Recent advances in cholera genetics

G. Uma[†], M. Chandrasekaran[†], Yoshifumi Takeda[#] and G. Balakrish Nair[‡,*]

[†]Department of Biotechnology, Cochin University of Science and Technology, Cochin, India
[#]Faculty of Human Life Sciences, Jissen Women's University, Tokyo 191-8510, Japan
[‡]International Centre for Diarrhoeal Disease Research, Dhaka 1212, Bangladesh

***Vibrio cholerae* is a comprehensively studied human pathogen. Genetic analyses conducted in the recent past show that the pathogenic forms of the microbe share certain virulence traits that distinguish epidemic strains from non-epidemic forms. The microbe has a dual lifestyle, primarily residing in the aquatic environs in association with zooplankton and other flora, but is also capable of transiently residing in the human intestine with the toxigenic forms causing cholera. The presence of virulence-associated genes in environmental strains provides interesting avenues to understand the pathogenicity of the disease; this is especially so because rampant genetic exchange has been demonstrated as the primary cause for the origination of new epidemic forms in many instances. Cholera research took a new turn with the completion of the sequencing of the whole genome of an El Tor biotype strain, opening new vistas to probe into the deepest pathogenic secrets of the organism. This review attempts to bring together some of the most important researches in recent times that have gone into understanding the genetics of the cholera germ.**

CHOLERA is a severely dehydrating diarrhoeal condition and is one disease in modern times that is epidemic, endemic and pandemic in nature. The epidemic forms of the disease are caused by the O1 and O139 serogroups of *Vibrio cholerae*, a bacterium that has interested microbiologists for long and continues to throw up surprises, despite the fact that much is known about its biology. The 2001 annual estimates from WHO indicate that *V. cholerae* was responsible for 1,84,000 cases of cholera around the world[1]; but this is only the tip of the iceberg, since several countries do not report or under-report cholera.

The recent advances in cholera genetics have revealed the extent of genetic flux in the bacterial genome, and large-scale sequencing techniques have been put to use in decoding the alphabets of the cholera genome. These have brought to fore the importance of gene transfer in the emergence of epidemic and pandemic strains of cholera. Many genes associated with virulence in cholera are now known to have been acquired through gene-transfer events that have mediated the transformation of avirulent strains to those capable of causing outbreaks of diarrhoea.

*For correspondence. (e-mail: gbnair@icddrb.org)

Classification

The recent issue of the *Bergey's Manual of Systematic Bacteriology* classifies the family Vibrionaceae into six genera, namely *Vibrio, Allomonas, Enhydrobacter, Listonella, Photobacterium* and *Salinivibrio*[2]. The vibrios are a group of Gram-negative, curved or straight motile rods that normally inhabit the aquatic environments. Currently, the genus consists of 51 species, of which at least 12 are known to be associated with human disease. *V. cholerae* is currently classified into more than 200 somatic antigen (O) serogroups[3]. Until 1992, the only serogroup known to cause epidemic cholera was O1. Strains belonging to the O1 serogroup are further classified into two biotypes, namely the classical and El Tor, and these can be differentiated by different phenotypic traits and more recently, by precise genetic markers[4]. Seven recorded pandemics of cholera have occurred globally and there is firm evidence that at least the fifth and sixth were caused by the classical biotype O1 strains[5]. The El Tor biotype is responsible for the ongoing seventh pandemic[5]. In 1992, yet another serogroup, namely O139 started causing outbreaks of cholera in India and Bangladesh[6,7]. Currently, these two serogroups are associated with endemic and epidemic cholera, while the other *V. cholerae* serogroups not associated with epidemics or pandemics are collectively referred to as non-O1, non-O139 *V. cholerae* or also as non-epidemic serogroups (Figure 1). Serogrouping is carried out using specific absorbed antisera or monoclonal antibodies against the 'O' antigen component of the bacterial lipopolysaccharide[3]. In addition to this, *V. cholerae* O1 is classified into three serotypes, namely Ogawa, Inaba, and Hikojima, the last of which is a rare and inadequately described serotype[3].

Shuffling the 'O's: Genetic exchanges in the 'O' antigen synthesis region that led to emergence of O139

The role of lateral gene transfer in the generation of new pathogenic strains is best elucidated by the O139 cholera outbreak. In 1992, an epidemic of cholera broke out in Chennai; this was unexpectedly caused by a non-O1 strain[6]. Outbreaks of cholera that subsequently broke out in Kolkata and Bangladesh were caused by the same strain that did not agglutinate with any of the existing

antisera to 138 serogroups of *V. cholerae* known at that time[6,7]. Consequently, the strain was designated as belonging to a new serogroup O139, and the synonym Bengal was attached to this new serogroup to indicate its first isolation from coastal cities off the Bay of Bengal[8]. After the initial outbreaks, the strain spread to other regions and imported cases were reported from other continents[9]. In India, O139 showed surges of preponderance over the existing O1 El Tor strains as the dominant cause of cholera[10].

