JOURNAL OF BACTERIOLOGY, Feb. 1998, p. 901–908 Vol. 180, No. 4
0021-9193/98/$04.00+0

Ordered Cloned DNA Map of the Genome of *Vibrio cholerae* 569B and Localization of Genetic Markers

SOMA CHATTERJEE, ASIM K. MONDAL,† NASIM A. BEGUM, SUSANTA ROYCHOUDHURY,* AND JYOTIRMOY DAS

Biophysics Division, Indian Institute of Chemical Biology, Calcutta 700 032, India

Received 25 July 1997/Accepted 6 December 1997

By using a low-resolution macrorestriction map as the foundation (R. Majumder et al., J. Bacteriol. 176:1105–1112, 1996), an ordered cloned DNA map of the 3.2-Mb chromosome of the hypertoxinogenic strain 569B of *Vibrio cholerae* has been constructed. A cosmid library the size of about 4,000 clones containing more than 120 Mb of *V. cholerae* genomic DNA (40-genome equivalent) was generated. By combining landmark analysis and chromosome walking, the cosmid clones were assembled into 13 contigs covering about 90% of the *V. cholerae* genome. A total of 92 cosmid clones were assigned to the genome and to regions defined by *Not*I, *Sfi*I, and *Ceu*I macrorestriction maps. Twenty-seven cloned genes, 9 *rrn* operons, and 10 copies of a repetitive DNA sequence (IS*1004*) have been positioned on the ordered cloned DNA map.

Vibrio cholerae, a noninvasive gram-negative bacterium and the causative agent of the diarrheal disease cholera, is serologically classified as belonging to the O antigenic group. Strains belonging to O group 1 (O1) are responsible for cholera. Strains other than O1 are called non-O1; they can cause only sporadic infections and do not have the potential to cause epidemics (31). Strains of serovar O1 consist of two biotypes, classical and El Tor. Only recently, an outbreak of cholera in India and Bangladesh which subsequently spread into several parts of the subcontinent was caused by a novel non-O1 strain, O139 Bengal (36). However, several pieces of evidence suggested that strain O139 Bengal closely resembles biotype El Tor of the serovar O1 (5, 43).

Construction of genetic maps is restricted to organisms for which genetic tools are available and experimental genetic transfers are feasible. Although a great deal is known about the biochemistry, physiology, and clinical microbiology of *V. cholerae* (23), the genetic analysis of this organism has been hindered, primarily because of the lack of demonstrable genetic exchange systems. There is no transducing phage of *V. cholerae*, and transformation of these cells by plasmid DNA only has been demonstrated (34). Conjugation is mediated by a factor, P (6), which unlike the F factor of *Escherichia coli* cannot integrate into the chromosome and hence cannot induce Hfr donors. Thus, the mobilization of chromosomal DNA is limited in this organism. The alternative to examining the organization of genomes in organisms for which a genetic map is not available is to construct a physical map which will allow the examination of the phylogenetic relationship between organisms and the variations of genome structure between different serovars and biotypes. Even for organisms with well-defined genetic maps, physical methods can provide additional details like the orientation of genes, rearrangements within a genome, acquisition of DNA from other organisms, and mapping of any sequence which can be used as a probe. A combined genetic and physical map of the 3.2-Mb genome of the classical O1 hypertoxigenic strain 569B (38) has recently been constructed by using the enzymes *Not*I (29), *Ceu*I (32), and *Sfi*I (unpublished observation). The availability of the macrorestriction map enabled examination of the organization of the genomes of *V. cholerae* strains belonging to different serovars and biotypes. One of the unique observations was intraspecies variation in the number of *rrn* operons in vibrios. Strains belonging to serovars O1 and O139 have 9 *rrn* operons, and those belonging to non-O1/non-O139 have 10 *rrn* operons (32). Genomes of *V. cholerae* strains belonging to different serovars and biovars, and particularly those of the pathogenic strains, are undergoing rapid rearrangements and exhibit extensive restriction fragment length polymorphism in the CTX genetic element locus (5). While the linkage maps are conserved within biovars, they vary substantially between biovars (32).

The macrorestriction maps are of relatively low resolution and permit detection of gross chromosomal aberrations, and they allow qualitative evaluation of intraspecies genetic variations and identification of individual isolates of a species by comparison of their macrorestriction patterns. The ordered cloned DNA map of the genome generated from a set of overlapping phage or cosmid clones that cover the whole genome, on the other hand, has much greater potential as a tool to study genome structure and reshuffling of genes (14, 20). The phage or cosmid libraries provide a readily renewable source of DNA, which is important particularly for pathogenic microbes like *V. cholerae*. The ordered cloned DNA map also provides direct access to a given chromosomal locus, permitting surrogate genetics (14) to be conducted, leading to the identification of virulence determinant genes and protective antigens. The ordered cloned DNA library can be used to examine the modulation of transcription of sets of genes that are specifically expressed following exposure to environmental fluctuations (13, 41). A functional description of the bacterial genome can be extended to the protein level by cloning the DNA insert from each cosmid clone into a suitable vector from which controlled expression can be achieved (40). Ordered cloned DNA maps have been constructed for the genomes of relatively few organisms, such as *E. coli* (26), *Mycoplasma pneumonia* (44), *Desulfovibrio vulgaris* (15) *Haloferax volcanii*

* Corresponding author. Mailing address: Biophysics Division, Indian Institute of Chemical Biology, 4, Raja S. C. Mullick Rd., Calcutta 700 032, India. Phone: 91-33-473 0350/5197/5368. Fax: 91-33-473 0350/5197/0284. E-mail: biophy@cal.vsnl.net.in.

† Present address: Heritable Disorder Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892.

(11), *Mycobacterium leprae* (16), *Bacillus subtilis* (1), *Helicobacter pylori* (9), *Myxococcus xanthus* (21), and *Rhodobacter capsulatus* (19). The present report describes the construction of an overlapping cloned DNA map of the genome of *V. cholerae* 569B, done by using the low-resolution macrorestriction map as the foundation. Twenty-seven homologous and heterologous genes, 9 *rrn* operons, and 10 copies of a repetitive DNA sequence, IS*1004*, have been positioned on the map.

