Identification of VCR, a Repeated Sequence Associated with a Locus Encoding a Hemagglutinin in Vibrio cholerae O1

ANDREW BARKER, CHRISTOPHER A. CLARK, AND PAUL A. MANNING*

Microbial Pathogenesis Unit, Department of Microbiology and Immunology, University of Adelaide, Adelaide, South Australia 5005, Australia

Received 9 November 1993/Accepted 14 June 1994

We have determined the nucleotide sequence of a 6.3-kb BamHI fragment of the chromosome of Vibrio cholerae 569B that includes the sequence of the mannos-fucose-resistant hemagglutinin reported previously (V. L. Franzon, A. Barker, and P. A. Manning, Infect. Immun. 61:3032–3037, 1993). This region contains nine copies of a 124-bp direct repeat, here named VCR, of imperfect dyad symmetry, that are shown by Southern hybridization to occur at least 60 to 100 times in the V. cholerae O1 chromosome. Large-scale chromosomal mapping suggests that the repeats are confined to about 10% of the chromosome. Related sequences are also found in non-O1 V. cholerae but not in other members of the family Vibrionaceae. However, VCR is unrelated to other previously described repetitive sequences.

Vibrio cholerae of the O1 serotype is the causative agent of the disease Asiatic cholera. Pathogenesis occurs as a number of sequential, interdependent steps: bacteria are ingested via contaminated food or water (38), colonize the intestinal epithelium, and elaborate a number of extracellular proteins, leading to profuse watery diarrhea and causing high mortality if untreated (34, 35, 45). Since colonization is a prerequisite to establishing a productive infection (48), the factors responsible have been the subject of much research. A number of potential colonization factors, including fimbriae such as TCP (54) and other hemagglutinins (24), have been identified, but to date only TCP has been shown to be a virulence determinant in humans (25). The mannos-fucose-resistant hemagglutinin (MFRHA) (18, 55, 56) has also been implicated as a virulence determinant (17), but its exact role in colonization remains to be determined.

Repetitive sequences have been identified in a wide variety of eubacterial chromosomes (33), the best characterized of these being the REP or PU sequences identified in members of the family Enterobacteriaceae (19, 52), with related sequences being detected in diverse bacterial species (57). There is considerable controversy as to whether such sequences have any function: identification of interactions in vitro with a number of nucleoid-associated proteins, such as DNA gyrase, the histone-like protein HU, and DNA polymerase I, have led to proposals that REP sequences might represent a site for the assembly of scaffold proteins on DNA (20, 59, 60). These repeats are similar to the eukaryotic matrix association regions, which contain repetitive sequence elements that bind topoisomerase II (4, 51). However, there are other repetitive sequences also involved in scaffold attachment that do not bind topoisomerase II (41). REP sequences often occur in clusters called REP elements (21, 26), at least 90 of which are found randomly distributed in the Escherichia coli chromosome (12). This approximates the number of supercoiled domains in the chromosome (20, 59, 60) and also the number of “high-affinity” gyrase cleavage sites (9). Consequently, it has been suggested that REP sequences are involved in the maintenance of chromosomal structure (20, 59, 60). Alternatively, the proposal that REP sequences are selfish DNA (26) has not been discounted and remains a possibility (33).

Another role attributed to REP sequences is that they act as the endpoints of RecA-independent recombination events, including large-scale DNA rearrangements and deletions (49). This can have important implications for bacterial virulence, as chromosomally encoded virulence determinants, at least in pathogenic E. coli, are located in “pathogenicity islands” rather than being randomly distributed throughout the genome (23). It was also noted that these islands could be deleted, leading to a loss of virulence genes. Although no mechanism has been proposed to mediate this phenomenon, it is worth noting that large-scale DNA rearrangements can be mediated by DNA gyrase in the absence of both extensive sequence homology at the recombinational endpoints and recA function (40) and that large deletions whose end points are REP sequences have been characterized in Salmonella typhimurium (49). The high degree of genetic instability in Streptomyces species (32) and spontaneous deletions leading to loss of the pigment phenotype in Yersinia pestis (14) are also associated with deletions whose endpoints are repetitive sequences.

