Microbiology (1996), **142**, 2165–2174 Printed in Great Britain

A putative integrase gene defines the distal end of a large cluster of ToxR-regulated colonization genes in *Vibrio cholerae*

Michael E. Kovach, Megan D. Shaffer and Kenneth M. Peterson

Author for correspondence: Kenneth M. Peterson. Tel: +1 318 675 5753. Fax: +1 318 675 5764. e-mail: kpeter©pop3.lsumc.edu

Department of Microbiology and Immunology, Louisiana State University Medical Center-Shreveport, Shreveport, LA 71130, USA

A large cluster of virulence genes encoding proteins involved in *Vibrio cholerae* accessory colonization factor (ACF) expression and toxin-coregulated pilus (TCP) biogenesis is flanked by sequences that resemble bacteriophage attachment (*att*) half-sites. Adjacent to the *attL*-like site is a gene (*int*) that encodes a protein related to the integrase family of site-specific recombinases. The putative vibrio integrase appears to be most closely related to the *Escherichia coli* cryptic prophage (CP4-57) integrase protein (52% identity, 73% similarity). Genomic analysis of numerous *V. cholerae* strains (O1, non-O1 and O139) revealed that only vibrios capable of causing epidemic Asiatic cholera possess the TCP-ACF colonization gene cluster in association with the integrase. The fact that the integrase gene is absent in avirulent strains suggests that epidemic strains of *V. cholerae* obtained the TCP-ACF colonization gene cluster via horizontal transfer.

Keywords: *Vibrio cholerae*, colonization, integrase, ToxR regulon, mobile genetic element

INTRODUCTION

Vibrio cholerae, a Gram-negative bacterium, is the etiological agent of the severe diarrhoeal disease Asiatic cholera. This syndrome results from the ingestion of food or water contaminated with the bacterium, followed by colonization of the small bowel and the subsequent secretion of cholera toxin (CTX) by the vibrios. Intraintestinal survival of *V. cholerae* is a poorly understood process requiring the expression of a number of coordinately regulated virulence factors (DiRita, 1992) in order to overcome the host's natural defence mechanisms. Prior to intestinal colonization the vibrios must survive exposure to the gastric acid of the stomach, localize in the small bowel, and penetrate the mucous gel protecting the intestinal microvilli. Following penetration of the protective mucous gel, the vibrios colonize the epithelial cell surface of the microvilli. At some point in the colonization process, the synthesis of CTX, toxin-coregulated pilus (TCP) (Taylor *et al.*, 1987) and an additional accessory colonization factor (ACF) (Peterson & Mekalanos, 1988) occurs. Production of TCP and ACF enhance the ability of *V. cholerae* to successfully colonize the intestinal epithelium (Taylor *et al.*, 1987; Peterson & Mekalanos, 1988). Recent studies have shown that the *tcp* and *acf* gene clusters are physically linked on the *V. cholerae* chromosome, suggesting that the TCP and ACF factors may interact in some manner to promote successful colonization of the small bowel (Everiss *et al.*, 1994a, b). The action of CTX triggers the characteristic, profuse diarrhoea associated with this disease, promoting dissemination of the organisms back into the environment. Synthesis of CTX, TCP and ACF are controlled at the transcriptional level by a DNA-binding protein, ToxR, in response to environmental stimuli (Miller & Mekalanos, 1988). Transcriptional activation of these *V. cholerae* virulence genes can occur either directly by ToxR binding to the heptanucleotide TTTTGAT (Miller & Mekalanos, 1984), or indirectly via ToxR-regulated control of a second transcriptional activator, ToxT (DiRita *et al.*, 1991; DiRita, 1992). ToxT is a member of the AraC family of transcriptional activators that regulates the expression of a number of genes involved in *V. cholerae* virulence (*acfA*, *tcpA*, *tcpC* and *tcpI*) as well as several genes whose roles in virulence have not been established (*aldA* and *tagA*) (DiRita *et al.*, 1991; DiRita, 1992; Higgins *et al.*, 1992; Parsot & Mekalanos, 1991, 1992). The activation of *V. cholerae* virulence factors by ToxT, a transcriptional activator dependent on ToxR for expression, represents

Abbreviations: ACF, accessory colonization factor; CTX, cholera toxin; TCP, toxin-coregulated pilus.

The GenBank accession number for the nucleotide sequence reported in this paper is U02372.

an environmentally responsive regulatory cascade that is responsible for the coordinate activation of genes involved in *V. cholerae* colonization of the small bowel (DiRita, 1992). These genes belong to the ToxR regulon and are referred to as ToxR-activated genes (TAGs) (Peterson & Mekalanos, 1988).

In numerous pathogenic bacteria virulence genes have been identified on mobile genetic elements, which may facilitate transfer of these genes between bacterial species (Finlay & Falkow, 1989). Work by Pearson *et al.* (1993) has demonstrated that the *V. cholerae* CTX genes encode two additional toxins: the zonula occludens toxin (*zot*) (Fasano *et al.*, 1991; Baudry *et al.*, 1992) and an accessory cholera enterotoxin (*ace*) (Trucksis *et al.*, 1993). Both these toxins, as well as a core encoded pilin (*cep*), are located within a mobile genetic element. Karaolis *et al.* (1995) have inferred, based on polymorphism within the *V. cholerae* housekeeping gene encoding the aspartate-semialdehyde dehydrogenase (*asd*), that horizontal transfer of the *V. cholerae* O antigen genes may occur, suggesting that *V. cholerae* could show a high level of intraspecies genetic exchange. Additionally, the atypical *V. cholerae* codon usage for the genes within the TCP locus has led to the proposal that the TCP gene cluster may be part of a larger genetic element (Ogierman & Manning, 1992; Kaufman *et al.*, 1993).

In this report, we describe the identification of a *V. cholerae* gene, *int*, that encodes a protein related to the integrase family of site-specific recombinases. The *V. cholerae int* is the distal marker for a DNA segment of greater than 45 kbp encoding the *V. cholerae* ToxR-activated colonization determinants TCP and ACF. The data presented herein showing the physical linkage of the integrase gene to the ACF and TCP gene clusters supports the hypothesis that these colonization factors are part of a large genetic element involved in intestinal colonization that may have been, at one time, part of a mobile element acquired by *V. cholerae*.

METHODS

Bacterial strains, phages, plasmids, media and reagents. The bacterial strains, phages and plasmids used in this study are listed in Table 1. Non-O1 strains of *V. cholerae* other than O139 have been referred to as 'nonagglutinating' or 'noncholera' vibrios and are typically not associated with epidemic cholera (Morris, 1990). *Escherichia coli* and *Vibrio cholerae* strains were maintained at −70 °C in Luria-Bertani (LB) medium containing 20% (v/v) glycerol (Miller, 1972). *E. coli* strains were cultured in LB broth or on LB agar at 37 °C. *V. cholerae* strains were cultured on LB agar at 37 °C or in LB broth pH 6·5 at 30 °C or LB broth pH 8·4 at 37 °C as previously described (Peterson & Mekalanos, 1988). Antibiotics, when necessary, were used at the following concentrations: ampicillin (Ap) 100 μg ml^{-1}, kanamycin (Km) 30 μg ml^{-1} and streptomycin (Sm) 100 μg ml^{-1}. Restriction endonucleases were purchased from New England Biolabs and T4 DNA ligase was purchased from Promega. Enzymes were used according to the manufacturers' recommendations. Isotopes, [^{35}S]dATPαS and [^{32}P]dATP were purchased from Amersham. All other biochemical reagents were purchased from Sigma unless otherwise stated. Restriction analysis, plasmid construction and molecular biology techniques were performed according to standard procedures (Ausubel *et al.*, 1989; Sambrook *et al.*, 1989). Reagents used in RNA preparation and analysis were treated with diethyl pyrocarbonate (DEPC) prior to use.