MATERIALS AND METHODS

Construction of cosmid library. The *V. cholerae* 569B used in this study was obtained from the National Institute of Cholera and Enteric Diseases, Calcutta, India. *V. cholerae* cells were grown in a gyratory shaker at 37°C in nutrient broth (NB) containing 0.1 M NaCl (pH 8.0) and maintained as described previously (12, 28, 37). Genomic DNA was prepared by the method of Wilson (45). Five micrograms of genomic DNA was partially digested with *Mlu*I and size fractionated in 0.9% low-melting-point (GTG) agarose (FMC, Rockland, Maine). DNA from the 30- to 45-kb region was eluted from the gel, extracted with phenol-chloroform, and ethanol precipitated. The precipitate was dissolved in TE buffer (10 mM Tris-HCl, 1 mM EDTA [pH 8]) and preserved at a final concentration of 200 ng/ml at 4°C.

The cosmid Lorist M, having the phage λ origin of replication (obtained from R. L. Charlebois, University of Ottawa, Ottawa, Ontario, Canada) was used for the construction of the library. About 5 μg of the cosmid DNA was digested with *Mlu*I, dephosphorylated by using calf intestinal phosphatase (New England Biolabs, Beverly, Mass.), ethanol precipitated, washed with 70% ethanol, and dissolved in 5 μl of TE. Two micrograms of size-fractionated genomic DNA was ligated to 5 μg of vector DNA by using 1 U of T4 DNA ligase (Boehringer Mannheim, Indianapolis, Ind.) in a final volume of 20 μl at 16°C for 16 h. The ligation mixture was diluted to 100 μl with SM buffer (0.58% NaCl, 0.2% $MgSO_4$, 100 mM Tris-HCl, 2% gelatin [pH 7.5]) and packaged with phage λ packaging extract prepared from *E. coli* BHB 2688 and BHB 2690 cells (22). The packaged phage particles were absorbed for 30 min at 37°C to *E. coli* ED8767 cells grown to logarithmic phase in terrific broth (TB) containing 1.2% tryptone, 2.4% yeast extract, and 0.4% glycerol and spread on TB agar plates containing 30 μg of kanamycin sulfate per ml. About 4,000 recombinant clones were picked and grown overnight at 37°C in 96-well microtiter plates containing 200 μl of TB containing kanamycin sulfate. Ninety microliters of 50% glycerol was added, and the mixture was stored at −70°C. Cosmid clones were divided into three batches (A, B, and C) each having about 1,350 clones and were numbered A1 to A1350 for batch A, and so on.

Grouping of cosmid clones. Restriction fragment-specific cosmid clones were identified by hybridizing clones with different *Not*I, *Ceu*I, and *Sfi*I fragments of *V. cholerae* genomic DNA. Batches of 500 cosmid clones were grown on Hybond nylon membrane (Amersham, Amersham, England), and colony blot hybridizations were performed by using restriction fragments, labelled by random priming (18), as probes. Hybridization was carried out at 60°C for 12 h, and filters were washed at the desired stringency, dried, and autoradiographed.

Landmark analysis and chromosome walking. The enzymes *Bam*HI, *Sal*I, *Stu*I, and *Nco*I, having on average one site per 50 kb of *V. cholerae* genomic DNA, were chosen as rare-cutter enzymes for landmark analysis. Ten microliters of DNA digested with 2.5 U of *Mlu*I and 2.5 U of one of the rare-cutter enzymes in a 20-μl volume at 37°C for 4 to 5 h was loaded on a 45-well 0.9% agarose gel (23 by 25 cm) and electrophoresed at 4°C for 18 h. The overlapping clones were identified manually by analyzing the restriction digestion profiles of cosmid clones. For chromosome walking, RNA probes of the terminal clones of the desired contig were prepared from T7 and SP6 promoters by using a Promega kit (Promega Corp., Southampton, United Kingdom) and hybridized with DNA by dot blotting or colony blotting to obtain candidate extenders.

RESULTS

Construction of cosmid library. A cosmid library of the genome of the hypertoxinogenic strain 569B of *V. cholerae* was constructed by cloning genomic DNA partially digested with the enzyme *Mlu*I into the cosmid vector Lorist M. Among the enzymes tested to generate genomic DNA fragments, *Mlu*I was chosen as the cloning enzyme because it did not produce fragments larger than 25 kb. The optimal conditions for partial digestion of the genomic DNA with *Mlu*I were established by digesting DNA with various amounts of enzyme and for different times to generate DNA fragments between 30 and 45 kb. The size of the library was about 4,000 clones carrying inserts of >35 kb, which contained more than 120 Mb of *V. cholerae* DNA (40-genome equivalent).

TABLE 1. Grouping of clones

Fragment name	Size (kb)	No. of colonies screened	No. of colonies hybridized
N1	364	2,400	140
N2	324	1,000	240
N4	189	500	80
N7	166	500	80
N12, N13	112, 106	500	60
S2	296	500	60
C3	325	500	45
C4	275	500	65
C5	180	500	60
C6	120	400	40
C7	78	300	35
C8	72	300	40
N8, S8	150, 175	300	120

Grouping of cosmids into subsets. By taking advantage of the macrorestriction maps of the *V. cholerae* 569B genome, the clones of the cosmid library were grouped into subsets. Batches of about 500 clones from the library were transferred onto nylon filters and hybridized with labelled *Not*I, *Sfi*I, or *Ceu*I fragments of *V. cholerae* genome separated by pulsed-field gel electrophoresis (PFGE). The fragments that are clearly resolved in PFGE and can be eluted from the gel without contamination by adjacent fragments were used for grouping the clones (Table 1). The number of clones belonging to any particular restriction fragment was sufficient to cover at least five times the size of the fragment. Of 37 *Not*I (29) and 9 *Ceu*I (32) fragments of the *V. cholerae* genome, the *Not*I fragments N1, N2, N4, N7, N8, N12, and N13, covering about 43% of the genome, and the *Ceu*I fragments C3 to C8, covering another 42% of the genome, were used for grouping the clones. Another 8% of the genome was covered by *Sfi*I fragments S2 and S8. The ambiguities arising from clones hybridizing with more than one restriction fragment due to the presence of internal repeat sequences were resolved by hybridizing *Not*I-digested genomic DNA with riboprobes prepared from the ends of inserts of these cosmid clones. Altogether, 1,065 of 4,000 cosmid clones were used in subsequent analysis. In each group, identical clones were eliminated by digestion with three restriction enzymes and one representative clone was used for further studies. This allowed the reduction of the number of clones for contig assembly to 665.