To aid our investigation of the role of the MFRHA in the virulence of V. cholerae, we have sequenced the region surrounding the corresponding genetic locus. We have shown that this section of the chromosome contains a large number of copies of a 124-bp direct repeat originally detected by van Dongen et al. (56), and we discuss the possible roles that these repeats could have in the expression of closely linked genes and on the structure and evolution of the V. cholerae chromosome.

MATERIALS AND METHODS

Bacterial strains and media. The E. coli strains used were JM101 ([F traD36 proAB lacZAM15 supE thi Δ(lac-proAB) λ C]) (58) and DH5α [F ϕ80d lacZ ΔM15 Δ(lac-argF)U169 recA1 endA1 hsdR17(rK− mK− s−) supE44 thi-1 gyrA96 relA1 λ−] (Bethesda Research Laboratories). Other strains used are listed in Tables 1 through 4. Unless otherwise stated, the growth medium used was nutrient broth (Oxoid), except for phage propagation, for which M9 minimal medium and 2xTY broth (39) were used.

DNA sequencing. The nucleotide sequence of the 6.3-kb
BamHI fragment was determined by subcloning small, overlapping restriction fragments derived from plasmid pMP471 (18) into the phage vectors M13 mp18 and mp19 (60). Sequences were determined with chain terminators (47). Sequence or Sequenase II enzyme according to the manufacturer’s specifications (United States Biochemical Corporation), and an “oligo-walk” strategy with primers generated from partial sequence and manufactured on an Applied Biosystems model 381A DNA synthesizer. When cloned restriction fragments did not overlap, the junction in pMP471 was sequenced by using supercoiled plasmid template (7) with primers binding near the restriction site (not shown). The nucleotide sequence from bp 2727 to 3380 has been reported previously (17). Sequences were compared with entries in the GenBank and EMBL databases with the BLAZE algorithm (6).

Other DNA manipulations. DNA was manipulated for subcloning by standard procedures (46) with restriction endonucleases and other DNA-modifying enzymes purchased from New England Biolabs, Bethesda Research Laboratories, Boehringer Mannheim, or Amersham. For Southern blots, chromosomal DNAs were prepared as described before (36), digested with BamHI or HindIII, and run on standard 0.8 or 1% agarose gels, respectively. Samples for pulsed-field gels were prepared by a modification of published procedures (43, 50) as follows. A 20-ml culture of *V. cholerae* in M9 medium (39) was grown at 30°C to an A<sub>600</sub> of 0.6 and incubated with chloramphenicol (final concentration, 180 μg/ml) for 60 min. Cells were then pelleted by centrifugation in a bench centrifuge at 5,000 × g for 10 min, washed in 10 ml of 1 M NaCl–10 mM Tris-HCl (pH 8.0), resuspended in 4 ml of the same buffer, and warmed to 45°C. Agarose beads were formed by adding 5 ml of 1% LGT agarose (UltraPure; Bethesda Research Laboratories) in 1 M NaCl–10 mM Tris-HCl (pH 8.0)–20 ml of sterile mineral oil (Primol 352; ESSO), both at 45°C, to the cells, emulsifying the mixture, pouring it into 100 ml of ice-cold 1 M NaCl–10 mM Tris-HCl (pH 8.0), and stirring the mixture for 5 min. Beads were then pelleted by centrifugation at 4°C and 500 × g in 50-ml tubes for 10 min, and the mineral oil was removed. The tubes were recentrifuged, and the beads in the pellet were combined in a single tube and then centrifuged as before. The beads were then resuspended in 15 ml of 1 mg/ml protease K–1% sodium n-lauroyl sarcosine–0.5 M EDTA (pH 9.2) warmed to 65°C and incubated at 65°C for at least 48 h. A lysozyme treatment step was omitted because of excessive nicking of the DNA during this step (1), presumably by secreted DNases (15). After protease K digestion, the beads were recovered by centrifugation as before and washed a minimum of five times for a total of 2 h in 10 mM Tris-HCl (pH 7.5)–0.1 mM EDTA at room temperature before storage at 4°C. Restriction endonuclease digestion of agarose beads was performed under standard conditions, and digests were separated on a 1% agarose gel with a Biorad CHEF II apparatus and run conditions of initial A time, 1 min, final A time, 10 min, and 200 V for 18 h at 14°C in 0.5× TBE. All gels were transferred to nitrocellulose filters by standard conditions and were probed with radioactive or digoxigenin (DIG)-labeled probes as described below, with stringent washes and autoradiography done as described before (46). Stringent washes for filters probed with oligodeoxynucleotide probes were done with tetramethyl ammonium chloride (58). DIG-labeled probes were detected with anti-DIG–horseradish peroxidase conjugate followed by the addition of enhanced chemiluminescence substrate (Amersham) and exposure to X-ray film. Densitometer analysis of autoradiographs was performed on a Molecular Dynamics model 300A computing densitometer. Filters were stripped for reprobing with boiling water and checked for stripping by autoradiography before reprobing. The oligodeoxynucleotide probe used was 5'CGCGGCTCAATGGGACTGAA-3' and was synthesized on an Applied Biosystems model 381A DNA synthesizer. Fragment probes were the 0.7-kb *Clai*-XbaI fragment of