The O139 Bengal strain and the O1 serogroup outbreak strains of both the classical and El Tor biotypes show many similarities, but their major differences are significant. The O139 strain is capsulated unlike the O1 strains and has significant dissimilarities in the 'O' antigen component of the bacterial lipopolysaccharide[11]. Genetic studies point to the possibility that a seventh pandemic El Tor O1 strain could have given rise to the O139 strain by genetic shuffling. In fact, analysis of the gene cluster encoding 'O' antigen biosynthesis in the O1 and O139 strains, namely the *wbf* gene cluster, showed that there was a replacement of a 22 kbp region coding for O1-antigen specificity by a larger genetic element of 35 kbp that encoded O139 specificity[12]. The entire O-antigen biosynthesis region of an O139 strain was cloned and used to probe reference strains of 192 other *V. cholerae* serogroups[13]. One cluster of genes did not react with any of the existing serogroups, suggesting that this fragment may have originated from another bacterium. The lower GC content of these genes in comparison to the overall *V. cholerae* GC content also supported this hypothesis. The

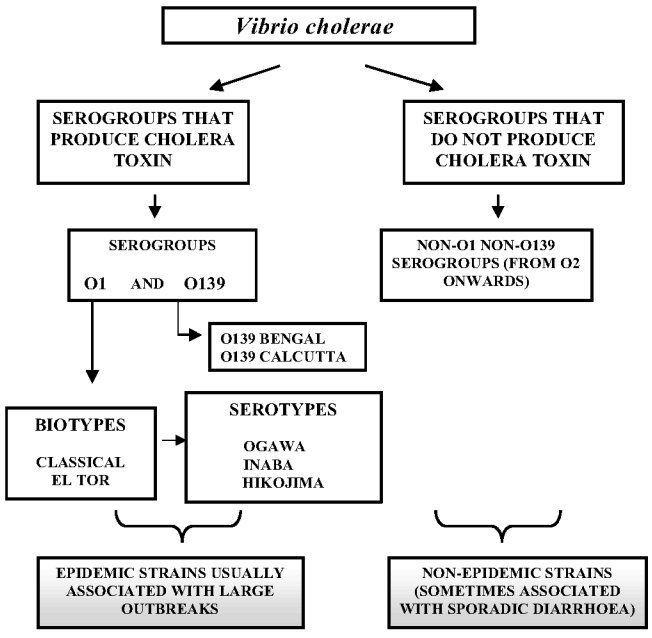

Figure 1. Classification of *Vibrio cholerae* serogroups into epidemic and non-epidemic groups is shown; the respective biotypes within a serotype are also represented. The O1 and O139 serogroups are currently the only ones associated with epidemic cholera.

O139 *wbf* gene cluster linked to O-antigen biosynthesis was found to bear maximum similarity to that of serogroup[13] O22. Later studies involving the screening of 300 non-O1, non-O139 isolates for the presence of virulence genes established that non-O1, non-O139 *V. cholerae* strains with pathogenic potential could emerge by the exchange of O-antigen biosynthesis genes. Non-O1, non-O139 serogroups were seen to possess O1-like genetic backbones in the *wbf* region, while other virulence encoding genes were found to be diverse.

The presence of sequences bearing similarity to known insertion elements has highlighted the probable role of transposition events in the swapping of O1 for O139-specific gene clusters. Later analyses elucidated that the genes responsible for O139 specificity are similar to those of the serogroup O22, wherefrom the O139 serotype could have emerged by horizontal gene transfer[12,13].

Virus-mediated virulence: The CTX phage and cholera toxin

The origination of pathogenic strains of *V. cholerae* from their innocuous marine ancestors remains an enigma in the field of cholera research. Gradual evolutionary change marked by the acquisition of virulence elements and their adaptation to the *V. cholerae* host are now considered probable episodes that could have generated pathogenic strains. Undoubtedly, gene-transfer events have mediated the most significant evolutionary milestones that marked what we choose to refer to as the 'pathogenization' of *V. cholerae*.

This point is well exemplified by the principal virulence determinant in toxigenic *V. cholerae* – the cholera toxin (CT) that is directly responsible for the severe secretory diarrhoea that characterizes cholera. CT is a two-subunit toxin encoded by the genes *ctxA* and *ctxB*[14,15]. These genes are not part of the innate bacterial genome. They are instead borne on the genome of a filamentous lysogenic phage – the CTXΦ[15]. Thus, the most important genes of *V. cholerae* were acquired by the organism through the involvement of a phage. Intriguingly, the *ctxAB* genes apparently seem to have no function for the phage as well and a comparative analysis of all the phage genes suggests that the phage itself might have acquired the toxin genes from elsewhere[16].

CT is an ADP ribosylating toxin that is well characterized in terms of sequence and structure[14,17]. When the regions neighbouring the *ctxAB* genes were sequenced and analysed, it became evident that they were borne on a distinctive genetic element, believed at that time to be a transposon-mediated cassette and named as the CTX genetic element[18,19]. In the 5′ region of the toxin genes other genes like *zot* and *ace* were identified, which were also initially reported to encode toxins (zonula occludens toxin and accessory cholera enterotoxin)[20,21]. The *cep* gene was

so designated because of its pilin-encoding property (core-encoded protein)[19].

The path-breaking discovery of Waldor and Mekalanos[15] that the CTX element corresponded to a phage genome, came much later in time. Their set of elegant experiments showed that the CTX element was transmissible and this led them to prove that under appropriate conditions, the CTX element could give rise to viral particles which could be visualized by electron microscopy[15]. The phage genome consists of two regions, RS2 and core; the former represents a site-specific recombination system that allows the lysogenic phage to integrate at specific sites on the host chromosome, and the latter consists of a retinue of genes, including the *ctxA, ctxB, zot, ace, psh, cep* and *orfU*[15,22].

