
3. In addition, sequence data could be useful for the construction of modern genetic tools, like microarrays chip, to study the virulence of *V. cholerae*.
4. However, in reality, the extensive information on the genome of *V. cholerae* has not played a direct role in public health and in mitigating the burden of the disease cholera. In fact, the global burden of cholera has risen in the past 5 years according to recent World Health Organization estimates. A challenge for the future will be how to convert this trillion bytes of information on the *V. cholerae* genome into public health tools that will enable us to combat the disease in the front line and also prevent the explosive outbreaks of cholera that we witness today.

Concluding Remarks

V. cholerae is a highly pathogenic and infectious bacterial species. Several representative genome sequences are now available in the public domain, which provide a unique opportunity to elucidate fundamental questions linked to the pathogenicity and evolution of pathogens. The pathogen is continuously evolving to adapt to the changing environment. Comparative genome sequencing clearly indicate that IMGEs are the key player for their successful evolution. Several different basic questions regarding the evolution of *V. cholerae* still need to be addressed. These include:

1. Emergence of new CTXΦ variants: Current *V. cholerae* isolates harbor novel CTXΦ carrying genetic material of different origins. Mechanism of shuffling of genomic components between phage variants is a long-standing question of whether it is generated by homologous recombination in the host cells carrying different phage variants or some unknown site-specific recombination systems exist in the host that help the acquisition of *ctxAB* allele in the pre-CTX element. More interestingly, why is CTXΦ constantly evolving in the environment?

Additionally, what is the function of the *ctxAB* genes to the phage genome and is it in the phage genome to be ferried about from strain to strain?

2. The genomes of recent epidemic isolates are quite heterogeneous. Several GIs present in one variant are deleted in other isolates. Questions arise about their importance in host fitness. Secondly, GIs are specifically integrated at the tRNA loci. It is important to know why they specifically select such locations. Is it because such targets are numerous in the bacterial genomes and located in the highly transcribed operons, thus ensuing their optimal cellular level, or whether the special tRNA transcription machineries interact with integrase that make bias for the target selection? Some GIs have orientation biasness, further raising the question of their mechanism of integration.
3. ICEs are present in all current epidemic isolates. All ICEs do not harbor antibiotic resistance traits but are still maintained in the genomes. What is the driving force ensuring their stable inheritance?

Acknowledgments The authors acknowledge all members of the F.X. Barre lab for helpful suggestions during the preparation of this chapter. BD is supported by the CNRS postdoctoral research fellowship, Government of France. GBN acknowledges the support of the Indian Council of Medical Research, New Delhi, India.

References

Barker A, Manning PA (1997) VlpA of *Vibrio cholerae* O1: the first bacterial member of the alpha 2-microglobulin lipocalin superfamily. Microbiology 143(Pt 6):1805–1813

Burrus V, Marrero J, Waldor MK (2006a) The current ICE age: biology and evolution of SXT-related integrating conjugative elements. Plasmid 55(3):173–183

Burrus V, Pavlovic G, Decaris B, Guedon G (2002) Conjugative transposons: the tip of the iceberg. Mol Microbiol 46(3):601–610

Burrus V, Quezada-Calvillo R, Marrero J, Waldor MK (2006b) SXT-related integrating conjugative element in New World *Vibrio cholerae*. Appl Environ Microbiol 72(4):3054–3057

Cambray G, Guerout AM, Mazel D (2010) Integrons. Annu Rev Genet 44:141–166

Casjens S (1998) The diverse and dynamic structure of bacterial genomes. Annu Rev Genet 32:339–377

Ceccarelli D, Salvia AM, Sami J, Cappuccinelli P, Colombo MM (2006) New cluster of plasmid-located class 1 integrons in *Vibrio cholerae* O1 and a dfrA15 cassette-containing integron in Vibrio parahaemolyticus isolated in Angola. Antimicrob Agents Chemother 50(7):2493–2499

Chen Y, Johnson JA, Pusch GD, Morris JG Jr, Stine OC (2007) The genome of non-O1 Vibrio cholerae NRT36S demonstrates the presence of pathogenic mechanisms that are distinct from those of O1 *Vibrio cholerae*. Infect Immun 75(5):2645–2647