Contig assembly. To generate contigs, overlapping cosmid clones were identified primarily by landmark analysis (10). This involves comparison of gel patterns of different clones digested with the cloning enzyme and the double digest of the cloning enzyme and a rare-cutting enzyme. Restriction enzymes having on average one site per 5 kb in the genome are normally used as the cloning enzymes so that the complete digestion of the cloned DNA yields about six to eight fragments. The second enzyme selected for landmark analysis should have on average one site per 50 kb. Thus, among the several fragments produced following complete digestion of the cloned DNA by the cloning enzyme, at least one will have a site for the second enzyme. This fragment will disappear following digestion with the second enzyme, producing new fragments. If two cosmid clones are overlapping, the common bands produced on complete digestion with the cloning enzyme will disappear upon digestion with the second enzyme and reappear as equal-sized fragments in both the clones.

In the present study, *Mlu*I was chosen as the cloning enzyme and *Bam*HI, *Sal*I, *Stu*I, and *Nco*I were chosen as rare-cutting

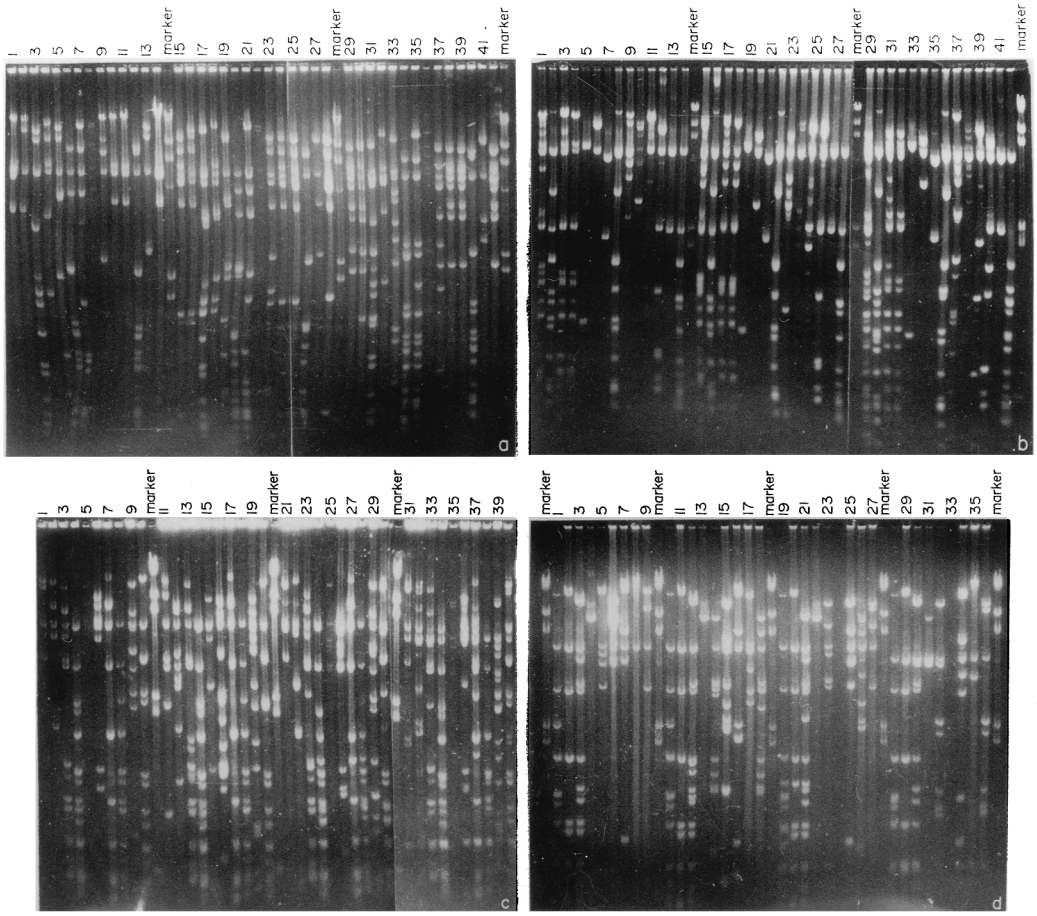

FIG. 1. Landmark analysis of cosmids for identifying overlapping clones. (a and b) Digestion patterns of cosmid clones with *Mlu*I (lanes 1 to 14), *Mlu*I plus *Bam*HI (lanes 15 to 28), and *Mlu*I plus *Sal*I (lanes 29 to 42). (c and d) Digestion patterns of cosmid clones with *Mlu*I (lanes 1 to 10), *Mlu*I plus *Bam*HI (lanes 11 to 20), *Mlu*I plus *Sal*I (lanes 21 to 30), and *Mlu*I plus *Stu*I (lanes 31 to 40).

enzymes. Cosmid clones from different groups, selected randomly, were subjected to landmark analysis to generate contigs. About 50 cosmid clones from any particular group were digested with *Mlu*I and with *Mlu*I and one of the rare-cutting enzymes, and fragments were separated in agarose gels (Fig. 1). Any two clones having at least one *Mlu*I fragment in common which disappears following digestion with any of the second enzymes are overlapping clones, and the disappearing common fragment is the landmark and is a measure of the extent of the overlap. For example, the clones A42 and A90 have a 15-kb common fragment following *Mlu*I digestion (Fig. 2A). When these clones were digested with *Mlu*I and *Sal*I, the common 15-kb fragment was cleaved, producing four fragments of 9.3, 2.4, 2.1, and 1.2 kb (Fig. 2A). Thus, the 15-kb fragment is a landmark and the clones A42 and A90 have an overlap of 15 kb. Similarly, a comparison of the *Mlu*I, *Mlu*I-plus-*Bam*HI and *Mlu*I, *Mlu*I-plus-*Sal*I digestion profiles of the clones A14 and A42 (Fig. 2A) showed that these two clones have a 4-kb overlap. Cosmid clones A90 and A104 (Fig. 2A) have two landmarks of 9 and 1 kb and hence have a 10-kb overlap. Thus, from the landmark analysis of four clones, a contig of A14, A42, A90, and A104 was assembled (Fig. 2B). More than 80% of the overlaps were determined by using the landmark strategy alone.