The RS2 region consists of the *rstR, rstA* and *rstB* genes. The *rstA* gene product is responsible for phage DNA replication, *rstB* is involved in site-specific integration of the phage and *rstR* is a repressor of *rstA* function[22]. The *rstR* gene is transcribed in a direction opposite to that of other genes of the phage. The *rstR* gene is flanked by two intergenic regions, ig-1 and ig-2. The ig-2 region carries the *rstA* promoter. It has been experimentally inferred that the *rstR* proteins are biotype specific, i.e. classical *rstA* can be repressed by classical *rstR* but not by El Tor *rstR*[23]. The *rstR* genes exhibit the phenomenon of heteroimmunity, which means that El Tor *V. cholerae* resist superinfection with another El Tor-derived CTXΦ, but classical *V. cholerae* will allow the integration of an El Tor-derived CTXΦ[23]. The RS2 module of the phage is preceded immediately by a similar element[22,24], the RS1, which has three genes identical in sequence to the corresponding genes in the RS2 module, these are the *rstR*1, *rstA*1, *rstB*1. There is one gene in RS1 that has no counterpart in the RS2 module: this is the *rstC* gene[22,24]. The RS1 was experimentally proven a satellite phage that could exploit the morphogenesis genes of the CTXΦ to exist as a phage itself[25]. The function of RstC was recently elucidated to be that of an antirepressor which could counteract the activity of the *rstR* gene product. RstC was also shown to induce *ctxAB* expression, thus contributing to virulence in *V. cholerae*[26].

The work by Waldor and Mekalanos[15] redefined the roles of the core genes of the CTX genome to those encoding the morphogenesis and structural proteins of the phage rather than toxins; thus, even while their names are retained as such, they currently seem to be misnomers, unless proven to be proteins of dual function.

The expression of *ctxAB* is a coordinated multi-step process that requires the elaboration of a number of virulence factors[27]. One such cardinal step is the successful colonization of the host epithelium by toxigenic bacteria, and this is mediated by their filamentous surface structures: the pili or fimbriae. The toxin co-regulated pilus is a well-characterized surface organelle of *V. cholerae*. It is a bundle-forming pilus which is coordinately expressed with CT from which it derives the name toxin co-regulated pilus. TCP expression is regulated by ToxR, a transmembrane component that activates CT in response to appropriate environmental signals[28]. A cluster of 15 genes comprises the Tcp operon[29]. TcpA, the protein that constitutes the pilus structure, bears homology to bacterial type-IV pilins involved in colonization[30]. The other genes in the cluster are involved in the biogenesis, secretion and export of TcpA[29]. Another potential colonization factor of *V. cholerae* designated as *acf* for Accessory Colonization Factor, also regulated by ToxR, is present immediately downstream of the TCP group of genes; this also includes a cluster of genes[31]. The *ctx* module is synchronized in its function by a regulatory arrangement designated as the cascade system. The regulatory cascade induces the *ctx* module in response to environmental stimuli, and thus serves to control virulence in response

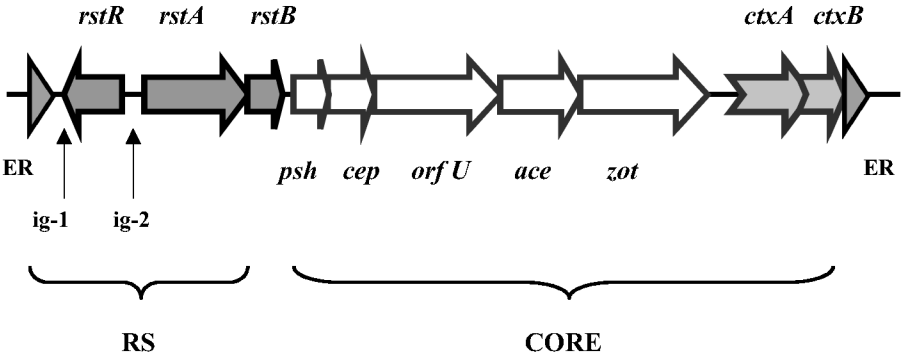

Figure 2. Schematic of CTXΦ phage genome. Genes encoding cholera toxin, *ctxAB* appear at the 3′ (right) end of the genome. The *rstR*, *rstA* and *rstB* genes constitute the RS region (shaded arrows) that is responsible for the site-specific integration and replication of the phage. The CORE region of the phage encodes phage coat and morphogenesis proteins. ER represents End Repeat, a 14 bp sequence that flanks either end of the phage genome and represents the site-specificity of the phage in integrating to the host chromosome.

to environmental conditions like temperature, pH, oxygen concentration, etc. The regulation involves the expression of the toxin genes, the toxin co-regulated pilus and metabolic genes like *aldA* and *tag*. The *tox* genes (*toxT*, *toxS* and *toxR*) are part of the cascade. ToxS is a 'sensory' membranous protein that activates ToxR, also a membrane-spanning protein. ToxT is present in the cytoplasmic matrix and is induced by ToxR; ToxT regulates the expression of the *ctx* and *tcp* genes[27-29].