Genomic DNA isolation and analysis. Genomic DNA was isolated from 1·5 ml overnight cultures of *V. cholerae* (Ausubel *et al.*, 1989). *V. cholerae* cells were harvested by centrifugation and resuspended in 567 μl TE (10 mM Tris pH 8·0/1 mM EDTA). SDS [30 μl, 10% (w/v) solution] and proteinase K (3 μl, 20 mg ml^{-1} solution) were added, giving final concentrations of 100 μg proteinase K ml^{-1} and 0·5% SDS in 600 μl TE. After incubation at 37 °C for 1 h, one-sixth vol. 5 M NaCl was added and the sample was mixed by inversion. Cetyltrimethylammonium bromide (CTAB)/NaCl solution [80 μl, 10% (w/v) CTAB in 0·7 M NaCl] was added to the sample, mixed by inversion and incubated at 65 °C for 10 min. The solution was extracted once with an equal volume of chloroform/isoamyl alcohol, and once with an equal volume of phenol/chloroform. The DNA was precipitated by adding 0·6 vol. isopropyl alcohol and the chromosomal DNA was collected on a glass rod. The DNA was washed once with 70% (v/v) ethanol and resuspended in TE buffer pH 8·0.

DNA manipulations and analysis. Chromosomal DNA used in slot blot analysis was denatured at 100 °C in 0·4 M NaOH/10 mM EDTA for 10 min and applied to Hybond N+ filters (Amersham) under alkaline conditions (Ausubel *et al.*, 1989). Chromosomal DNA used in Southern analysis was digested with *Dra*I, size-fractionated on a 1·2% (w/v) agarose gel in 1 × TAE (40 mM Tris, 40 mM acetic acid, 2 mM EDTA, pH 8·0) and transferred to a BioBlot-N+ membrane (Costar) by capillary transfer in 10 × SSC (1 × SSC: 0·15 M sodium chloride, 0·015 M sodium citrate, pH 7·0). Filters were baked for 2 h *in vacuo* at 80 °C and prehybridized for 2 h in 6 × SSPE (1 × SSPE: 150 mM NaCl, 10 mM sodium phosphate, 1 mM EDTA, pH 7·4), 5 × Denhardt's solution [0·1% Ficoll (Type 400), 0·1% polyvinylpyrrolidone, 0·1% bovine serum albumin (Fraction V)], 0·5% SDS, 50% (v/v) formamide and 100 μg denatured herring sperm DNA ml^{-1} at 42 °C. The DNA probes used in this study were derived as follows: *tagD*-specific probe, 0·7 kbp *Dra*I fragment from pTAGD-1; *tcpI*-specific probe, ~ 1·0 kbp *Eco*RV fragment from pTCPI-1; *toxT*-specific probe, ~ 0·6 kbp *Hin*dIII/*Eco*RI fragment from pB53.2; *acf*-specific probe, ~ 4·2 kbp *Sac*I fragment from pMEK2; *int*-specific probe, ~ 1·3 kbp *Eco*RV fragment from pMEK3; *ssrA*-specific probe, ~ 1·3 kbp *Dra*I fragment from pMEK3; and the probe derived from *V. cholerae* SG34 downstream of the *ssrA* gene that was used as a probe, ~ 3·0 kbp *Sac*I/*Bgl*II fragment from ϕ19. DNA restriction fragments used as probes in Southern analysis were isolated by agarose gel electrophoresis, purified using GeneClean (Bio101) and labelled with [^{32}P]dATP using a random primer labelling kit (Promega). The unincorporated nucleotides were removed from the oligonucleotide probes by G-50 spin-column purification (Sambrook *et al.*, 1989). The radiolabelled probe was heated for 5 min at 100 °C and added to the prehybridizing filters. Hybridization was continued overnight at 42 °C. The filters were washed as follows: 5 min in 2 × SSC/0·5% SDS at room temperature, 15 min in 2 × SSC/0·1% SDS at room temperature, 2 h in 0·1 × SSC/0·5% SDS at 68 °C and 30 min in 0·2 × SSC/0·1% SDS at 68 °C. The washed filters were air-dried and exposed to Kodak X-OMAT film at −70 °C with an intensifying screen (Fisher Scientific) for 4–16 h.

RNA isolation and analysis. Overnight cultures of *V. cholerae* were grown under ToxR-repressing conditions, in LB pH 8·4 at

Table 1. Bacteria, phages and plasmids used in this study

Strain or plasmid	Relevant genotype or characteristics	Reference or source
Strains		
V. cholerae		
O395	O1, classical/Ogawa, Sm^R	Laboratory stock
KP1.17	O395 *toxR*::pVM55, Ap^R, Sm^R	Laboratory stock
C6709	O1, El Tor/Inaba, Sm^R	Peru*
0872	O1, El Tor/Inaba	Peru*†
ANCO223	O1, El Tor/Ogawa	Peru*
P27459	O1, El Tor/Inaba, Sm^R	Bangladesh*
E7946	O1, El Tor/Ogawa, Sm^R	Bahrain*
AI4456	Non-O1, Sm^R, non-motile	Bangladesh*†
AI1837	O139, Sm^R	Bangladesh*
MO2	O139, Sm^R	India*
MO10	O139	India*
SG34	O56	Waldor & Mekalanos (1994)
VO7	O37	Waldor & Mekalanos (1994)
1074-78	O1, El Tor/Ogawa, nonpathogenic	Levine *et al.* (1982)
1196-78	O1, El Tor/Ogawa, nonpathogenic	Levine *et al.* (1982)
E. coli		
JM109	F′ *traD36 lacI*^{q} Δ(*lacZ*)M15 *proAB*/*recA1 endA1 gyrA96* (Nal^R) *thi hsdR17* (r_k^-, m_k^+) *supE44* (*mcrA*) *relA1* Δ(*lac*–*proAB*)	Promega
Plasmids		
pUC18	*lacZ*′, Ap^R	Yanisch-Perron *et al.* (1985)
pBluescript II KS^+	*lacZ*′, Ap^R	Stratagene
pMEK1	pUC18 containing 16 kbp *Sal*I fragment from φA13, Ap^R	This study
pMEK2	pUC18 containing ~ 4·2 kbp *acfC*$^+$ *Sac*I fragment from φA13, Ap^R	This study
pMEK3	pBluescript II KS^+ containing ~ 4·4 kbp *Bam*HI fragment from pMEK1, Ap^R	This study
pB53.2	pUC18 containing 11 kbp *Eco*RV/*Bam*HI fragment containing the Tn*phoA* fusion junction from KP3.51(*acfB*::Tn*phoA*), Ap^R, Km^R	Everiss *et al.* (1994a)
pTAGD-1	pBluescript II containing *V. cholerae tagD*	Hughes *et al.* (1994)
pTCPI-1	pBluescript II containing *V. cholerae tcpI*	Harkey *et al.* (1994)
pVM55	pJM703.1::*Eco*RI–*Hpa*I(*toxR*), Ap^R	Miller & Mekalanos (1988)
Bacteriophage		
λGEM-11	Lambda replacement cloning vector	Promega
φA13	λGEM-11 containing 16 kbp *acf*$^+$ partial *Sau*3A fragment from *V. cholerae* O395	This study
φ19	λGEM-11 containing ~ 16 kbp *ssrA*$^+$ partial *Sau*3A fragment from *V. cholerae* SG34	This study
φD12-2	λGEM-11 containing ~ 16 kbp *tagD*$^+$ partial *Sau*3A fragment from *V. cholerae* O395.	This study

* Recovered from patients with cholera.