For some clones, the common *Mlu*I fragment(s) did not disappear following digestion with any of the four rare-cutting enzymes used and thereby did not allow the identification of the landmarks. To overcome this problem, one option is to use more rare-cutting enzymes, which is labor intensive. The other option, which was adopted in the present study, is chromosome walking with riboprobes generated from the two ends of the clone to determine overlapping clones. This approach was used for clones with one *Mlu*I common fragment. Clones having multiple *Mlu*I common fragments were directly taken as overlapping clones, since it is unlikely that two nonoverlapping clones will generate multiple similar-sized fragments. In cases where all the expected reappearing fragments of the landmark following digestion with the second enzyme could not be detected in the gel, the disappearance of the common *Mlu*I fragment(s) was taken as evidence that two clones were overlapping.

Map integration. To generate a relational map, the assembled contigs were positioned on the macrorestriction map (29). This involved the following steps. (i) Cosmid clones containing *Not*I site(s) were identified. *V. cholerae* 569B genomic DNA was digested with *Not*I, end labelled, and subsequently digested with *Hin*dIII to generate probes specific for ends of a *Not*I fragment. All the assembled cosmid clones were hybridized with these probes, and the clones that lit up were digested with *Not*I to confirm the presence of a *Not*I site (Fig. 3A). (ii) Contigs were positioned in the *Not*I map. To position the contig with respect to the junction between two *Not*I frag-

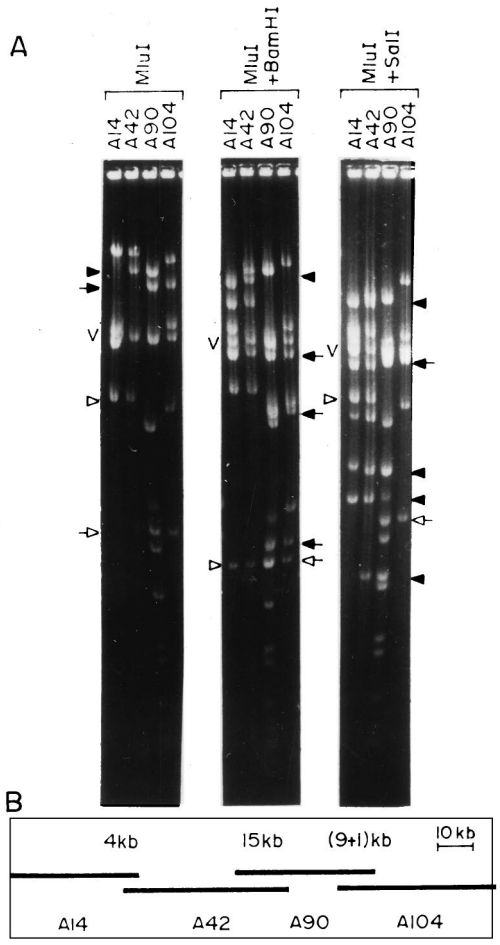

FIG. 2. Identification of overlapping clones and contig assembly by landmark analysis. (A) *Mlu*I, *Mlu*I-plus-*Bam*HI, and *Mlu*I-plus-*Sal*I digestion patterns of four cosmid clones. The closed arrowhead in the *Mlu*I digest represents fragments common to cosmid clones A42 and A90, which is not cleaved by *Bam*HI but produced four fragments (closed arrowheads) following *Sal*I digestion. The closed and open arrows represent two fragments common to cosmid clones A90 and A104. In the *Mlu*I-*Bam*HI double digest, both the fragments disappeared and identical new fragments appeared (closed and open arrows). In the *Mlu*I-*Sal*I double digest, only the fragment identified by the closed arrow disappeared and identical new fragments appeared (closed arrow). The open arrowhead represents a fragment common to A42 and A14 in the *Mlu*I digest which disappeared following *Bam*HI digestion, producing identical new fragments (open arrowhead). The *Mlu*I fragment common to A42 and A14 did not disappear upon digestion with *Mlu*I plus *Sal*I (open arrow). V, vector DNA. (B) Assembled contig comprising four overlapping cosmids, A14, A42, A90, and A104. The extent of overlap between the clones is marked above each overlap.

ments, clones having a *Not*I site(s) in the contig were used as probes in Southern blot hybridization of *Not*I-digested *V. cholerae* 569B genomic DNA. For example, the clone A1044 hybridized with *Not*I fragments N3 and N11 (Fig. 3B), which are linked. Similarly, the clone A606 hybridized with N20 and N23 (Fig. 3B). The clone A793 hybridized with three fragments, N19, N28, and N20 (Fig. 3B), which are linked (29). Whenever required, the positions of the contigs on the macrorestriction map were confirmed by hybridizing *Not*I site-containing cosmids with *Ceu*I-digested *V. cholerae* 569B genomic DNA. *Ceu*I has nine sites in the genome, and all the sites are located in the *rrn* operons (27). The clone A1044, having one *Not*I site and one *Ceu*I site, strongly hybridized with the *Ceu*I fragments C6 and C5 (Fig. 3B), which span the junction of N3 and N11 (29). Because of the presence of an *rrn* operon in the clone, all the other *Ceu*I fragments also hybridized with it, though relatively weakly. The clone A793, having no *Ceu*I site, hybridized only with *Ceu*I fragment C2, which spans N19-N28-N20 of the *Not*I map (Fig. 3B). The positioning of the contigs in the combined *Not*I-*Ceu*I (Fig. 4) map was further confirmed by identifying the cosmid clones with *Sfi*I sites in conformity with the combined *Sfi*I-*Not*I-*Ceu*I macrorestriction map.