A swarm of virulence genes: The *Vibrio* Pathogenicity Island

An interesting aspect of toxigenic *V. cholerae* is that the receptor for CTXΦ entry into the bacterium – the TCP, which is involved in the colonization of the bacteria on the human gut epithelium as well, and the ACF are among the genetic modules that have been acquired by the *V. cholerae* genome from other bacterial donors by horizontal gene transfer[32]. TCP and ACF are borne on the genetic island designated as the Vibrio Pathogenicity Island (VPI) – a hallmark of epidemic and pandemic *V. cholerae* strains. Investigations into the relationship between pathogenic and non-pathogenic strains of *V. cholerae* revealed that a pathogenicity island (PAI) is present in the toxigenic strains alone[32]. The ~ 40 kb VPI has a low GC content of 35%, suggesting acquisition from another source by horizontal gene transfer. Putative integrase and transposase genes and flanking *att* sites were present in VPI and could possibly be of phage origin[32].

Polymerase chain reaction (PCR) and Southern hybridization assays revealed that this gene cluster was absent in non-toxigenic environmental strains of *V. cholerae*, but were invariably present in all epidemic and pandemic strains tested[33]. Having established that the VPI is a necessary element for epidemic and pandemic strains, it was postulated that the evolution of toxigenic strains from non-toxigenic ones must be a multi-step process, the initial step of which would be the acquisition of the VPI[34]. This would lead to the expression of the *tcp*, which would in turn facilitate the acquisition of the CTXΦ, thus providing the genes for cholera toxin. The extent of transfer of these virulence-related elements was demonstrated convincingly when Chakraborty *et al.*[35] reported the presence of the toxin co-regulated pilus as well as the cholera toxin genes in environmental isolates that contained neither the O1 nor the O139 antigen of *V. cholerae*. This study proved that virulence genes were not exclusively associated with clinical strains and also set the tone for the hypothesis that environmental strains could act as reservoirs for virulence genes. This trail of thought was re-emphasized in a later study by Mukhopadhyay *et al.*[36], where intact or nearly intact VPI islands were found by PCR assays in a set of environmental isolates tested. Apart from this, the studies also reported the finding of three putative unreported *rstR* repressor genes, as well as a new *tcpA* allele[36].

TcpA variants in toxigenic non-O1, non-O139 serogroup isolates have been analysed[37]. Since TcpA is involved in interaction with the immune system of the host, selection favours diversity at the exposed regions of the protein. This is reflected in the sequence diversity of various *tcpA* variants at the carboxyl region. An assessment of the diverse variants has, however, revealed that they are all capable of functioning as colonization factors. Serogroups like O141, O37 and O8 have been shown to possess three novel *tcpA* sequences[37].

Recently, a novel PAI has been detected in O1 serogroup isolates[38]. This 57.3 kb island has been designated as VPI-2 and is conspicuously absent in non-toxigenic isolates. The presence of a phage–like integrase, insertion at a tRNA site, direct repeats at the integration sites, and a low GC content of 42% are features of this island.

A gene snare in *V. cholerae*: The Integron Island

Rampant virulence gene-acquisition was yet again demonstrated when a unique class of integrons was found to

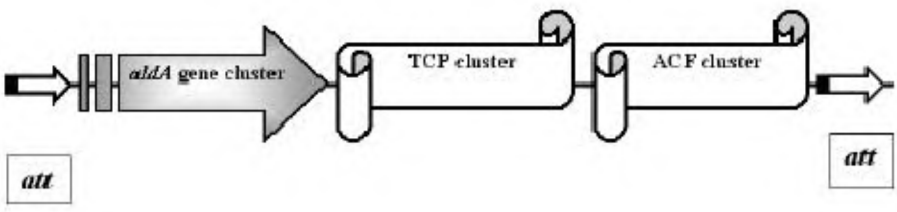

Figure 3. Schematic of the Vibrio Pathogenicity Island; a genomic landmark generally found in epidemic strains. Analyses have shown that some of the genes in this cluster are present, as are the cholera toxin-encoding genes; in many environmental isolates. The boxed 'att' regions represent attachment sites characterized by signature sequences. The *aldA* gene cluster consists of a set of genes found upstream of the TCP cluster in epidemic *V. cholerae*. The TCP cluster comprises a group that encodes the biosynthesis, secretion and structure of the proteins associated with the Toxin-Coregulated Pilus. The ACF cluster represents a set of genes that help the organism in effective adherence and colonization to the gut epithelial surface.

be present in the *V. cholerae* genome. Among bacterial pathogens, the spread of antibiotic-resistance genes is extensive and horizontal transfer is a successful means of achieving this. Molecular analyses of the resistance genes in various microbial pathogens have revealed the presence of a class of genetic elements called integrons[39]. These elements have been frequently reported from antibiotic-resistant strains and have the unique capability of capturing and functionally exploiting drug-resistance genes. In general, integrons are chromosomal elements that have the ability to capture open reading frames and convert them to functional genes[39,40]. The insertion of a gene or reading frame occurs by means of site-specific recombination mediated by the enzyme integrase that is encoded by the recipient. Integron-infused gene cassettes possess a characteristic structure. A central reading frame is flanked on both sides by imperfect repeat elements. The boundaries of the integrated cassettes are defined by sequences called core sites. These display the amino acid sequence GTTRRRY and are the targets of recombination[40,41]. Mazel *et al.*[41] have identified the presence of a previously unreported integrase gene in *V. cholerae*. Repeat sequences designated as VCR clusters were identified in many species of the genus *Vibrio*. The VCRs are constituted by a group of 123–126 bp sequences of imperfect dyad symmetry[41]. Analysis of the VCRs proved that they are similar in organization to integron structures associated with drug resistance. The VCRs are present in a broad range of vibrios, and it appears that they evolved much before the emergence of antibiotic resistance. This finding raises interesting questions regarding the evolution and role of integrons[41,42].