† Serogroup has not been confirmed.

37 °C. These cultures were washed once with an equivalent volume of PBS (Ausubel *et al.*, 1989) pH 7·2, resuspended in an equal volume of LB pH 6·5 and diluted 1:100 in 50 ml LB pH 6·5 (optimal conditions for the expression of ToxR-activated genes) (Peterson & Mekalanos, 1988) and grown with aeration for 4 h at 30 °C (OD_{600} ~ 1·0). Total RNA was isolated from these cultures using the hot phenol method (Aiba *et al.*, 1981). Isolated RNA was purified by centrifugation through a CsCl

gradient (Ausubel *et al.*, 1989), quantitated spectrophotometrically and stored in diethyl-pyrocarbonate-treated water at −70 °C. Total RNA from each *V. cholerae* strain was denatured in 50% formamide at 65 °C for 10 min and applied to nitrocellulose filters under high salt conditions (Ausubel *et al.*, 1989). Nitrocellulose filters containing the transferred RNA were equilibrated in 5× SSC for 15 min, air-dried, baked *in vacuo* at 80 °C for 2 h and prehybridized for 1 h in 6× SSPE, 0·5% SDS, 5× Denhardt's solution, 50% formamide with 100 μg herring sperm DNA ml^{-1} at 42 °C. Radiolabelled probes were added and hybridization continued overnight at 42 °C. The filters were washed for 5 min at room temperature in 1× SSC/0·1% SDS, followed by three 20 min washes at 68 °C in 0·2× SSC/0·1% SDS. The washed filters were air-dried and exposed to Kodak X-OMAT film at −70 °C with an intensifying screen for 4–16 h.

DNA sequencing. The DNA sequence was determined by the dideoxy chain-termination method (Sanger *et al.*, 1977), using synthesized complementary oligonucleotide primers (Integrated DNA Technologies), [^{35}S]dATPαS and a Sequenase kit (United States Biochemical) following the manufacturers' protocols. The nucleic acid and the deduced protein sequence were analysed using the MacVector sequence analysis software (International Biotechnologies). Protein and nucleic acid similarity searches were performed against the non-redundant databases using the National Center for Biotechnology Information (NCBI) network server and the BLAST algorithm (Altschul *et al.*, 1990).

RESULTS

Transcriptional analysis of genes downstream of the ACF gene cluster

We have previously shown that the TCP and ACF gene clusters are physically linked on the *V. cholerae* chromosome and encompass greater than 20 kbp of contiguous DNA sequence encoding at least 21 ToxR-activated genes involved in intestinal colonization (Everiss *et al.*, 1994a). To identify ToxR-activated genes located downstream of the previously characterized ACF genes, a series of overlapping fragments isolated from cloned DNA downstream of the ACF gene cluster were used to perform Northern blot analysis on RNA isolated from the fully virulent classical *V. cholerae* strain O395. ToxR-activated genes are expressed in *V. cholerae* O395 but are not expressed in KP1.17 due to an insertional disruption in the transcriptional activator ToxR that prevents expression of genes belonging to the ToxR regulon. Except for the *toxR* null allele, KP1.17 is otherwise isogenic to strain O395. A recombinant clone was isolated from an O395 genomic library that contained a 16 kbp DNA fragment of the *V. cholerae* chromosome encompassing the *acf* gene cluster and ~ 10 kbp of downstream DNA (Fig. 1). Plasmid pMEK1 DNA (Fig. 1) was digested with *Bam*HI and *Eco*RV, and each of the resulting six DNA fragments were isolated, labelled with [^{32}P]dATP, and used as probes in RNA slot blot analysis. When the probes generated from the 4·4 kbp *Bam*HI DNA fragment (encompassing probes 5 and 6 in Fig. 2) were used in Northern analysis, a hybridization signal was produced in RNA samples isolated from *V. cholerae* O395 and the ToxR-null strain KP1.17 (Fig. 2). The probes hybridizing to the RNA isolated from the ToxR-null strain indicated

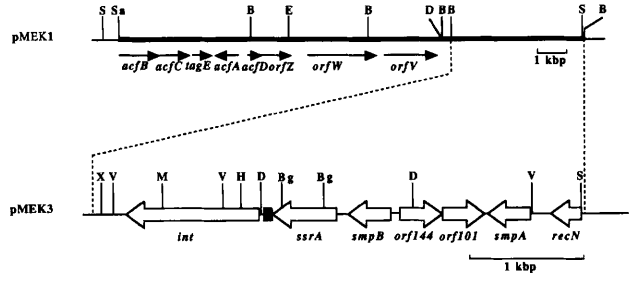

Fig. 1. Physical map of the distal region of the ACF gene cluster. The location of *acf* and additional genes as determined by DNA sequence analysis are indicated. Arrows denote the direction of transcription and the black box indicates the *att* site. Restriction sites: B, *Bam*HI; Bg, *Bgl*II; D, *Dra*I; E, *Eco*RI; H, *Hin*dIII; M, *Msc*I; S, *Sal*I; Sa, *Sau*3A; V, *Eco*RV; X, *Xho*I.

the existence of transcriptionally active, non-ToxR-activated genes downstream of the *acf* gene cluster. These constitutively expressed genes mark the distal limit of the > 45 kbp gene cluster encompassing the ToxR-activated TCP and ACF colonization determinants.

Identification, isolation and nucleotide sequence determination of the *V. cholerae* O395 *int* gene

The 4·4 kbp *Bam*HI DNA fragment from pMEK1 was isolated and ligated into *Bam*HI-digested pBluescript II KS^+ generating pMEK3 (Fig. 1). The complete nucleotide sequences of pMEK1 and pMEK3 were determined using a series of subclones, complementary oligonucleotide primers and exonuclease III deletions (data not shown). Analysis of the nucleotide sequence of pMEK3 revealed a 422 amino acid ORF (Fig. 3) that could encode a protein with a predicted molecular mass of 48·4 kDa and a pI of 9·5. A polypurine region centred 10 nt upstream of the putative initiation codon was identified (Fig. 3) that may function as a ribosome-binding site (RBS) (Shine & Dalgarno, 1974; Steitz & Jakes, 1975) in *V. cholerae*. Computer-assisted homology searches revealed similarity between the protein encoded by this putative ORF and members of the integrase family of site-specific recombinases. The highest degree of similarity detected, 73% (52% identity), was to the *E. coli* cryptic prophage integrase protein, SlpA (Fig. 4) (Kirby *et al.*, 1994). Based on the level of similarity between the predicted protein and the integrase family of site-specific recombinases, we have designated this *V. cholerae* gene *int* (integrase). Additionally, the DNA sequence upstream of the *V. cholerae* integrase gene showed identity with the proposed *E. coli* CP4-57 *attL* site at nineteen of twenty nucleotides (Fig. 5) and showed a significant degree of similarity to the region of the *E. coli* chromosome reported by Miczak *et al.* (1991) that encodes the 10Sa RNA (*ssrA*), *smpB*, *smpA* and *recN* gene products (Fig. 1) (Chauhan & Apirion, 1989; Rostas *et al.*, 1987; Komine *et al.*, 1994). The chromosomal region immediately upstream of the putative *V. cholerae int* shows 79·8% nucleotide identity with the *E. coli ssrA*, and 89% and 88% amino acid