Closing of gaps in the map. To close or reduce the gaps between the contigs generated by landmark analysis, chromosome walking was performed. Riboprobes generated by using T7 or SP6 promoters of the cosmid Lorist M from the ends of the terminal clones of each contig were hybridized to clones belonging to a particular group. Chromosome walking allowed identification of about 20% of the overlaps in the contig assembled. While chromosome walking allowed identification of overlapping clones, it could not provide information about the

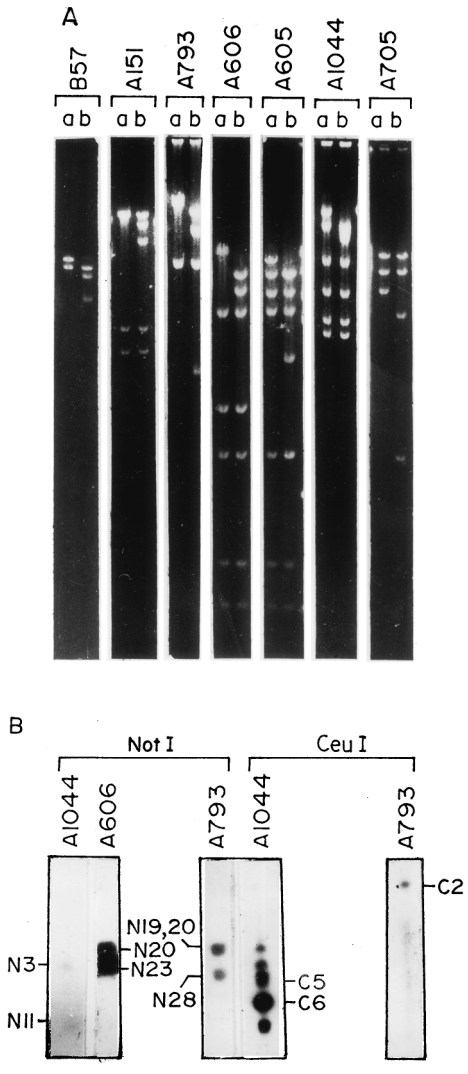

FIG. 3. (A) Identification of *Not*I linking clones. Cosmid clones hybridizing with probes generated from the ends of *Not*I-digested *V. cholerae* genomic DNA were digested with *Bam*HI (lanes a) and *Bam*HI and *Not*I (lanes b). (B) Southern blot hybridization of PFGE-separated *Not*I- and *Ceu*I-digested *V. cholerae* 569B genomic DNA with *Not*I linking cosmid clones A1044, A606, and A793 as probes. The linked *Not*I and *Ceu*I fragments are marked. The clone A793, having no *Ceu*I site, hybridized only with *Ceu*I fragment C2.

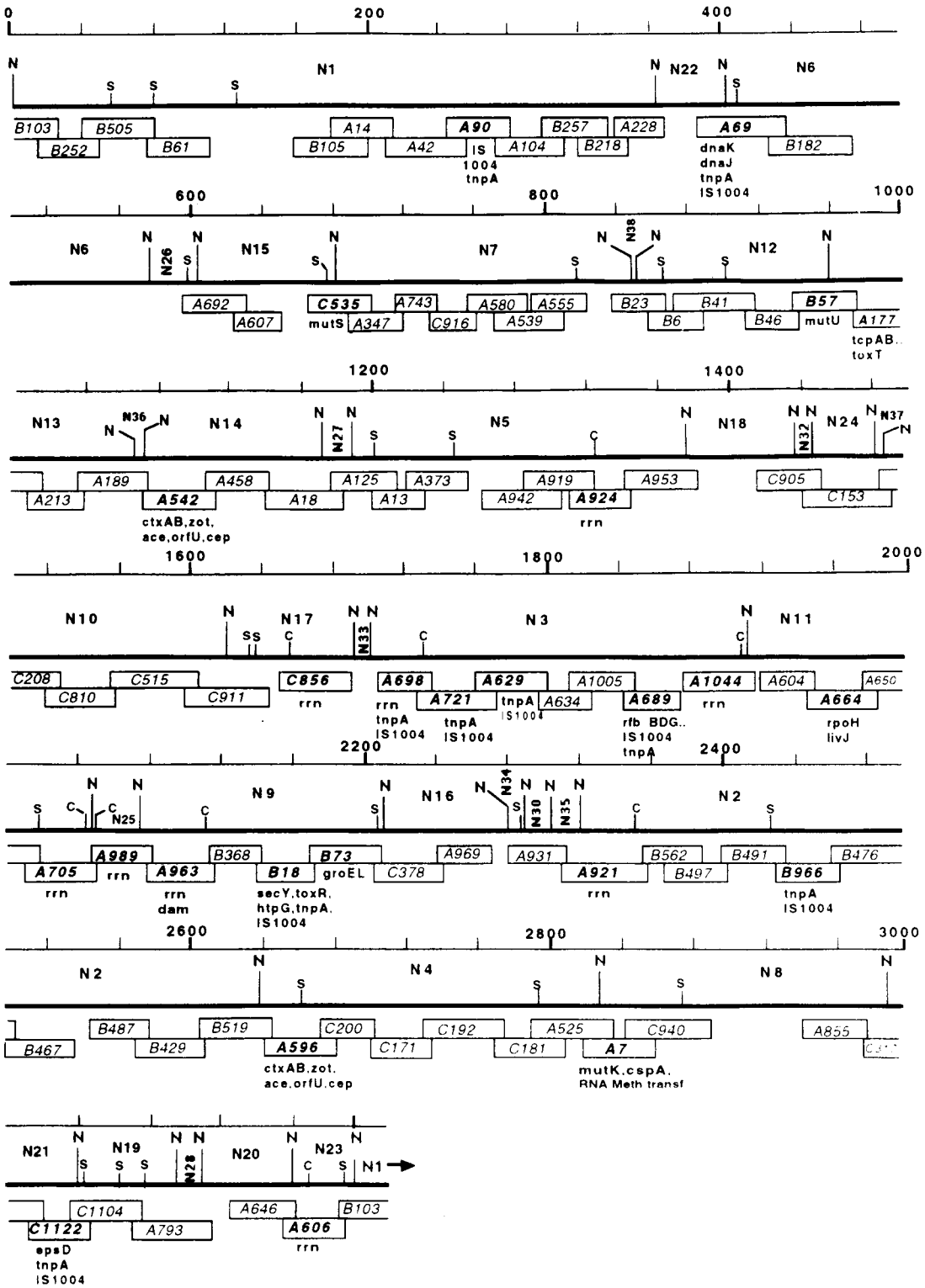

FIG. 4. Linearized ordered cloned DNA map of the 3.2-Mb circular chromosome of *V. cholerae* 569B and positioning of genetic markers on the map. The thick lines represent a composite macrorestriction map consisting of the *Not*I (N), *Sfi*I (S), and *Ceu*I (C) restriction sites. The linkages between different *Not*I fragments were taken from the published physical map (29). The rightmost end of each thick line is contiguous with the leftmost end of the following line. Since the genome is circular, *Not*I sites in the far upper left and far lower right are the same. Each cosmid is represented by an open rectangular box with an identification number in the center. The lengths of the boxes reflect their sizes in kilobases and also the extents of overlap between any two overlapping cosmids. Positions of the genetic markers are shown below the cosmids they belong to. The thin line represents the scale in kilobases, where the first *Not*I site is taken as zero.