**TABLE 1. Characteristics of *V. cholerae* O1 strains**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Biotype</th>
<th>Serotype</th>
<th>Source (yr of isolation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>569B</td>
<td>Classical</td>
<td>Inaba</td>
<td>India* (1946)</td>
</tr>
<tr>
<td>CA401</td>
<td>Classical</td>
<td>Inaba</td>
<td>Texas* (1953)</td>
</tr>
<tr>
<td>CA411</td>
<td>Classical</td>
<td>Ogawa</td>
<td>Texas* (1953)</td>
</tr>
<tr>
<td>C21</td>
<td>Classical</td>
<td>Ogawa</td>
<td>Holland* (1957)</td>
</tr>
<tr>
<td>AA14041</td>
<td>Classical</td>
<td>Ogawa</td>
<td>Bangladesh* (1985)</td>
</tr>
<tr>
<td>Z17561</td>
<td>Classical</td>
<td>Inaba</td>
<td>Bangladesh* (1985)</td>
</tr>
<tr>
<td>O17</td>
<td>El Tor</td>
<td>Ogawa</td>
<td>India* (pre-1965)</td>
</tr>
<tr>
<td>C5</td>
<td>El Tor</td>
<td>Ogawa</td>
<td>Holland* (1957)</td>
</tr>
<tr>
<td>C31</td>
<td>El Tor</td>
<td>Ogawa</td>
<td>Holland* (1957)</td>
</tr>
<tr>
<td>AA14073</td>
<td>El Tor</td>
<td>Ogawa</td>
<td>Bangladesh* (1985)</td>
</tr>
<tr>
<td>BM69</td>
<td>El Tor</td>
<td>Inaba</td>
<td>India* (1985)</td>
</tr>
</tbody>
</table>

* K. Bhaskaran, Central Drug Institute, Lucknow, India.

* J. Berry, University of Texas, Austin.

* P. Guinée, Rijksinstituut voor Volksgezondheid en Milieuhygiëne, Bilthoven, Holland (deceased).

* B. Kay, International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh.

* S. Pal, National Institute of Cholera and Enteric Diseases, Calcutta, India.

**TABLE 2. *V. cholerae* non-O1 strains**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Isolate type</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCV165</td>
<td>Environmental</td>
<td>India*</td>
</tr>
<tr>
<td>N34</td>
<td>Human (septicemia)</td>
<td>Australia*</td>
</tr>
<tr>
<td>N41</td>
<td>Human (traveler's diarrhea)</td>
<td>Australia*</td>
</tr>
<tr>
<td>N50</td>
<td>River water</td>
<td>Australia*</td>
</tr>
<tr>
<td>N125</td>
<td>Wallaby</td>
<td>Australia*</td>
</tr>
<tr>
<td>BV7</td>
<td>Environmental</td>
<td>Maryland*</td>
</tr>
<tr>
<td>BV22</td>
<td>Environmental</td>
<td>Maryland*</td>
</tr>
<tr>
<td>BV41</td>
<td>Environmental</td>
<td>Maryland*</td>
</tr>
</tbody>
</table>

* See Table 1, footnote a.

* P. Desmarchelier, University of Queensland, Brisbane, Australia.

* M. Voll, University of Maryland, College Park.