Chin CS, Sorenson J, Harris JB, Robins WP, Charles RC, Jean-Charles RR et al (2011) The origin of the Haitian cholera outbreak strain. N Engl J Med 364(1):33–42

Chun J, Grim CJ, Hasan NA, Lee JH, Choi SY, Haley BJ et al (2009) Comparative genomics reveals mechanism for short-term and long-term clonal transitions in pandemic *Vibrio cholerae*. Proc Natl Acad Sci USA 106(36):15442–15447

Coetzee JN, Datta N, Hedges RW (1972) R factors from Proteus rettgeri. J Gen Microbiol 72(3):543–552

Das B, Bischerour J, Barre FX (2011) VGJphi integration and excision mechanisms contribute to the genetic diversity of *Vibrio cholerae* epidemic strains. Proc Natl Acad Sci USA 108(6): 2516–2521

Das B, Bischerour J, Val ME, Barre FX (2010) Molecular keys of the tropism of integration of the cholera toxin phage. Proc Natl Acad Sci USA 107(9):4377–4382

Das B, Halder K, Pal P, Bhadra RK (2007) Small chromosomal integration site of classical CTX prophage in Mozambique *Vibrio cholerae* O1 biotype El Tor strain. Arch Microbiol 188(6): 677–683

Davis BM, Kimsey HH, Chang W, Waldor MK (1999) The *Vibrio cholerae* O139 Calcutta bacteriophage CTXphi is infectious and encodes a novel repressor. J Bacteriol 181(21):6779–6787

Dziejman M, Balon E, Boyd D, Fraser CM, Heidelberg JF, Mekalanos JJ (2002) Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc Natl Acad Sci USA 99(3):1556–1561

Falero A, Caballero A, Ferran B, Izquierdo Y, Fando R, Campos J (2009) DNA binding proteins of the filamentous phages CTXphi and VGJphi of *Vibrio cholerae*. J Bacteriol 191(18):5873–5876

Faruque SM, Albert MJ, Mekalanos JJ (1998) Epidemiology, genetics, and ecology of toxigenic *Vibrio cholerae*. Microbiol Mol Biol Rev 62(4):1301–1314

Faruque SM, Tam VC, Chowdhury N, Diraphat P, Dziejman M, Heidelberg JF et al (2007) Genomic analysis of the Mozambique strain of *Vibrio cholerae* O1 reveals the origin of El Tor strains carrying classical CTX prophage. Proc Natl Acad Sci USA 104(12):5151–5156

Feng L, Reeves PR, Lan R, Ren Y, Gao C, Zhou Z et al (2008) A recalibrated molecular clock and independent origins for the cholera pandemic clones. PLoS One 3(12):e4053

Fluit AC, Schmitz FJ (2004) Resistance integrons and super-integrons. Clin Microbiol Infect 10(4):272–288

Frost LS, Leplae R, Summers AO, Toussaint A (2005) Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol 3(9):722–732

Ghosh A, Ramamurthy T (2011) Antimicrobials & cholera: are we stranded? Indian J Med Res 133(2):225–231

Ghosh-Banerjee J, Senoh M, Takahashi T, Hamabata T, Barman S, Koley H et al (2010) Cholera toxin production by the El Tor variant of *Vibrio cholerae* O1 compared to prototype El Tor and classical biotypes. J Clin Microbiol 48(11):4283–4286

Goel AK, Jain M, Kumar P, Bhadauria S, Kmboj DV, Singh L (2008) A new variant of *Vibrio cholerae* O1 El Tor causing cholera in India. J Infect 57(3):280–281

Goldstein F, Gerbaud G, Courvalin P (1986) Transposable resistance to trimethoprim and 0/129 in *Vibrio cholerae*. J Antimicrob Chemother 17(5):559–569