TABLE 2. Positioning of cloned genes in the contigs of the ordered cloned DNA map

Gene(s)	Gene product(s) or function	Source	Cosmid(s)	Reference
mutS	DNA mismatch repair	*V. cholerae*	C535	4
mutL	DNA mismatch repair	*V. cholerae*	B57	4
mutK	DNA mismatch repair	*V. cholerae*	A7	Unpublished data
dam	Adenine methyltransferase	*V. cholerae*	A963	2
cspA	Cold shock protein	*V. cholerae*	A7	Unpublished data
RNA methyltransferase gene	RNA methyltransferase	*V. cholerae*	A7	Unpublished data
rpoH	σ^{32}	*V. cholerae*	A664	39
udhA	Unknown dehydrogenase	*V. cholerae*	A664	39
groEL	Hsp60	*E. coli*	B73	17
dnaK	Hsp70	*V. cholerae*	A69	Unpublished data
grpE	DNA synthesis	*V. cholerae*	A69	Unpublished data
dnaJ	Hsp40	*V. cholerae*	A69	Unpublished data
L15 L36	Ribosomal large-subunit proteins	*V. cholerae*	B18	7
secY	Inner membrane protein translocator	*V. cholerae*	B18	7
epsD	Protein secretion	*V. cholerae*	C1122	Unpublished data
ctxAB	Cholera toxin	*V. cholerae*	A542, A596	24
zot	Zonula occludens toxin	*V. cholerae*	A542, A596	3
ace	Accessory cholera enterotoxin	*V. cholerae*	A542, A596	42
cep	Core-encoded pilus	*V. cholerae*	A542, A596	35
orfU	Unknown open reading frame	*V. cholerae*	A542, A596	42
toxR	Virulence gene activator	*V. cholerae*	B18	43
htpG	Stress response protein	*V. cholerae*	B18	43
tcpAB	Toxin-coregulated pilus	*V. cholerae*	A177	25
toxT	Transcriptional activator	*V. cholerae*	A177	25
tnpA (IS*1004*)	Transposase	*V. cholerae*	A90, A69, A458, A698, A721, A629, B18, B966, A855, C1122	8
rfbBDEG	O antigen	*V. cholerae*	A689	30

extent of the overlap. The overlapping clones identified by chromosome walking were thus subjected to landmark analysis to estimate the length of overlap. By combining landmark analysis and chromosome walking, 92 cosmid clones in 13 contigs covering about 90% of the *V. cholerae* genome have been positioned in the overlapping cloned DNA map (Fig. 4). One 120-kb gap and 14 small gaps (ranging from 10 to 50 kb) are yet to be filled.

Positioning of *V. cholerae* genes on the cloned DNA map. Twenty-seven cloned genes and 10 copies of one IS element have been positioned on the ordered cloned DNA map of the *V. cholerae* 569B genome (Fig. 4) by hybridization using homologous and heterologous genes as probes (Table 2). The gene probes used comprised virulence determinant genes, DNA mismatch repair genes, stress response genes, and genes involved in protein translocation. The genes were positioned on the macrorestriction map (29) rather arbitrarily on fragments to which they hybridized, not reflecting their true order in the genome. It will be possible to determine the order of genes in the chromosome and the approximate distances between them from the ordered cloned DNA map. For example, in the low-resolution macrorestriction map, *tcp* and one of the *ctx* genetic elements were positioned in *Not*I fragment N14 (29). The high-resolution map showed that the *tcp* and *ctx* genes are located in two cosmids, A177 and A542, respectively, falling within *Not*I fragments N13 and N14, and that the distance between the two genes is about 50 to 80 kb. The *dam*, *secY*, and *groEL* genes, positioned in *Not*I fragment N9 in the macrorestriction map, are located in the cosmids A963, B18, and B73, respectively, and the order in which these genes are present in the chromosome is *dam-secY-groEL* (Fig. 4). Nine *rrn* operons were positioned in the map on cosmids having *Ceu*I sites. The *Ceu*I sites in the *V. cholerae* genome were taken as the positions of the *rrn* operons.

DISCUSSION

The present report describes the construction of a high-resolution overlapping cloned DNA map of the genome of hypertoxinogenic strain 569B of *V. cholerae*. Thirteen contigs covering 2.85 Mb (about 90% of the whole genome) have been assembled. The availability of the macrorestriction map of the *V. cholerae* genome was extremely useful in grouping the cosmid clones into defined subsets and reducing the number of clones to be analyzed. Besides, the knowledge of *Not*I, *Sfi*I, and *Ceu*I sites in the physical map helped in accurately positioning and orienting contigs containing clones having sites for one of these enzymes.

The success of generating an ordered cloned DNA map depends primarily on the efficiency of detecting overlaps. Several different approaches have been adopted by different investigators to identify overlapping clones. These include (i) restriction mapping of randomly selected clones (26), (ii) fingerprinting (33), (iii) chromosome walking, and (iv) identification of overlapping clones from shared landmarks (10). Each of these approaches has its own limitations, and to construct high-resolution maps of genomes of prokaryotic organisms it is always necessary to combine results obtained from two or more of these approaches. Although the landmark analysis was tested only with one organism, *H. volcanii*, to identify overlapping clones (11), this was preferred over the other strategies, in the present study, for several reasons. This approach allowed detection of small overlaps, and from a relatively small number of clones, an ordered cloned DNA map can be constructed. A minimal set of 92 overlapping clones was sufficient to generate contigs covering 90% of the *V. cholerae* genome by this approach. A total of 72% of the overlaps were less than 10 kb, and the length of none of the overlaps was more than 20 kb. Except in a few cases where chromosome walking was neces-

sary, four rare-cutting enzymes were adequate to identify landmarks. Furthermore, this method does not require extensive use of radioisotopes, which makes it less hazardous.