PCR and Southern hybridization assays over a range of *V. cholerae* serotypes from geographically and temporally distinct isolates, have shown that the VCR seems to be a ubiquitous element. This suggests that the acquisition event must have occurred before the speciation of *V. cholerae*. In fact, other species in the genus also show such genetic elements. It is highly interesting that even non-pathogenic serotypes of *V. cholerae* harbour the integron. The presence of the element in a *V. metschnikovii* isolate from the 1800s has reconfirmed the idea that the integrons are indeed of ancient origin[41,42]. A single ancestral VCR predecessor is suggested by the analysis of phylogenetic patterns of the members of the genus *Vibrio*[43]. Chromosomal integrons have been presumed to act as large gene pools for horizontal gene transfer among the members of the genus *Vibrio*. It has been suggested that the chromosomal super integron of *V. cholerae* could serve as the means to acquire new genes, thus helping the bacterium in adaptation to new environments and in its evolution as a pathogen. Capture of genes by the super-integron could ensure or enhance their expression, thus leading to a gradual change or increase in the protein repertoire of the bacterium, thus facilitating its evolution[44].

The *V. cholerae* genome: Two's company

The presence of integrons was reconfirmed by the cholera-sequencing project that was initiated[45] in 1996. When scientists started out on the sequencing work, the *V. cholerae* genome was understood to possess a single 3.2 kb circular chromosome based on the physical map of the genome constructed using the pulsed field gel electrophoresis (PFGE) technique[46]. Lack of adherence to that size while the sequencing work gained momentum led to some doubts on the number of base pairs of DNA in the cholera genome. At this point of time was reported a surprising and interesting observation: the *V. cholerae* genome was bi-chromosomal[47]. This was revealed in a simple PFGE experiment where total genomic DNA was electrophoresed alongside a restriction enzyme-digested sample of the same. The untreated DNA migrated to two mega-sized bands, while the latter appeared as 10 bands[47]. Subsequent analyses of a similar kind with other *V. cholerae* strains as well with *V. mimicus* revealed that all these carried dual-chromosome genomes[47]. Thus, dual chromosomes were probably a trait of the genus *Vibrio*, and this was confirmed in later analyses[48]. Of the two chromosomes, the larger one has a size of 2.4 Mb, while the smaller one is 1.6 Mb. This finding raised the question as to whether the smaller chromosome could be a megaplasmid[47,49]. The second replicon is quite large, and accounts for ~ 40% of the bacterial genome. Moreover, all representative serotypes contain both the replicons. These reasons, along with the fact that each replicon contains unique genes, and the fixed relative stoichometry of the two led to the conclusion that the smaller replicon was more likely to be a chromosome than a plasmid[47].

Decrypting the cholera code: The *V. cholerae* sequencing project

An important milestone in the field of cholera research was achieved when the whole genome of a seventh pandemic cholera strain was completed[49]. The complete genome sequence of both chromosomes of a prototype *V. cholerae* strain – N16961 was published by the Institute for Genomic Research (TIGR) in the year 2000. The larger chromosome (Chr I) contains most of the genes involved in cell division, transcription and translation. The major pathogenicity-associated elements like the VPI and CTX element, are located on this chromosome. The smaller chromosome has the majority of hypothetical genes, the functions of which are not yet well understood[49]. Sequencing confirmed the presence of the Integron Island, and this was seen to be located on Chr II, which also consists of genes generally associated with plasmids. There are over a hundred reading frames that have duplicate copies in both the chromosomes. The CTX element is present as a single copy in the sequenced El Tor strain, while the

classical biotype strains carry two copies of the CTX phage; one in each chromosome. Various features of the smaller chromosome led to the proposal that it could originally have been a megaplasmid captured by an ancestral *Vibrio* sp. Phylogenetic analysis shows that a gene near the putative replication origin of Chr II shows more affinity to plasmid cognates, unlike Chr I. Moreover, genes encoding ribosomal RNAs are found in Chr I only. These, together with the fact that Chr I contains most of the functional genes are best explained by the 'megaplasmid hypothesis'. However, the GC content of the two chromosomes is almost the same; this would mean that the two plasmids have been cohabitants for a considerably long evolutionary period. This conflicts the megaplasmid hypothesis. This led Waldor and Raychaudhuri[50] to suggest the possibility that the small chromosome could have arisen by excision from a single large ancestral chromosome. The presence of the two chromosomes could be evolutionarily advantageous for the organism, given the fact that it has a complex lifestyle involving the varied environments of the human gut and the estuarine habitat[50]. Heidelberg *et al.*[49] have also proposed that single chromosome cells may be generated in response to environmental signals, such that metabolic functions would continue, but valuable nutrients would not be used up in replicating the cells. Such metabolically active and replication-defective cells called 'drones' could aid the normal cells in bacterial biofilms. The *V. cholerae* genome has a large gamut of genes that encode proteins with a wide range of substrate specificities. The organism is greatly benefited by these, as it thrives in the varied environments of the gut and the estuary[49,50]. Chitin that is found in the zooplanktonic masses, with which the bacterium is closely associated in the open waters, is effectively utilized by the presence of a metabolic pathway that degrades it in a manner similar to that of its marine counterpart, *V. furnissi*[49,51].