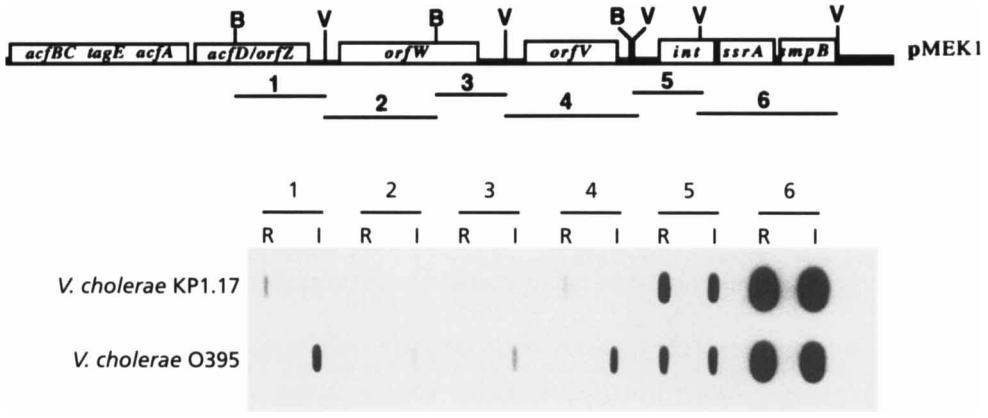

Fig. 2. Slot blot analysis of RNA isolated from *V. cholerae* O395 and KP1.17 under ToxR-regulon-inducing conditions (I) and ToxR-regulon repressing conditions (R). Regions of pMEK1 used as probes are numbered 1–6 and depicted under the physical map. Probes 5 and 6 identify non-ToxR-activated genes. Restriction sites: B, *Bam*HI; V, *Eco*RV.

```
ssrA >
AGATCTAATTCAGGTTAGCCATTCGTTAGCGTGTCGGTTCGCAGGCGGTGGTGAAATTAAAGATCGACTAAGCATGTAGTACCAAAGATGAATGGTTTTC   100
                attL
GGACGGGGGTTCAACTCCCCCCAGCTCCACCAAATCATAGTTTAAAGACGTCTCCGGGCGTCTTTTTTCTTATATAAAACAATTAAAAAATCAATAACTT   200

AGCATTCATAGCTATTCCACAAGATTCCATAGCAGCTTTGATTTTTAGGTACACTACAAGGTACACTGATTTTATCGTGTTAATCTGGTGTACCCATCAT   300
                                                                                                    M     1

GGCAAAAAACACAGTAAAGCTGACAGATAAAGAGATCAAAGCCTCCAAACCTAGAGAGAAAGAATATAATCTTTTCGATGGTGATGGCTTGCGACTTAGA   400
  A  K  N  T  V  K  L  T  D  K  E  I  K  A  S  K  P  R  E  K  E  Y  N  L  F  D  G  D  G  L  R  L  R      34

ATCAAACCTAACGGTTCAAAGCAGTGGATCTTCAACTACTATCGCCCCACAAACGGTAAACGAGCTAACTTAAGCTTTGGAATGTACCCAGACATCTCCC   500
  I  K  P  N  G  S  K  Q  W  I  F  N  Y  Y  R  P  T  N  G  K  R  A  N  L  S  F  G  M  Y  P  D  I  S      67

TTTTACAAGCAAGAAAAAGCTTACTTTCTGCCAAAGAGCTGATCGCACAAGGCATCGATCCACAAGACGATCGAAAACGCCAACAACAAGTACATAAAGA   600
L  L  Q  A  R  K  S  L  L  S  A  K  E  L  I  A  Q  G  I  D  P  Q  D  D  R  K  R  Q  Q  Q  V  H  K  E     101

GATCCATGAGCACACCTTTGAGAACGTCACAAAGGATTGGTTTGCCATCAAACAGAATGATGTGACACCGGATTACGCCCTTGATATCTGGCGTTCATTA   700
  I  H  E  H  T  F  E  N  V  T  K  D  W  F  A  I  K  Q  N  D  V  T  P  D  Y  A  L  D  I  W  R  S  L     134

GAGCTACATATCTTCCCGCACATTTCCGACAAACCGGTAAAAGCGATCACCGCTCCGGAGATCATAGAACTACTCAAACCGATTGAAGCCAAAGGTAGCC   800
  E  L  H  I  F  P  H  I  S  D  K  P  V  K  A  I  T  A  P  E  I  I  E  L  L  K  P  I  E  A  K  G  S     167

TAGAAACCGTCAAACGTTTAACACAGCGACTCAATGAAGTGATGAACTTCGCCACTAACTGCGGTCTTATCCAGGCCAATCCATTAACTGGCATCAAAGC   900
L  E  T  V  K  R  L  T  Q  R  L  N  E  V  M  N  F  A  T  N  C  G  L  I  Q  A  N  P  L  T  G  I  K  A     201

CGCCTTCAAGAAACCAAAAAAAGAAAACATGGCAGCACTGACACCAGCAGAACTGCCCGAACTCATGAGCGCAATCGCTAATGCCAGTATCAAACGCACT  1000
  A  F  K  K  P  K  K  E  N  M  A  A  L  T  P  A  E  L  P  E  L  M  S  A  I  A  N  A  S  I  K  R  T     234

ACCCGTTGTTTACTTGAGTGGCAACTACACACCATGACGCGCCCAGCGGAAGCATCTGGTGCACGTTGGGATGAAATTAATTGGGAAGAAAAGGTTTGGA  1100
  T  R  C  L  L  E  W  Q  L  H  T  M  T  R  P  A  E  A  S  G  A  R  W  D  E  I  N  W  E  E  K  V  W     267

CCATTCCTGCTGAAAGAATGAAAAAAAGAAGAGAACACCGCATCCCGCTCACAGAGCAGATGCTGGCGCTATTGGAAGTAATGAAACCGATAAGTGGCCA  1200
T  I  P  A  E  R  M  K  K  R  R  E  H  R  I  P  L  T  E  Q  M  L  A  L  L  E  V  M  K  P  I  S  G  H     301

TCGTGATTTTATTTTTCCGTCGGATAGAGCCCCGAAAAAGCCCTGCAACAGCCAAACGGCAAATATGGCATTGAAACGTATGGGGTTTGCTGGAAGGCTG  1300
  R  D  F  I  F  P  S  D  R  A  P  K  K  P  C  N  S  Q  T  A  N  M  A  L  K  R  M  G  F  A  G  R  L     334

GTTAGTCACGGTTTGCGCTCTCTAGCGAGTACCACACTCAACGAACAAGGCTTTGATCCAGACTTGGTTGAGTCGGCATTGGCTCACGTGGACGATAACC  1400
  V  S  H  G  L  R  S  L  A  S  T  T  L  N  E  Q  G  F  D  P  D  L  V  E  S  A  L  A  H  V  D  D  N     367

AAGTACGCAGTGCCTATAACCGCACTGACTACTTAGAACGCAGAAAACCAATGATGTGCTGGTGGAGCGAGCATATAGAAGAGGCAGCCAGAGGTAGCCT  1500
Q  V  R  S  A  Y  N  R  T  D  Y  L  E  R  R  K  P  M  M  C  W  W  S  E  H  I  E  E  A  A  R  G  S  L     401

ATCTGTGACTGGTACAAAGCAATTGAAAAATAAACTTCCACCGCCCCCCACATGGAAGCAGATGTGATCCTAACGGCATTCTTTTTGTTCTAAGAACTGG  1600
  S  V  T  G  T  K  Q  L  K  N  K  L  P  P  P  P  T  W  K  Q  M  *                                     422

TTGTCAGTGGAATACGCTTAACGCGGCGAGTATTTTTACCTCTAGCTCTGCACATCGTCGCTTTCAAAAGCGTCGAGATGCAAGTGTATTTGAACTAATT  1700

TGGCAAAATGGAATGATTAACAAAGGTATTGATTGGCGTTCTATATTAATGGATGGCTGCCAAACTAAGGTTCCTTTAGCTAGCTCAAAAAAGGCAAGAA  1800

CCCAACAGACAGAGTAAAACAAGACGTAAAACGCAGTTTGTTGACGGATGAAAATGGATCGCCACGTGCTGTTGTCAGTGACGGAGCCAACGTTCACGAC  1900

ATCAAATTTGTACTCAAAACTTTAGAAGCATTAGAGTGCTGTCGTCTGCTGGTACGTATCCCGATATCTCTTCGGCAGAAGCTCGAGAGAAAGCCTTGAG  2000

GATCGCAAACAGTTAGCGGATATCATGGATCC                                                                      2032
```