One of the problems encountered during the construction of the map was instability of cosmid clones. When maintained in *E. coli*, some of the clones were spontaneously deleted. The deletion of some of these clones could be due to the presence of toxic genes. This might be one of the reasons for the presence of the small gaps in the cloned DNA map. The other possibility is that the DNA segments in these regions are not represented in the library. A lambda clone library of *V. cholerae* genomic DNA is under construction, and this will be used to bridge the gaps in the ordered cosmid map and to get complete coverage.

It has been possible to refine and more accurately position genetic loci in the high-resolution map; in the macrorestriction map, in comparison, the genes were arbitrarily positioned on the restriction fragments to which they hybridized. Some more genes in addition to those placed in the macrorestriction map, viz., *grpE*, *dnaJ*, *mutK*, *cspA*, *epsD*, *tnpA*, *rfb*, and genes encoding RNA methyltransferase and ribosomal large-subunit proteins L15 and L36, have been positioned on the ordered cloned DNA map. A 628-bp repeat sequence, IS*1004*, has been reported to be present in the *V. cholerae* genome (8). The present study showed that there are 10 copies of this repeat sequence in the genome of strain 569B of *V. cholerae*, and their locations in the genome have been determined. Several clones other than those containing IS*1004* in the cosmid library hybridized with more than one *Not*I restriction fragment, suggesting the presence of yet-unidentified repeat sequences in those clones. With the addition of more genes, the utility of the map is expanding and its resolution is improving. This will lead to more insight into chromosome organization and help to identify new virulence determinant factors and to understand the molecular basis of pathogenicity of this important human pathogen.

ACKNOWLEDGMENTS

We thank R. L. Charlebois, University of Ottawa, Ottawa, Ontario, Canada, for providing the cosmid vector Lorist M and *E. coli* ED8767 and E. M. Bik, National Institute of Public Health and the Environment, Bithoven, The Netherlands, for providing *tnpA* and *rfbBDEG* genes. We also thank all members of the Biophysics Division, Indian Institute of Chemical Biology, for their kind cooperation and encouragement during this study.

S.C. and N.A.B. are grateful to the Council of Scientific & Industrial Research, New Delhi, India, for a predoctoral fellowship and pool officership, respectively. This work was supported by the Department of Biotechnology (grants BT/TF/15/03/91, BT/MB/05/12/94, and BT/R&D/PRO109/15/8/96) of the Government of India.

REFERENCES

1. **Azevedo, V., E. Alvarez, E. Zumstein, G. Damiani, V. Sgaramella, S. D. Ehrlich, and P. Serror.** 1993. An ordered collection of *Bacillus subtilis* DNA segments cloned in yeast artificial chromosome. Proc. Natl. Acad. Sci. USA **90:**6047–6051.
2. **Bandyopadhyay, R., and J. Das.** 1994. DNA adenine methyl-transferase encoding gene (*dam*) of *Vibrio cholerae*. Gene **140:**67–71.
3. **Baudry, B., A. Fasano, J. Ketley, and J. B. Kaper.** 1992. Cloning of a gene (*zot*) encoding a new toxin produced by *Vibrio cholerae*. Infect. Immun. **60:**428–434.
4. **Bera, T. K., S. K. Ghosh, and J. Das.** 1989. Cloning and characterization of the *mutL* and *mutS* genes of *Vibrio cholerae*: nucleotide sequence of the *mutL* gene. Nucleic Acids Res. **17:**6241–6251.
5. **Bhadra, R. K., S. Roychoudhury, R. K. Banerjee, S. Kar, S. Majumder, S. Sengupta, S. Chatterjee, G. Khetawat, and J. Das.** 1995. Cholera toxin (CTX) genetic element in *Vibrio cholerae* O139. Microbiology **141:**1977–1983.
6. **Bhaskaran, K., V. B. Sinha, and S. S. Iyer.** 1973. Chromosome mobilization in *Vibrio cholerae* (biotype El Tor) mediated by sex factor P. J. Gen. Microbiol. **78:**119–124.
7. **Bhattacharyya, D., and J. Das.** 1997. The *secY* gene of *V. cholerae*: identification, cloning and characterization. Gene **196:**261–266.
8. **Bik, E. M.** 1996. DNA fingerprinting of *Vibrio cholerae* strains with a novel sequence insertion element: a tool to identify epidemic strains. J. Clin. Microbiol. **34:**1453–1461.
9. **Bukanov, N. O., and D. E. Berg.** 1994. Ordered cosmid library and high-resolution physical-genetic map of *Helicobacter pylori* strain NCTC11638. Mol. Microbiol. **11:**509–523.
10. **Charlebois, R. L., J. D. Hofman, L. C. Schalkwyk, W. L. Lam, and W. F. Doolittle.** 1989. Genome mapping in halobacteria. Can. J. Microbiol. **15:**21–29.
11. **Charlebois, R. L.** 1991. Detailed physical map and set of overlapping clones covering the genome of the archaebacterium *Haloferax volcanii* DS2. J. Mol. Biol. **222:**509–524.
12. **Chaudhuri, K., R. K. Bhadra, and J. Das.** 1992. Cell surface characteristics of environmental and clinical isolates of *Vibrio cholerae* non-O1. Appl. Environ. Microbiol. **58:**3567–3573.
13. **Chuang, S., D. L. Daniels, and F. R. Blatner.** 1993. Global regulation of gene expression in *Escherichia coli*. J. Bacteriol. **175:**2026–2036.
14. **Cole, S. T., and I. S. Girons.** 1994. Bacterial genomics. FEMS Microbiol. Rev. **14:**139–160.
15. **Deckers, H. M., and G. Voordouw.** 1994. Identification of a large family of genes for putative chemoreceptor proteins in an ordered library of the *Desulfovibrio vulgaris* Hidenborough genome. J. Bacteriol. **176:**351–358.
16. **Eiglmeirer, K., N. Honore, S. A. Woods, B. Caydron, and S. T. Cole.** 1993. Use of an ordered cosmid library to deduce the genomic organization of *Mycobacterium leprae*. Mol. Microbiol. **7:**197–206.
17. **Fayet, O., T. Ziegelhoffer, and C. Georgopoulus.** 1989. The *groES* and *groEL* heat shock gene products of *Escherichia coli* are essential for bacterial growth at all temperatures. J. Bacteriol. **171:**1379–1385.
18. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132:**6–13.
19. **Fonstein, M., and R. Haselkorn.** 1993. Chromosomal structure of *Rhodobacter capsulatus* strain SB1003: cosmid encyclopedia and high-resolution physical and genetic map. Proc. Natl. Acad. Sci. USA **90:**2522–2526.
20. **Fonstein, M., and R. Haselkorn.** 1995. Physical mapping of bacterial genomes. J. Bacteriol. **177:**3361–3369.
21. **He, Q., H. Chen, A. Kuspa, Y. Cheng, D. Kaiser, and L. J. Shimkets.** 1994. A physical map of the *Myxococcus xanthus* chromosome. Proc. Natl. Acad. Sci. USA **91:**9584–9587.
22. **Hohn, B.** 1979. *In vitro* packaging of λ and cosmid DNA. Methods Enzymol. **68:**299–309.
23. **Kaper, J. B., G. Morris, and M. M. Levine.** 1995. Cholera. Clin. Microbiol. Rev. **8:**48–86.
24. **Kaper, J. B., H. Lockman, M. M. Baldini, and M. M. Levine.** 1984. A recombinant live oral cholera vaccine. Bio/Technology **2:**345–349.
25. **Keasler, S. P., and R. H. Hall.** 1993. Detecting and biotyping *Vibrio cholerae* O1 with multiplex polymerase chain reaction. Lancet **341:**1661.
26. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *Escherichia coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50:**495–508.
27. **Liu, S.-L., and K. E. Sanderson.** 1992. A physical map of the *Salmonella typhimurium* LT2 genome made by using *Xba*I analysis. J. Bacteriol. **174:**1662–1672.
28. **Lohia, A., S. Majumdar, A. N. Chatterjee, and J. Das.** 1985. Effect of changes in the osmolarity of the growth medium on *Vibrio cholerae* cells. J. Bacteriol. **163:**1158–1166.
29. **Majumder, R., S. Sengupta, G. Khetawat, R. K. Bhadra, S. Roychoudhury, and J. Das.** 1996. Physical map of the genome of *Vibrio cholerae* 569B and localization of genetic markers. J. Bacteriol. **176:**1105–1112.
30. **Manning, P. A., M. W. Heuzenroeder, J. Yeadon, D. I. Leabesley, P. R. Reeves, and D. Rowley.** 1986. Molecular cloning and expression in *Escherichia coli* K12 of the O antigens of the Ogawa and Inaba serotypes of the lipopolysaccharides of *Vibrio cholerae* O1 and their potential for vaccine development. Infect. Immun. **53:**272–277.
31. **Morris, J. G.** 1990. Non-O group I *Vibrio cholerae*: a look at the epidemiology of an occasional pathogen. Epidemiol. Rev. **12:**179–191.
32. **Nandi, S., G. Khetwat, S. Sengupta, R. Majumder, S. Kar, R. K. Bhadra, S. Roychoudhry, and J. Das.** 1997. Rearrangements in the genomes of *Vibrio cholerae* strains belonging to different serovars and biotypes. Int. J. Syst. Bacteriol. **47:**858–862.
33. **Olson, V. M., J. E. Dutchik, M. Y. Graham, G. M. BroDeur, C. Helms, M. Frank, M. MacCollin, R. Scheinman, and T. Frank.** 1986. Random clone strategy for genomic restriction mapping in yeast. Proc. Natl. Acad. Sci. USA **83:**7826–7830.
34. **Panda, D. K., U. Dasgupta, and J. Das.** 1991. Transformation in *Vibrio cholerae* by plasmid DNA. Gene **105:**107–111.
35. **Pearson, G. D. N., A. Woods, S. L. Chiang, and J. J. Mekalanos.** 1993. CTX genetic element encodes a size specific recombination system and an intestinal colonization factor. Proc. Natl. Acad. Sci. USA **90:**3750–3754.
36. **Ramamurthy, T., S. Garg, R. Sharma, S. K. Bhattacharya, G. B. Nair, T.**