Cholera chip: The *V. cholerae* microarray

The post-sequencing phase of *V. cholerae* genomics saw the development and application of a potentially effective tool for cholera research: a *V. cholerae* microarray based on the El Tor strain that was sequenced[52]. The genetic makeup of classical, other El Tor and non-toxigenic environmental isolates has been compared to the N16961 strain, a 7th pandemic clinical isolate. The analyses have revealed that certain genes or groups of genes are absent in some of these in comparison to the latter. These fall into four groups, namely (i) genes that are absent in classical strains but present in all El Tor strains; (ii) genes present only in pandemic strains; (iii) genes present only in the 7th pandemic El Tor strains and (iv) genes that are not seen in individual strains[52]. Seven genes were seen to be absent in the classical strains, of which two were on the large chromosome; in the region encoding the RTX toxin. Earlier studies had suggested that the classical strains were lacking in the *rtxC* gene; a gene that is presumed to encode an acylation enzyme that activates the toxin[53]. Microarray analysis revealed that the deletion extends into the *rtxA* and *B* genes. The other five are seen in the smaller chromosome, of which three are hypothetical and the other two are putative acetyl transferases that reside on the Integron Island. Since few genes uniquely distinguish the classical from the El Tor strains, despite the fact that the two are believed to have evolved from separate lineages, the ancestral environmental strain from which the two evolved may have a greater resemblance to the El Tor strain than previously understood[52]. This yet again underscores the importance of foreign gene-acquisition events that have aided the pathogenization process of *V. cholerae*.

Some of the genes differentiate the pandemic strains (classical and 7th pandemic) from environmental and prepandemic El Tor strains. The early El Tor and prepandemic strains carry the TCP island genes. A non-toxigenic Gulf Coast isolate was found to be TCP +, but CTXΦ-. Twenty-two genes that are absent in classical, prepandemic TCP + El Tor isolates and a prepandemic environmental El Tor isolate are seen in the 7th pandemic strains. These are seen in clusters, a block of 11 genes characterized by low GC content has been designated as VSP I (for Vibrio Seventh Pandemic Island I)[52]. Yet another chromosomal island uniquely found in the 7th pandemic isolates is also characterized by a low GC content and is designated as VSP II. Both these have been proposed to have been acquired by the 7th pandemic strains by horizontal transfer. Many of the genes in these islands are hypothetical or conserved hypothetical genes[52].

Conclusion

Recent accomplishments in the genetic front of cholera research have accelerated our understanding of the long history of cholera. Independent findings by various workers have established that the *V. cholerae* genome is dotted with genetic elements of 'foreign' origin. Horizontal gene transfer and its influence on pathogenicity have found important examples in the evolution of toxigenic strains of *V. cholerae*. It has also been proven that despite being an ancient bacterium, *V. cholerae* is an 'emerging' pathogen owing to the exchange of genetic material, as evidenced by the appearance of toxigenic O139 outbreaks. Many mysteries regarding the origin of pathogenic strains remain, and as cholera research now graduates to proteome research, we can expect some more insights into the origins of this pathogen, which was once an innocuous estuarine-dweller.