**TABLE 3. Other *Vibrio* spp. strains**

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th>Features</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>V. mimicus</em></td>
<td>V800</td>
<td>None described</td>
<td>Australia*</td>
</tr>
<tr>
<td><em>V. fluvialis</em></td>
<td>V564</td>
<td>None described</td>
<td>Australia*</td>
</tr>
<tr>
<td><em>V. parahaemolyticus</em></td>
<td>NCTC 1088</td>
<td>Kanagawa*</td>
<td>Australia*</td>
</tr>
</tbody>
</table>

* S. Attridge, University of Adelaide, Adelaide, Australia.

* Institute of Medical and Veterinary Science (Adelaide, Australia) culture collection.

**TABLE 4. *Aeromonas* spp. strains**

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. hydrophila</em></td>
<td>A006</td>
<td></td>
</tr>
<tr>
<td><em>A. sobria</em></td>
<td>A187</td>
<td></td>
</tr>
<tr>
<td><em>A. sobria</em></td>
<td>A191</td>
<td></td>
</tr>
<tr>
<td><em>A. caviae</em></td>
<td>V14</td>
<td></td>
</tr>
<tr>
<td><em>Aeromonas</em> spp.</td>
<td>ABI</td>
<td></td>
</tr>
</tbody>
</table>

* All strains were from M. Atkinson, University of South Australia, Adelaide.
plasmid pPM1127 (18), referred to in the text as the mrhB probe; the 0.2-kb XbaI-NsiI fragment of plasmid pPM1631 (2), referred to in the text as the VCR probe; and the 0.2-kb ApaI-PstI fragment of pPM1648 (1), referred to in the text as the ORF2 probe. Probes referred to in the text as to the proximal and distal portions of plasmid pPM471 were the 0.7-kb BamHI-HindIII fragment and the 0.15-kb EcoRI-BamHI fragments of pPM471 (18), respectively, subcloned into M13mp19 (61). Oligonucleotides were 5'-end labeled with [γ-32P]ATP (BRESATEC) in a polynucleotide kinase reaction (44), restriction fragments were gel purified and labeled with [α-32P]dCTP (BRESATEC) in a random-primed reaction (13), and M13mp19 subclones were labeled by extension from the primer 5'-CACCATTTCCACACAC-3' (61), incorporating 11-DIG-dUTP (Boehringer Mannheim).

Nucleotide sequence accession number. The DNA sequence described in this study is available under GenBank and EMBL accession no. X64097.

RESULTS

Nucleotide sequence and analysis. We initially set out to determine the nucleotide sequence of a 2.4-kb HindIII fragment of V. cholerae 569B encoding the MFRHA (18). A report of the sequence of part of this region appeared recently (17). However, the presence of three copies of a 124-bp direct repeat within this region (see below) led us to perform Southern hybridizations (not shown), which indicated the presence of multiple copies of this repeat within the original MFRHA clone, pPM471 (18). Therefore, the sequence of the entire insert of pPM471, corresponding to a 6.3-kb BamHI fragment of V. cholerae 569B, was determined and is represented schematically in Fig. 1.

Ten open reading frames (ORFs) are apparent within the sequenced region, eight of which are of the same predicted transcriptional polarity (Fig. 1). Two ORFs, encoding 7-kDa and 27-kDa proteins, respectively, are associated with MFRHA activity (17, 56), and we propose the gene names mrhA for the ORF encoding the 7-kDa protein and mrhB for the ORF encoding the 27-kDa protein (for mannose-fucose-resistant hemagglutinin). The sequences of all ORFs were compared with entries in the PIR (Release 33) and SWISS-PROT (Release 22) databases for similar entries. Significant matches were obtained for ORF2 (1); ORF3.1, which shows 32% identity to the amino-terminal 72 amino acids of DnaT (primosomal protein i) of E. coli (36); ORF3.2, which shows 56% identity with the hypothetical RelE protein of E. coli (4); and mrhB (1). The alignments for ORF3.1 and ORF3.2 are shown in Fig. 2. Attempts to examine the functions of ORF3.1 and ORF3.2 by engineering mutations in them or by subcloning were unsuccessful (not shown).