Shimada, T. Takeda, T. Karasawa, H. Kurazano, A. Pal, and Y. Takeda. 1993. Emergence of novel strain of *Vibrio cholerae* with epidemic potential in southern and eastern India. Lancet **341:**703–704.
37. **Roy, N. K., G. Das, T. S. Balganesh, S. N. Dey, R. K. Ghosh, and J. Das.** 1982. Enterotoxin, DNA and alkaline phosphatase of *Vibrio cholerae* before and after animal passage. J. Gen. Microbiol. **128:**1927–1932.
38. **Roychoudhury, S., R. K. Bhadra, and J. Das.** 1994. Genome size and restriction fragment length polymorphism analysis of *Vibrio cholerae* strains belonging to different serovars and biotypes. FEMS Microbiol. Lett. **115:** 329–334.
39. **Sahu, G. K., R. Chowdhury, and J. Das.** 1997. The *rpo*H gene encoding σ^{32} homolog of *Vibrio cholerae*. Gene **189:**203–207.
40. **Sankar, P., E. Hutton, R. A. VanBogelen, R. L. Clark, and F. C. Neidhard.** 1993. Expression analysis of cloned chromosomal segments of *Escherichia coli*. J. Bacteriol. **175:**5145–5152.
41. **Trieselmann, B. A., and R. L. Charlebois.** 1992. Transcriptionally active regions in the genome of the archaebacterium *Haloferax volcanii*. J. Bacteriol. **174:**30–34.
42. **Trucksis, M., J. E. Galen, J. Michalski, A. Fasano, and J. B. Kaper.** 1993. Accessory cholera enterotoxin (Ace), the third toxin of a *Vibrio cholerae* virulence cassette. Proc. Natl. Acad. Sci. USA **90:**5267–5271.
43. **Waldor, M., and J. J. Mekalanos.** 1994. ToxR regulates virulence gene expression in non-O1 strains of *Vibrio cholerae* that cause epidemic cholera. Infect. Immun. **62:**72–78.
44. **Wenzel, R., and R. Hermann.** 1989. Cloning of complete *Mycoplasma pneumoniae* genome. Nucleic Acids Res. **17:**7029–7043.
45. **Wilson, K.** 1994. Preparation of genomic DNA from bacteria, p. 2.4.3–2.4.5. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current Protocols in Molecular Biology, vol. 1. John Wiley and Sons, New York, N.Y.