Fig. 3. Nucleotide sequence of a *Bgl*II–*Bam*HI DNA fragment containing the partial *V. cholerae* O395 *ssrA* and *int* genes and the deduced integrase amino acid sequence. The underlined segment indicates the putative ribosome-binding site and the double-underlined region indicates the putative *attL* site. The translation stop codon (TGA) is indicated by *.

similarity to *E. coli* SmpB and SmpA, respectively (data not shown). Further analysis has shown that the non-ToxR-activated transcript detected in slot blot analysis (Fig. 2) originates immediately upstream of the integrase and corresponds to the *V. cholerae* equivalent of the 10Sa RNA (*ssrA*).

```
Int   MAKNTVKLTDKEIKASKPREKEYNLFDGDGLRLRIKPNGSKQWIFNYYRPTNGKRANLSF  60
      **||*  *** ****|**|| |*|*|****| * **||*** * * ****   |*|| **
SlpA  MARKTKPLTDTEIKAAKPKDADYQLYDGDGLTLLIKSSGSKLWQFRYYRPLTKQRTKQSF  60

Int   GMYPDISLLQARKSLLSAKELIAQGIDPQDDRKRQQQVHKEIHEHTFENVTKDWFAIKQN 120
      * ** |** |***    |* *|*||****|||* * |  |*   |**  *|  *  |*|
SlpA  GAYPAVSLSDARKLRAESKVLLAKDIDPQEHQKEQVRNSQEAKTNTFLLVAERWWNVKKT 120

Int   DVTPDYALDIWRSLELHIFPHISDKPVKAITAPEIIELLKPIEAKGSLETVKRLTQRLNE 180
       ** *** ******* |*** *|* ||  * *  ||  ||*  * * ****|** **|**
SlpA  SVTEDYADDIWRSLERDIFPAIGDISITEIKAHTLVKAVQPVQARGALETVRRLCQRINE 180

Int   VMNFATNCGLIQANPLTGIKAAFKKPKKENMAALTPAELPELMSAIANASIKRTTRCLLE 240
      ** |* * ***|* *   *  ** **|* **||| * |**|** ||  ***  |****|
SlpA  VMIYAQNTGLIDAVPSVNIGKAFEKPQKKNMPSIRPDQLPQLMHTMRTASISMSTRCLFM 240

          +
Int   WQLHTMTRPAEASGARWDEINWEEKVWTIPAERMKKRREHRIPLTEQMLALLEVMKPISG 300
      *** *|******| ******| |   * *** ***  *|* |**||  **|**|***|**
SlpA  WQLLTITRPAEAAEARWDEIDFNASEWKIPAARMKMNRDHTVPLSDGALAILEMMKPLSG 300

                                          +  +
Int   HRDFIFPSDRAPKKPCNSQTANMALKRMGFAGRLVSHGLRSLASTTLNEQGFDPDLVESA 360
       *|*****   *||* **** * **** *||* ********|***|***|** **||*|*
SlpA  GREFIFPSRIKPNQPMNSQTVNAALKRAGLGGVLVSHGLRSIASTALNEEGFPPDVIEAA 360

             +
Int   LAHVDDNQVRSAYNRTDYLERRKPMMCWWSEHIEEAARGSLSVTGTKQLKNKLPPPPTWK 420
      ***** *|** ****|****|*|*** **|| |  *  **| |*   |||
SlpA  LAHVDKNEVRRAYNRSDYLEQRRPMMQWWADLVKAADSGSIVLTHLSKIR             410

Int   QM                                                            422
```

Fig. 4. Computer-assisted alignment of the *V. cholerae* O395 integrase (Int) predicted amino acid sequence with the *E. coli* CP4-57 cryptic prophage integrase (SlpA) amino acid sequence (Kirby *et al.*, 1994). Numbers on the right of the amino acid residues indicated the residue number. Identical matches (52%) are indicated by * and conservative changes (21%) are indicated by |. The amino acids that are conserved in the P4 family of integrase proteins that are believed to be part of the catalytic domain (Argos *et al.*, 1986; Abremski & Hoess, 1992) are indicated by +.

```
E. coli CP4-57          AACTCCCGCCAGCTCCACCA
                                  |
V. cholerae O395        AACTCCCCCCAGCTCCACCA
V. cholerae 1074-78     AACTCCCCCCAGCTCCACCA
                                  ‡
V. cholerae SG34        AAATCCCCCCAGCTCCACCA
```

Fig. 5. Comparison of the putative *attL* sites from *E. coli* and *V. cholerae* O395, 1074-78 and SG34. The base pair mismatch between *V. cholerae* O395 and *E. coli* is indicated by | and the base pair mismatch between *V. cholerae* O1 (O395 and 1074-78) and non-O1 *V. cholerae* SG34 is indicated by ‡.