FIG. 1. Genetic organization of a 6.3-kb BamHI fragment of V. cholerae 569B. ORFs are represented by open boxes, with the polarity of transcription indicated by arrows, and the approximate predicted molecular weights (in thousands) of the products are shown. The sequence of the MFRHA (mrhB) has been reported previously (17) and encompasses bp 2727 to 3380 of this sequence. The locations of nine copies of a 124-bp direct repeat (VCR) are indicated by solid boxes. Some restriction sites in the region are indicated.

FIG. 2. Pairwise alignments of homologous protein sequences. (A) Alignment between ORF3.1 and the amino-terminal 72 amino acids of DnaT (primosomal protein i) of E. coli. Full-length DnaT is 179 amino acids long (37). (B) Alignment between ORF3.2 and the hypothetical RelE protein of E. coli (3). Alignments were performed with the BLAZE algorithm (6). Identical amino acids are joined with dots. The boundaries of the region of similarity are indicated by an X above and below the sequences.
The most striking feature of the sequence presented is the occurrence of nine 124-bp repeats, all of which are in the same orientation relative to one another, and all of which are found only outside of predicted ORFs, that is, in an extragenic location (Fig. 1). The sequences of these repeats, along with four repeats sequenced by others (53, 56), show an overall similarity of about 92% (Fig. 3A). Several stretches of dyad symmetry can be identified within the repeats, and the whole repeat itself shows imperfect dyad symmetry, which if transcribed would give a stable secondary structure. In the case of
HindIII-digested

5454

BARKER

with biotypes
test indicated the
occurrence shown in
repeats VCR, the consensus
hybridizing large number of bands
detected strain C31, of the chromosomal
digestion reprobed and under conditions of
lanes 7, C5; lanes 8, C31; lanes 9, AA14073; lanes 10, BM69. A 6.3-kb BamHI fragment (A) and 2.4-kb HindIII fragment (B) that correspond to the hybridizing fragments in Fig. 5 are indicated by an arrowhead.

FIG. 4. Hybridization of a VCR-specific oligonucleotide probe to V. cholerae O1 chromosomal DNA. (A) BamHI-digested DNA; (B) HindIII-digested DNA. V. cholerae O1 strains: lanes 1, 569B; lanes 2, CA401; lanes 3, CA411; lanes 4, AA14041; lanes 5, Z17561; lanes 6, O17; lanes 7, C5; lanes 8, C31; lanes 9, AA14073; lanes 10, BM69. A 6.3-kb BamHI fragment (A) and 2.4-kb HindIII fragment (B) that correspond to the hybridizing fragments in Fig. 5 are indicated by an arrowhead.

the consensus sequence, a loop of free energy of −58.9 kCal/mol is predicted (Fig. 3B). We propose to name these repeats VCR, for V. cholerae repetitive DNA sequence.

Chromosomal distribution of VCR elements. Because of the number of repeats found in the region of the V. cholerae chromosome shown in Fig. 1, we investigated the possibility of the occurrence of further copies of VCR by probing 10 different V. cholerae O1 strains of both the classical and El Tor biotypes with a VCR-specific oligodeoxynucleotide probe. This test indicated the presence of a large number of copies of VCR (Fig. 4). To aid quantitation of VCR and to ensure that the large number of bands detected were not due to partial digestion of the chromosomal DNAs, the filters were stripped and reprobed with the mrhB-specific probe (Fig. 5). Except for strain C31, in which this region is deleted (55) (see below), this probe binds to a single fragment of the size expected from sequence analysis (Fig. 1). Given that the VCR-specific probe used will bind to six of the nine VCR sequences of pPM471 under conditions of high stringency (not shown), an estimate of the number of VCRs in the chromosome was obtained by quantitating the density of the bands in Fig. 4A (not shown). This gives an estimate of 90 to 100 copies of VCR in strains of the classical biotype and 60 to 80 copies of VCR in strains of the El Tor biotype. The patterns obtained with other probes which could be expected to identify all copies of the repeat were not markedly different, and so these estimates represent a reasonable approximation.