1. WHO weekly epidemiological record, 2002, vol. 77, pp. 257–268.

2. Garrity, G. M., Winters, M. and Searies, D. B., Taxonomic outline of the Procaryotes. *Bergey's Manual of Systematic Bacteriology Release1.0*., Springer–Verlag, 2001.
3. Shimada, T. *et al.*, Extended serotyping scheme for *Vibrio cholerae*, *Curr. Microbiol.*, 1994, **28**, 175–178.
4. Kaper, J. B., Morris, Jr. J. G. and Levine, M. M., *Cholera Clin. Microbiol. Rev.*, 1995, **8**, 48–86.
5. Blake, P. A., *Vibrio cholerae and Cholera: Molecular to Global Perspectives* (eds Wachsmuth, I. K., Blake, P. A. and Olsvik, Ø.), ASM Press, 1994, pp. 293–295.
6. Ramamurthy. T. *et al.*, Emergence of novel strain of *Vibrio cholerae* with epidemic potential in southern and eastern India. *Lancet*, 1993, **341**, 703–704.
7. Albert, M. J., Siddique, A. K., Islam, M. S., Faruque, A. S. G., Ansaruzzaman, M., Faruque, S. M. and Sack, R. B., Large outbreak of clinical cholera due to *Vibrio cholerae* non-O1 in Bangladesh. *Lancet*, 1993, **341**, 704.
8. Shimada, T., Nair, G. B., Deb, B. C., Albert, M. J., Sack, R. B. and Takeda, Y., Outbreaks of *Vibrio cholerae* non-O1 in India and Bangladesh. *Lancet*, 1993, **341**, 1347.
9. Swerdlow, D. L. and Isaäcson, M., In ref. 5, pp. 297 307.
10. Mukhopadhyay, A. K. *et al.*, Molecular epidemiology of reemergent *Vibrio cholerae* O139 Bengal in India. *J. Clin. Microbiol.*, 1998, **36**, 2149–2152.
11. Johnson, J. A., Salles, C. A., Panigrahi, P., Albert, M. J., Wright, A. C., Johnson, R. J. and Morris, Jr. J. G., Vibrio cholerae O139 synonym Bengal is closely related to Vibrio cholerae El Tor but has important differences. *Infect. Immunol.*, 1994, **62**, 2108–2110.
12. Yamasaki, S., Shimizu, T., Hoshino, K., Ho, S., Shimada, T., Nair, G. B. and Takeda, Y, The genes responsible for O-antigen synthesis of *Vibrio cholerae* O139 are closely related to those of *Vibrio cholerae* O22. *Gene*, 1999, **237**, 321–332.
13. Li, M., Shimada, T., Morris, Jr. J. G., Sulakvelidze, A. and Sozhamannan, S., Evidence for the emergence of non-O1 and non-O139 *Vibrio cholerae* strains with pathogenic potential by exchange of O-antigen biosynthesis regions. *Infect. Immunol.*, 2002, **70**, 2441–2453.
14. Zhang, R. G., Scott, D. L., Westbrook, M. I., Nance, S., Spangler, B. D., Shipley, G. G. and Westbrook, E. M., The three-dimensional crystal structure of cholera toxin. *J. Mol. Biol.*, 1995, **251**, 563–573.
15. Waldor, M. K. and Mekalanos, J. J., Lysogenic conversion by a filamentous bacteriophage encoding cholera toxin. *Science*, 1996, **272**, 1910–1914.
16. Boyd, E. F., Heilpern, A. J. and Waldor, M. K., Molecular analyses of a putative CTXΦ precursor and evidence for independent acquisition of distinct CTXΦs by toxigenic *Vibrio cholerae*. *J. Bacteriol.*, 2000, **182**, 5530–5538.
17. Hirst, T. R., Cholera toxin and Escherichia coli heat labile enterotoxin. In *The Comprehensive Sourcebook of Bacterial Protein Toxins* (eds Alouf, J. E. and Freer, J. H.), Academic Press, 1999, pp. 104–129.
18. Mekalanos, J. J., Duplication and amplification of toxin genes in *Vibrio cholerae*. *Cell*, 1983, **35**, 253–263.
19. Pearson, G. D. N., Woods, A., Chiang, S. L. and Mekalanos, J. J., CTX genetic element encodes a site – specific recombination and an intestinal colonization factor. *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 3750–3754.
20. Fasano, A., Baudry, B., Pumplin, D. W., Wasserman, S. S., Tall, B. D., Ketley, J. M. and Kaper, J. B., *Vibrio cholerae* produces a second enterotoxin which affects intestinal tight junctions. *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 5242–5246.
21. Trucksis, M., Galen, J. E., Michalski, J., Fasano, A. and Kaper, J., Accessory cholera enterotoxin (Ace), the third toxin of a *Vibrio cholerae* virulence cassette. *Proc. Natl. Acad. Sci. USA*, 1993, **90**, 5267–5271.
22. Waldor, M. K., Rubin, E. J., Pearson, G. D. N., Kimsey, H. and Mekalanos, J. J., Regulation, replication and integration functions of the *Vibrio cholerae* CTXΦ are encoded by region RS2. *Mol. Microbiol.*, 1997, **24**, 1910–1914.
23. Kimsey, H. H. and Waldor, M. K., CTXΦ immunity: application in the development of cholera vaccines. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 7035–7039.
24. Campos, J., Fando, R., Silva, A., Rodriguez, B. L. and Benitez, J. A., Replication function of the RS1 element associated with *Vibrio cholerae* CTXΦ prophage 1998. *FEMS Microbiol. Lett.*, 1998, **164**, 141–147.
25. Faruque, S. M. *et al.*, RS1 element of *Vibrio cholerae* can propagate horizontally as a filamentous phage exploiting the morphogenesis genes of CTXΦ, *Infect. Immunol.*, 2002, **70**, 163–170.
26. Davis, B. M., Kimsey, H., Kane, A. V. and Waldor, M. K., A satellite phage – encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. *EMBO J.*, 2002, **21**, 4240–4249.
27. Peterson, K. M. and Mekalanos, J. J., Characterization of the *Vibrio cholerae* ToxR regulon: Identification of novel genes involved in intestinal colonization. *Infect. Immunol.*, 1988, **56**, 2822–2829.
28. Miller, V. L. and Mekalanos, J. J., Synthesis of cholera toxin is positively regulated at the transcriptional level by toxR. *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 3471–3475.
29. Kaufman, M. R. and Taylor, R. K., The Toxin co regulated Pilus: Biogenesis and function. In ref. 5, pp. 187–202.
30. Shaw, C. E. and Taylor, R. K., *Vibrio cholerae* O395 tcpA pilin gene sequence and comparison of predicted protein structural features to those of type 4 pilins. *Infect. Immunol.*, 1990, **58**, 3042–3049.
31. Brown, R. C. and Taylor, R. K., Organization of tcp, acf and toxT genes within a ToxT-dependent operon. *Mol. Microbiol.*, 1995, **16**, 425–439.
32. Karaolis, D. K. R., Johnson, J. A., Bailey, C. C., Boedeker, E. C., Kaper, J. B. and Reeves, P. R., A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 3134–3139.
33. Vital Brazil, J. M., Alves, R. M., Rivera, I. N., Rodrigues, D. P, Karaolis, D. K. and Campos, L. C., Prevalence of virulence-associated genes in clinical and environmental *Vibrio cholerae* strains isolated in Brazil between 1991 and 1999. *FEMS Microbiol. Lett.*, 2002, **215**, 15–21.
34. Boyd, E. F., Moyer, K., Shi, L. and Waldor, M., Infectious CTXΦ and the Vibrio Pathogenicity Island prophage in *Vibrio mimicus*: evidence for recent horizontal transfer between *V. mimicus* and *V. cholerae*. *Infect. Immunol.*, 2000, **68**, 1507–1513.
35. Chakraborty, S. *et al.*, Virulence genes in environmental strains of *Vibrio cholerae*, *Appl. Environ. Microbiol.*, 2000, **66**, 4022–4028.
36. Mukhopadhyay, A. K., Chakraborty, S., Takeda, Y., Nair, G. B. and Berg, D. E., Characterization of VPI pathogenicity island and CTX phi prophage in environmental strains of *Vibrio cholerae*. *J. Bacteriol.*, 2001, **183**, 4737–4746.
37. Boyd, E. F. and Waldor, M. K., Evolutionary and functional analyses of variants of the toxin-co regulated pilus protein TcpA from toxigenic *Vibrio cholerae* non-O1/non–O139 serogroup isolates. *Microbiology*, 2002, **148**, 1655–1666.
38. Jermyn, W. S. and Boyd, E. F., Characterization of a novel *Vibrio* pathogenicity Island (VPI-2) encoding neuraminidase (nanH) among toxigenoic *Vibrio cholerae* isolates. *Microbiology*, 2002, **148**, 3681–3693.
39. Rowe Magnus, D. A., Guérout, A. M. and Mazel, D., Super-integrons. *Res. Microbiol.*, 1999, **150**, 641–651.
40. Rowe-Magnus, D. A., Guérout, A. M and Mazel. D., Bacterial