Identification of *int* sequences in El Tor and O139 strains of *V. cholerae*

To ascertain if the integrase gene and surrounding sequences were unique to *V. cholerae* classical strain O395, slot blot analysis was performed on 13 additional *V. cholerae* strains. DNA fragments specific for *tagD* (Hughes *et al.*, 1994), *tcpI* (Harkey *et al.*, 1994), *toxT* (Higgins *et al.*, 1992), *acf* (Peterson & Mekalanos, 1988) as well as the *int* and *ssrA* genes were used to determine if these sequences were conserved in recent *V. cholerae* O1 El Tor and non-O1 isolates. Chromosomal DNA from eight O1 (seven El Tor, one classical), three O139 and three non-O1 strains were analysed. Slot blot analysis (Fig. 6) demonstrated that DNA isolated from nine of the *V. cholerae* strains tested generated a hybridization signal when analysed with all six probes. Four of the five *V. cholerae* strains that failed to hybridize with all but the *ssrA*-specific probe are nonpathogenic (Levine *et al.*, 1982; Morris, 1990). Analysis of the *ssrA* region from the nonpathogenic O1, El Tor *V. cholerae* strain 1074-78 revealed that numerous base pair substitutions, insertions and deletions have occurred in the DNA sequence between the *ssrA* and *int* genes (data not shown). The *ssrA* region of the non-O1 *V. cholerae* SG34 chromosome has also been cloned (data not shown) and the nucleotide sequence of the region downstream of the *ssrA* gene determined. A single C to A transversion was noted between the *V. cholerae* O395 and SG34 *att*-like sites (Fig. 5). The nucleotide sequence alignment between O395 and SG34 ends at the last A residue of the proposed *att* site and no *int* sequences were observed (data not shown). In slot blot analysis a probe derived from ϕ19 (Table 1), encompassing the DNA downstream of the SG34 *ssrA* gene, hybridized to chromosomal DNA from

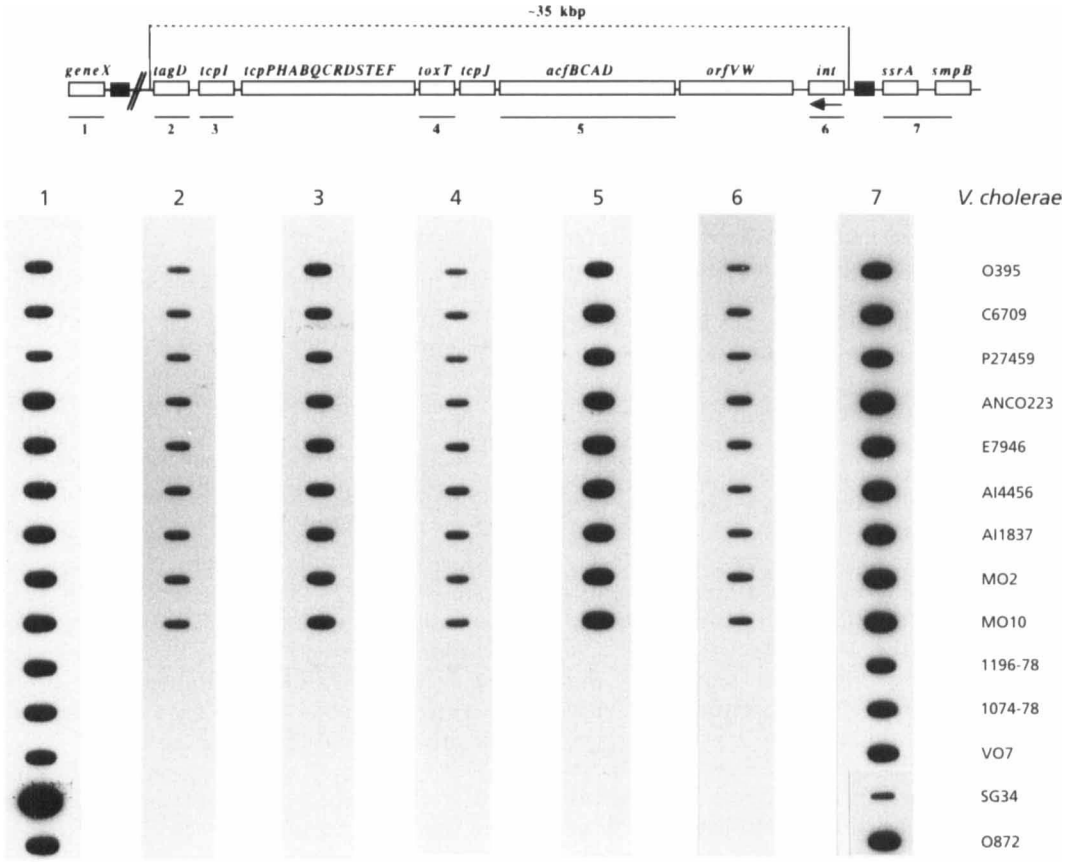

Fig. 6. Physical map depicting the regions of the *V. cholerae* chromosome that were used as probes in the slot blot analysis on chromosomal DNA isolated from 14 *V. cholerae* strains. Probes 2–7 were derived from *V. cholerae* O395 and probe 1 was derived from *V. cholerae* SG34. The black boxes indicate the *att*-like sequences. The arrow indicates the direction of transcription of the integrase gene.

all 14 *V. cholerae* strains tested (Fig. 6, probe 1). This probe did not hybridize with a recombinant phage, ϕD12-2 (Table 1) that contains ~ 10 kbp of *V. cholerae* DNA sequences upstream of *tagD* (data not shown). Since there is ~ 35 kbp of DNA sequence between *tagD* and *int*, the absence of hybridization between the SG34 derived probe and ϕD12-2 indicates that the *V. cholerae* chromosomal element encompassing the ToxR-regulated TCP-ACF colonization determinants is greater than 45 kbp. We are currently isolating overlapping recombinant bacteriophage clones in order to identify the proximal end of the TCP-ACF gene cluster.

DISCUSSION

Our efforts to define the distal limit of the *V. cholerae* ToxR-regulated gene cluster consisting of the TCP and ACF colonization determinants have led to the identification of a linked gene that belongs to the integrase family of site-specific recombinases. The *V. cholerae* integrase lies next to a gene (*ssrA*) encoding the *V. cholerae* equivalent of the *E. coli* 10Sa RNA. Based on the constitutive level of expression of *ssrA*, we have defined this region as the end of the ToxR-regulated TCP-ACF gene cluster. Slot blot analysis of DNA isolated from *V. cholerae* O1 and non-O1 strains has revealed that the region of the *V. cholerae* chromosome encoding TCP, ACF and the integrase is present in the Asiatic-cholera-causing strains but absent from the non-cholera vibrio strains tested.

The putative *V. cholerae* integrase shows significant amino acid similarity with the P4 family of integrases (Argos *et al.*, 1986). These catalyse site-specific integration via DNA nicking and strand exchange that is mediated through a conserved catalytic domain within the carboxy-terminus of the integrase protein (Argos *et al.*, 1986; Landry, 1989). The *V. cholerae* integrase contains the four highly conserved catalytic domain residues (Fig. 4) (Argos *et al.*, 1986; Abremski & Hoess, 1992) and shows the highest degree of amino acid similarity (73%) with the *E. coli* CP4-57 cryptic prophage integrase, SlpA. In addition to the extensive amino acid similarity between these two proteins, *V. cholerae* and *E. coli* show nucleotide identity at nineteen of twenty positions within the *attL*-like site upstream of the *V. cholerae* integrase and the CP4-57 *attL* site. Both the *V. cholerae* and *E. coli* CP4-57 integrase

genes are located at very similar chromosomal sites i.e. downstream of the gene (*ssrA*) that encodes the 10Sa small stable RNA, further reinforcing the similarity between these systems and these two proteins.