The pattern of hybridization shown in Fig. 4B shows remarkable similarity to the pattern observed when a number of MFRHA-positive cosmid clones were probed with the 2.4-kb HindIII fragment indicated in Fig. 5B (16). As this fragment contains three copies of VCR, the cross-hybridization observed (16) was probably due to these copies of VCR, and this result suggests that the copies of VCR observed in the chromosome of V. cholerae (Fig. 4) are linked. This possibility was investigated by probing digested DNA separated on a pulsed-field gel with the VCR-specific oligodeoxynucleotide used above (Fig. 6A). Two hybridizing bands are observed for strain 569B, of 285 and 195 kb for NotI-digested DNA and of 265 and 205 kb for SfiI-digested DNA. For strain O17,

FIG. 5. Hybridization of an mrhB (MFRHA)-specific probe to V. cholerae O1 chromosomal DNA. (A) BamHI-digested DNA; (B) HindIII-digested DNA. Lanes are the same as in Fig. 4. Hybridizing fragments are indicated by an arrowhead.
two hybridizing bands are observed with NotI-digested DNA, of 265 and 125 kb, and a single band of 300 kb is observed with SfiI-digested DNA. Finally, strain C31 shows single bands of 245 and 235 kb with NotI- and SfiI-digested DNA, respectively (Fig. 6A). The filter was stripped and reprobed with the mrrB-specific probe used above, and the result of this is shown in Fig. 6B. In this case, bands of 285 and 195 kb are detected for 569B DNA digested with NotI, bands of 265 and 205 kb are detected for 569B DNA digested with SfiI, and bands of 265 and 300 kb are detected for O17 DNA digested with NotI and SfiI, respectively. As expected, no hybridizing fragments were detected for C31. The detection of two hybridizing bands for 569B with an mrrB probe (Fig. 6B) is unexpected, as a single hybridizing band is detected for both BamHI- and HindIII-digested 569B DNA when probed with the same probe (Fig. 5), indicating that the larger band is probably due to incomplete digestion of the samples. In case the oligodeoxynucleotide probe used in Fig. 6A was not hybridizing with some copies of VCR (see Discussion), leading to a biased result, the filter was again stripped and reprobed with the VCR-specific fragment probe. No additional bands were apparent (not shown), indicating that the oligonucleotide probe has not “missed” significant bands through mismatch. Therefore, most copies of VCR lie within a region of approximately 200 kb in strain 569B, 300 kb in strain O17, and 240 kb in strain C31. Given a genome size of approximately 2,800 kb for V. cholerae O1 (30), this correlates with a region encompassing approximately 10% of the chromosome.

Specificity of VCR within the family Vibrionaceae. To see whether VCR-related sequences could be detected in other members of the family Vibrionaceae, a number of different species were probed with the VCR-specific oligodeoxynucleotide used above. Hybridization was observed in a number of non-O1 serotype V. cholerae, but no hybridization was observed for the other Vibrio or Aeromonas spp. tested (Fig. 7). These filters were stripped and reprobed with the mrrB-specific probe used above (Fig. 5), but a positive signal was obtained only for V. cholerae O1 strains 569B and C21 (not shown).

Is VCR involved in deletion events? In an attempt to define the deletion in strain C31, BamHI- and HindIII-digested chromosomal DNAs were probed with probes corresponding to the ends of plasmid pPM471. The distal probe hybridizes to a single band for 569B (6.3 kb for BamHI-digested DNA and 6.5 kb for HindIII-digested DNA) and no band for strain C31 (Fig. 8B), but the proximal fragment hybridizes to two bands for 569B (11 and 6.3 kb for BamHI-digested DNA and 2.8 and 1.2 kb for HindIII-digested DNA) and one band for strain C31 (10 kb for BamHI-digested DNA and 5.4 kb for HindIII-digested DNA) and a faint additional band of 3.9 kb for HindIII-digested C31 DNA (Fig. 8A). Next, the same samples were probed with an ORF2-specific probe, which again gives two bands of 6.3 and 4.5 kb for BamHI-digested 569B DNA and a single band of 11.5 kb for BamHI-digested C31 DNA (Fig. 8C). Bands were not apparent with HindIII-digested DNA in this case. This indicates that some sequences contained within pPM471 have been duplicated but that the deletion in C31 is larger than the cloned DNA. Therefore, further mapping of the deletion endpoints was not undertaken as part of this study.
The nucleotide sequence reported here extends that reported previously (17, 56) to include an additional eight ORFs linked to the genes encoding the MFRHA. The sequence of mhrA presented here is identical to that of van Dongen et al. (56), but that of mhrB shows six changes at the nucleotide level (17). Comparison of ORF sequences with entries in the databases revealed homology between ORF3.1 and the amino-terminal 72 amino acids of DnaT (primosomal protein i) of E. coli (38) and between ORF3.2 and the hypothetical E. coli RelE protein, presumed to be the translational inhibitor accumulated in relB mutants during amino acid starvation (3, 31). Attempts to engineer mutations in these ORFs by deleting portions of plasmid pPM471 (18) were unsuccessful, presumably because the desired products were lethal (not shown). Therefore, their functions remain unclear. Homology was also detected for ORF2 and mhrB and will be described elsewhere (1). None of the other ORFs reveal significant similarity to known protein sequences. Given that E. coli clones in which these other ORFs have been deleted retain hemagglutinating activity (17, 18, 55, 56), it seems unlikely that they are directly involved in this phenotype. Thus, the MFRHA appears to be an adhesin encoded by one or two ORFs, unlike previously characterized fimbrial adhesins, which are usually encoded by extensive operons (28).