resistance evolution by recruitment of super-integron gene cassettes. *Mol. Microbiol.*, 2002, **43**, 1657–1669.
41. Mazel, D., Dychinco, B., Webb, V. A. and Davies, J., A distinctive class of integron in the *Vibrio cholerae* genome. *Science*, 1998, **280**, 605–607.
42. Mazel, D. and Davies, J., Gene capture in *Vibrio cholerae*: Response. *Trends Microbiol.*, 1999, **93**, 95.
43. Clark, C. A., Purins, L., Kaewrakon, P., Focareta, T. and Manning, P. A., The *Vibrio cholerae* O1 chromosomal integron. *Microbiology*, 2000, **146**, 2605–2612.
44. Rowe-Magnus, D. A., Guérout, A., Ploncard, P., Dychinco, B., Davies, J. and Mazel, D., The evolutionary history of superintegrons provides an ancestry for multiresistant integrons. *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 652–657.
45. Lewis, R., TIGR introduces *Vibrio cholerae* genome. *Scientist*, 2000, **14**, 8 (news item featuring an interview with Claire Fraser).
46. Majumder, R., Sengupta, S., Khetawat, G., Bhadra, R. K., Roychoudhary, S. and Das, J., Physical Map of the genome of *Vibrio cholerae* 569B and localization of genetic markers. *J. Bacteriol.*, 1996, **178**, 1105–1112.
47. Trucksis, M., Michalski, J., Deng, Y. K. and Kaper, J. B., Accessory cholera enterotoxin (Ace), the third toxin of a *Vibrio cholerae* virulence cassette. *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 14464–14469.
48. Tagomori, K., Iida, T. and Honda, T., Comparison of genome structures of vibrios, bacteria possessing two chromosomes. *J. Bacteriol.*, 2002, **184**, 4351–4358.
49. Heidelberg, J. F. *et al.*, DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature*, 2000, **406**, 477–483.
50. Waldor, M. K. and Raychaudhuri, D., Treasure trove for cholera research. *Nature*, 2000, **406**, 469–470.
51. Schoolnik, G. K. and Yildiz, F. H., The complete genome sequence of *Vibrio cholerae*: a tale of two chromosomes and of two lifestyles. *Genome Biol.*, 2000, **1**, 1016.1–1016.3.
52. Dziejman, M., Balon, E., Boyd, D., Fraser, C. M., Heidelberg, J. F. and Mekalanos, J. J., Comparative genomic analysis of *Vibrio cholerae*: Genes that correlate with cholera endemic and pandemic disease. *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 1556–1561.
53. Lin, W. *et al.*, Identification of a *Vibrio cholerae* RTX toxin gene cluster that is tightly linked to the cholera toxin prophage. *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 1071–1076.

Received 26 June 2003; accepted 17 September 2003