Several lines of evidence suggest that the *V. cholerae tcp*, *acf* and *int* are part of a genetic element that encompasses > 45 kbp of the *V. cholerae* chromosome. The atypical codon bias reported by Ogierman & Manning (1992) for the TCP genes led them to suggest that the TCP gene cluster was either recently acquired or was non-*V. cholerae* in origin. Our studies indicate that the TCP-ACF gene cluster is located in a region of the *V. cholerae* chromosome that is analogous to the *E. coli* CP4-57 cryptic prophage integration site (Kirby *et al.*, 1994). When chromosomal DNA from 14 strains of *V. cholerae* was examined by slot blot analysis for the presence of *tcp*, *acf* and *int* sequences, an interesting pattern of hybridization was observed. The *V. cholerae* O1 and O139 strains that have been associated with Asiatic cholera hybridized to the *tcp*-, *acf*- and *int*-specific probes whereas the non-O1 and nonpathogenic O1 strains hybridized only to the area of the *V. cholerae* chromosome that flanked the ToxR-regulated TCP and ACF gene clusters. When the *ssrA* region of the chromosome from *V. cholerae* 1074-78, a nonpathogenic O1 strain that did not hybridize to the *tcp*-, *acf*- or *int*-specific probes, was examined, only the 5′ portion of the integrase gene was present in this strain and numerous base pair substitutions, insertions and deletions were noted between the *ssrA* and *int*, as well as within the portion of the *int* gene that was present. The presence of multiple base pair changes within the intervening DNA sequence of the integrase and 10Sa RNA is similar to that which frequently occurs when prophage undergo partial deletion or illegitimate recombination (Campbell, 1992, 1994). Our sequence data indicate that a similar phenomenon may have occurred in *V. cholerae* 1074-78. *V. cholerae* SG34 is a non-cholera non-O1 strain that lacks *tcp*, *acf* and any residual portions of the integrase gene near the *ssrA* region. This strain however, contains an intact *att* site suggesting that SG34 has never acquired the *tcp*, *acf* and *int* genes. These observations, taken together with the fact that the TCP-ACF gene cluster possesses a gene (*toxT*) encoding a transcriptional activator (ToxT) necessary for TCP-ACF gene expression leads us to believe that the TCP-ACF gene cluster was at one time part of a large mobile genetic element.

Karaolis *et al.* (1995) have proposed, on the basis of polymorphisms within the *V. cholerae* housekeeping gene encoding the aspartate-semialdehyde dehydrogenase (*asd*), that the O1 serogroup may arise from non-toxigenic, non-O1 *V. cholerae* via the horizontal transfer of the O antigen. A similar horizontal transfer of the element containing the TCP and ACF colonization determinants may have occurred, possibly via a mechanism that is similar to that shown for the 17 kbp mobile genetic element (SLP1) of *Streptomyces coelicolor*. The SLP1 element can undergo excision, conjugal transfer and subsequent site-specific integration into the chromosome of recipient *S. coelicolor* cells (Omer & Cohen, 1984). We are currently attempting to develop an *in vitro* system to determine if transfer of the *tcp–acf–int* region from the *V. cholerae* chromosome can occur between *V. cholerae* strains in a similar manner. Preliminary mating experiments between *V. cholerae* O395 (*tcp*$^+$ *acf*$^+$ *int*$^+$) and *V. cholerae* VO7 (*tcp acf int*) have not resulted in the transfer of the TCP-ACF gene cluster from pathogenic to nonpathogenic vibrios. These results may not be surprising since we have failed to detect an *int* transcript under a variety of different culture conditions (data not shown). In *E. coli*, a small basic protein (AlpA) has been identified that can act as a positive regulator of *slpA* (integrase) expression. Overexpression of AlpA results in an increase in transcription of *slpA*, which subsequently leads to excision of the CP4-57 prophage from the *E. coli* chromosome (Trempy *et al.*, 1994; Kirby *et al.*, 1994). A similar activator protein may be required for expression of the *V. cholerae* integrase and the subsequent excision or transfer of this element from the *V. cholerae* chromosome.

Future studies will attempt to overexpress *int* in the donor strain (*tcp*$^+$ *acf*$^+$) to determine if the integrase protein can promote excision of the TCP-ACF gene cluster from the *att*-like site downstream of the *V. cholerae* equivalent of the *ssrA* gene. Additional studies will focus on delineating the proximal end of the TCP-ACF element. Isolation of the intact TCP-ACF element using a modified form of the *in vivo* excision system developed by Pósfai *et al.* (1994), which allows the cloning of large (50–100 kbp) regions of the bacterial chromosome, may permit the size and limits of this element to be clearly defined and help identify the number of genes present within the TCP-ACF element. Isolation of the intact TCP-ACF element will also allow studies to be performed in which these colonization factors are introduced into non-O1, noncholera vibrios that possess *toxR* but lack TCP-ACF sequences such as *V. cholerae* VO7 (Waldor & Mekalanos, 1994). These recombinant *V. cholerae* strains can then be tested for their ability to express TCP and ACF, and colonize the intestines of infant mice. The introduction of the TCP-ACF colonization element into nonreactogenic strains of *V. cholerae* may thus provide an avenue for generating efficacious live vaccine candidates.

ACKNOWLEDGEMENTS

We thank D. T. Beattie, M. Waldor and J. B. Kaper for the kind gifts of *V. cholerae* strains, C. S. Harwood for pHRP309, S. Gottesman for pJK5 and pJT27, D. Friedman for sharing results prior to publication, and V. J. DiRita and P. Zuber for helpful comments. This work was supported by Public Health services grant AI28502 from the National Institutes of Health and the LSUMC Center for Excellence in Cancer Research, Treatment and Education.

REFERENCES

Abremski, K. E. & Hoess, R. H. (1992). Evidence for a second conserved arginine residue in the integrase family of recombination proteins. *Prot Eng* **5**, 87–91.

Aiba, H., Adhya, S. & deCrombrugghe, B. (1981). Evidence for two functional *gal* promoters in intact *Escherichia coli* cells. *J Biol Chem* **256**, 11905–11910.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

Argos, P., Landy, A., Abremski, K., Egan, J. B., Haggard-Ljungquist, E., Hoess, R. H., Kahn, M. L., Kalionis, B., Narayana, S. V. L., Pierson, L. S., III, Sternberg, N. & Leong, J. M. (1986). The integrase family of site-specific recombinases: regional similarities and global diversity. *EMBO J* **5**, 433–440.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K. (1989). *Current Protocols in Molecular Biology*. New York: John Wiley.

Baudry, B., Fasano, A., Ketley, J. & Kaper, J. B. (1992). Cloning of a gene (*zot*) encoding a new toxin produced by *Vibrio cholerae*. *Infect Immun* **60**, 428–434.

Campbell, A. M. (1992). Chromosomal insertion sites for phages and plasmids. *J Bacteriol* **174**, 7495–7499.

Campbell, A. (1994). Comparative molecular biology of lambdoid phages. *Annu Rev Microbiol* **48**, 193–222.

Chauhan, A. K. & Apirion, D. (1989). The gene for a small stable RNA (10Sa RNA) of *Escherichia coli*. *Mol Microbiol* **3**, 1481–1485.

DiRita, V. J. (1992). Co-ordinate expression of virulence genes by ToxR in *Vibrio cholerae*. *Mol Microbiol* **6**, 451–458.

DiRita, V. J., Parsot, C., Jander, G. & Mekalanos, J. J. (1991). Regulatory cascade controls virulence in *Vibrio cholerae*. *Proc Natl Acad Sci USA* **88**, 5403–5407.