Grim CJ, Choi J, Chun J, Jeon YS, Taviani E, Hasan NA et al (2010) Occurrence of the *Vibrio cholerae* seventh pandemic VSP-I island and a new variant. OMICS 14(1):1–7

group Cw (1993) Large epidemic of cholera-like disease in Bangladesh caused by *Vibrio cholerae* O139 synonym Bengal. Cholera Working Group, International Centre for Diarrhoeal Diseases Research, Bangladesh. Lancet 342(8868):387–390

Hacker J, Kaper JB (2000) Pathogenicity islands and the evolution of microbes. Annu Rev Microbiol 54:641–679

Halder K, Das B, Nair GB, Bhadra RK (2010) Molecular evidence favouring step-wise evolution of Mozambique *Vibrio cholerae* O1 El Tor hybrid strain. Microbiology 156(Pt 1):99–107

Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ et al (2000) DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature 406(6795): 477–483

Hochhut B, Lotfi Y, Mazel D, Faruque SM, Woodgate R, Waldor MK (2001) Molecular analysis of antibiotic resistance gene clusters in *vibrio cholerae* O139 and O1 SXT constins. Antimicrob Agents Chemother 45(11):2991–3000

Hochhut B, Waldor MK (1999) Site-specific integration of the conjugal *Vibrio cholerae* SXT element into prfC. Mol Microbiol 32(1):99–110

Jermyn WS, Boyd EF (2002) Characterization of a novel Vibrio pathogenicity island (VPI-2) encoding neuraminidase (nanH) among toxigenic *Vibrio cholerae* isolates. Microbiology 148(Pt 11): 3681–3693

Kaper JB, Morris JG Jr, Levine MM (1995) Cholera. Clin Microbiol Rev 8(1):48–86

Karaolis DK, Johnson JA, Bailey CC, Boedeker EC, Kaper JB, Reeves PR (1998) A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. Proc Natl Acad Sci USA 95(6):3134–3139

Kimsey HH, Waldor MK (2004) The CTXphi repressor RstR binds DNA cooperatively to form tetrameric repressor-operator complexes. J Biol Chem 279(4):2640–2647

Li M, Shimada T, Morris JG Jr, Sulakvelidze A, Sozhamannan S (2002) Evidence for the emergence of non-O1 and non-O139 *Vibrio cholerae* strains with pathogenic potential by exchange of O-antigen biosynthesis regions. Infect Immun 70(5):2441–2453

Mazel D (2006) Integrons: agents of bacterial evolution. Nat Rev Microbiol 4(8):608–620

Mazel D, Dychinco B, Webb VA, Davies J (1998) A distinctive class of integron in the *Vibrio cholerae* genome. Science 280(5363):605–608

Mitsuhashi S, Harada K, Hashimoto H, Egawa R (1961) On the drug-resistance of enteric bacteria 4 Drug-resistance of *Shigella* prevalent in Japan. Jpn J Exp Med 31:47–52

Mukhopadhyay AK, Basu I, Bhattacharya SK, Bhattacharya MK, Nair GB (1998) Emergence of fluoroquinolone resistance in strains of *Vibrio cholerae* isolated from hospitalized patients with acute diarrhea in Calcutta, India. Antimicrob Agents Chemother 42(1):206–207

Murphy RA, Boyd EF (2008) Three pathogenicity islands of *Vibrio cholerae* can excise from the chromosome and form circular intermediates. J Bacteriol 190(2):636–647

Nair GB, Faruque SM, Bhuiyan NA, Kamruzzaman M, Siddique AK, Sack DA (2002) New variants of *Vibrio cholerae* O1 biotype El Tor with attributes of the classical biotype from hospitalized patients with acute diarrhea in Bangladesh. J Clin Microbiol 40(9):3296–3299

Nair GB, Ramamurthy T, Bhattacharya SK, Mukhopadhyay AK, Garg S, Bhattacharya MK et al (1994) Spread of *Vibrio cholerae* O139 Bengal in India. J Infect Dis 169(5):1029–1034

Nusrin S, Gil AI, Bhuiyan NA, Safa A, Asakura M, Lanata CF et al (2009) Peruvian *Vibrio cholerae* O1 El Tor strains possess a distinct region in the Vibrio seventh pandemic island-II that differentiates them from the prototype seventh pandemic El Tor strains. J Med Microbiol 58(Pt 3): 342–354