Everiss, K. D., Hughes, K. J. & Peterson, K. M. (1994a). The accessory colonization factor and toxin-coregulated pilus gene clusters are physically linked on the *Vibrio cholerae* O395 chromosome. *DNA Seq* **5**, 51–55.

Everiss, K. D., Hughes, K. J., Kovach, M. E. & Peterson, K. M. (1994b). The *Vibrio cholerae acfB* colonization determinant encodes an inner membrane protein that is related to a family of signal-transducing proteins. *Infect Immun* **62**, 3289–3298.

Fasano, A., Baudry, B., Pumplin, D. W., Wasserman, S. S., Tall, B. D., Ketley, J. M. & Kaper, J. B. (1991) *Vibrio cholerae* produces a second enterotoxin which affects intestinal tight junctions. *Proc Natl Acad Sci USA* **88**, 5242–5246.

Finlay, B. B. & Falkow, S. (1989). Common themes in microbial pathogenicity. *Microbiol Rev* **53**, 210–230.

Harkey, C. W., Everiss, K. D. & Peterson, K. M. (1994). The *Vibrio cholerae* toxin-coregulated-pilus gene *tcpI* encodes a homolog of methyl-accepting chemotaxis proteins. *Infect Immun* **62**, 2669–2678.

Higgins, D. E., Nazareno, E. & DiRita, V. J. (1992). The virulence gene activator ToxT from *Vibrio cholerae* is a member of the AraC family of transcriptional activators. *J Bacteriol* **174**, 6974–6980.

Hughes, K. J., Everiss, K. D., Harkey, C. W. & Peterson, K. M. (1994). Identification of a Vibrio *cholerae* ToxR-activated gene (*tagD*) that is physically linked to the toxin-coregulated pilus (*tcp*) gene cluster. *Gene* **148**, 97–100.

Karaolis, D. K. R., Lan, R. & Reeves, P. R. (1995). The sixth and seventh cholera pandemics are due to independent clones separately derived from environmental, nontoxigenic, non-O1 *Vibrio cholerae*. *J Bacteriol* **177**, 3191–3198.

Kaufman, M. R., Shaw, C. E., Jones, I. D. & Taylor, R. K. (1993). Biogenesis and regulation of the *Vibrio cholerae* toxin-coregulated pilus: analogies to other virulence factor secretory systems. *Gene* **126**, 43–49.

Kirby, J. E., Trempy, J. E. & Gottesman, S. (1994). Excision of a P4-like cryptic prophage leads to Alp protease expression in *Escherichia coli*. *J Bacteriol* **176**, 2068–2081.

Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K. & Inokuchi, H. (1994). A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. *Proc Natl Acad Sci USA* **91**, 9223–9227.

Landry, A. (1989). Dynamic, structural and regulatory aspects of λsite-specific recombination. *Annu Rev Biochem* **58**, 913–949.

Levine, M. M., Black, R. E., Clements, M. L., Cisneros, L., Saah, A., Nalin, D. R., Gill, D. M., Craig, J. P., Young, C. R. & Ristaino, P. (1982). The pathogenicity of nonenterotoxigenic *Vibrio cholerae* serogroup O1 biotype El Tor isolated from sewage water in Brazil. *J Infect Dis* **145**, 296–299.

Miczak, A., Chauhan, A. K. & Apirion, D. (1991). Two new genes located between 2758 and 2761 kilobase pairs on the *Escherichia coli* genome. *J Bacteriol* **173**, 3271- 3272.

Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Miller, V. L. & Mekalanos, J. J. (1984). Synthesis of cholera toxin is positively regulated at the transcriptional level by *toxR*. *Proc Natl Acad Sci USA* **81**, 3471–3475.

Miller, V. L. & Mekalanos, J. J. (1988). A novel vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in *Vibrio cholerae* requires *toxR*. *J Bacteriol* **170**, 2575–2583.

Morris, J. G. (1990). Non-O group 1 *Vibrio cholerae*: a look at the epidemiology of an occasional pathogen. *Epidemiol Rev* **12**, 179–191.

Ogierman, M. A. & Manning, P. A. (1992). Homology of TcpN, a putative regulatory protein of *Vibrio cholerae*, to the AraC family of transcriptional activators. *Gene* **116**, 93–97.

Omer, C. A. & Cohen, S. N. (1984). Plasmid formation in *Streptomyces*: excision and integration of the SLP1 replicon at a specific chromosomal site. *Mol Gen Genet* **196**, 429–438.

Parsot, C. & Mekalanos, J. J. (1991). Expression of the *Vibrio cholerae* gene encoding aldehyde dehydrogenase is under control of ToxR, the cholera toxin transcriptional activator. *J Bacteriol* **173**, 2842–2851.

Parsot, C. & Mekalanos, J. J. (1992). Structural analysis of the *acfA* and *acfD* genes of *Vibrio cholerae*: effects of DNA topology and transcriptional activators on expression. *J Bacteriol* **174**, 5211–5218.

Pearson, G. D. N., Woods, A., Chiang, S. L. & Mekalanos, J. J. (1993). CTX genetic element encodes a site-specific recombination system and an intestinal colonization factor. *Proc Natl Acad Sci USA* **90**, 3750–3754.

Peterson, K. M. & Mekalanos, J. J. (1988). Characterization of the *Vibrio cholerae* ToxR regulon: identification of novel genes involved in intestinal colonization. *Infect Immun* **56**, 2822–2829.

Pósfai, G., Koob, M., Hradecná, Noaman, H., Filutowicz, M. & Szybalski, W. (1994). *In vivo* excision and amplification of large segments of the *Escherichia coli* genome. *Nucleic Acids Res* **22**, 2392–2398.

Rostas, K., Morton, S. J., Picksley, S. M. & Lloyd, R. G. (1987). Nucleotide sequence and LexA regulation of the *Escherichia coli recN* gene. *Nucleic Acids Res* **15**, 5041–5049.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.

Shine, J. & Dalgarno, L. (1974). The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci USA* **71**, 1342–1346.

Steitz, J. A. & Jakes, K. (1975). How ribosomes select initiator

regions in mRNA: base pair formation between the 3′ terminus of 16S rRNA and the mRNA during initiation of protein synthesis in *Escherichia coli*. *Proc Natl Acad Sci USA* **72**, 4734–4738.

Taylor, R. K., Miller, V. L., Furlong, D. B. & Mekalanos, J. J. (1987). Use of *phoA* gene fusions to identify a pilus colonization factor coordinately regulated with cholera toxin. *Proc Natl Acad Sci USA* **84**, 2833–2837.

Trempy, J. E., Kirby, J. E. & Gottesman, S. (1994). Alp suppression of Lon: dependence on the *slpA* gene. *J Bacteriol* **176**, 2061–2067.

Trucksis, M., Galen, J. E., Michalski, J., Fasano, A. & Kaper, J. B. (1993). Accessory cholera enterotoxin (Ace), the third toxin of a *Vibrio cholerae* virulence cassette. *Proc Natl Acad Sci USA* **90**, 5267–5271.

Waldor, M. K. & Mekalanos, J. J. (1994). ToxR regulates virulence gene expression in non-O1 strains of *Vibrio cholerae* that cause epidemic cholera. *Infect Immun* **62**, 72–78.

Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.

Received 16 October 1995; revised 20 February 1996; accepted 13 March 1996.