Ogawa A, Takeda T (1993) The gene encoding the heat-stable enterotoxin of *Vibrio cholerae* is flanked by 123-base pair direct repeats. Microbiol Immunol 37(8):607–616

O'Shea YA, Finnan S, Reen FJ, Morrissey JP, O'Gara F, Boyd EF (2004) The Vibrio seventh pandemic island-II is a 26.9 kb genomic island present in *Vibrio cholerae* El Tor and O139 serogroup isolates that shows homology to a 43.4 kb genomic island in V. vulnificus. Microbiology 150(Pt 12):4053–4063

Quilici ML, Massenet D, Gake B, Bwalki B, Olson DM (2010) *Vibrio cholerae* O1 variant with reduced susceptibility to ciprofloxacin, Western Africa. Emerg Infect Dis 16(11):1804–1805

Ramamurthy T, Garg S, Sharma R, Bhattacharya SK, Nair GB, Shimada T et al (1993) Emergence of novel strain of *Vibrio cholerae* with epidemic potential in southern and eastern India. Lancet 341(8846):703–704

Safa A, Nair GB, Kong RY (2010) Evolution of new variants of *Vibrio cholerae* O1. Trends Microbiol 18(1):46–54

Sack DA, Sack RB, Nair GB, Siddique AK (2004) Cholera. Lancet 363(9404):223–233

Siddique AK, Nair GB, Alam M, Sack DA, Huq A, Nizam A et al (2010) El Tor cholera with severe disease: a new threat to Asia and beyond. Epidemiol Infect 138(3):347–352

Stewart-Tull DE, Ollar RA, Scobie TS (1986) Studies on the *Vibrio cholerae* mucinase complex. I. Enzymic activities associated with the complex. J Med Microbiol 22(4):325–333

Threlfall EJ, Rowe B, Huq I (1980) Plasmid-encoded multiple antibiotic resistance in *Vibrio cholerae* El Tor from Bangladesh. Lancet 1(8180):1247–1248

Thungapathra M, Amita SKK, Chaudhuri SR, Garg P, Ramamurthy T et al (2002) Occurrence of antibiotic resistance gene cassettes aac(6')-Ib, dfrA5, dfrA12, and ereA2 in class I integrons in non-O1, non-O139 *Vibrio cholerae* strains in India. Antimicrob Agents Chemother 46(9): 2948–2955

Trucksis M, Michalski J, Deng YK, Kaper JB (1998) The *Vibrio cholerae* genome contains two unique circular chromosomes. Proc Natl Acad Sci USA 95(24):14464–14469

Val ME, Kennedy SP, El Karoui M, Bonne L, Chevalier F, Barre FX (2008) FtsK-dependent dimer resolution on multiple chromosomes in the pathogen *Vibrio cholerae*. PLoS Genet 4(9):e1000201

Waldor MK, Mekalanos JJ (1996) Lysogenic conversion by a filamentous phage encoding cholera toxin. Science 272(5270):1910–1914

Waldor MK, Rubin EJ, Pearson GD, Kimsey H, Mekalanos JJ (1997) Regulation, replication, and integration functions of the *Vibrio cholerae* CTXphi are encoded by region RS2. Mol Microbiol 24(5):917–926

Waldor MK, Tschape H, Mekalanos JJ (1996) A new type of conjugative transposon encodes resistance to sulfamethoxazole, trimethoprim, and streptomycin in *Vibrio cholerae* O139. J Bacteriol 178(14):4157–4165

Wozniak RA, Fouts DE, Spagnoletti M, Colombo MM, Ceccarelli D, Garriss G et al (2009) Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. PLoS Genet 5(12):e1000786

Xu Q, Dziejman M, Mekalanos JJ (2003) Determination of the transcriptome of *Vibrio cholerae* during intraintestinal growth and midexponential phase in vitro. Proc Natl Acad Sci USA 100(3):1286–1291