Van Dongen et al. (56) identified two copies of a repeat sequence that in this study has been shown to be present in multiple copies in the chromosome of V. cholerae O1. An estimate of 60 to 100 copies of VCR in V. cholerae O1 was obtained. It is known that the oligonucleotide binds to only six of the nine copies of VCR present on pPM471 at the degree of stringency used (not shown). However, the VCR fragment probe (random primed) that is capable of binding to all copies of VCR does not reveal any major new bands when used to probe a filter identical to that shown in Fig. 3A. VCR shows several regions of dyad symmetry, and the whole repeat shows imperfect symmetry that if transcribed would form a stable secondary structure. This raises the possibility that VCR might play a role in transcription termination in V. cholerae. Transcription has been shown to terminate within at least one of the repeats, but it seems unlikely that all copies of VCR act as transcriptional terminators, because RNase protection experiments with VCR-specific probes of either orientation reveal multiple bands consistent with the presence of full-length copies of VCR within V. cholerae total RNA (2).

Furthermore, the repeat upstream of mhrA is unlikely to act as a terminator because of the polarity of the flanking ORFs. Similarly, only a subset of REP sequences have been shown to act as transcriptional terminators (22). Therefore, some other role must lead to their maintenance.

Repetitive extragenic sequences have been characterized in a number of bacterial genera but do not show any recognizable sequence similarities (32). REP sequences found in E. coli and in diverse genera (57) have been shown to bind DNA polymerase 1 (20), DNA gyrase (59), and IHF (5, 42) as well as the histone-like protein HU (60). In S. typhimurium, REP sequences have been shown to be the endpoints of recombination events that occur in a recA-independent manner (49), and in E. coli, DNA gyrase (topoisomerase II) has been shown to catalyze illegitimate recombination (40). However, REP sequences are distributed more or less randomly throughout the E. coli chromosome (12), whereas VCR sequences are restricted to a region of approximately 10% of the V. cholerae chromosome. Another class of repetitive elements designated ERIC (29) have also been identified, with related sequences in V. cholerae. However, although they have a similar length and the potential to form a stable stem-loop structure, like VCR, there is no sequence homology.

The most striking feature of VCR is its localization to a limited region of the chromosome. An upper limit on the region occupied by VCR has been determined by pulsed-field gel electrophoresis, but the lower limit is still unknown. Assuming a density of VCR repeats equivalent to that seen on pPM471 (i.e., nine in 6 kb), a lower limit of about 42 to 70 kb (depending on the strain) can be predicted.

The orientation of the nine copies of VCR within pPM471 is such that recombination events between them will lead to deletion of the intervening DNA; the presence of multiple copies of VCR on multicopy plasmids and M13 vectors leads to a very high rate of deletion events in recA+ strains (1, 10, 16). Such a “deletogenic” activity would make VCR an effector sequence in V. cholerae for deletion events associated with pathogenicity islands. Despite this, the VCR region of V. cholerae O1 seems to be stable within biotypes, with strains isolated over a 40-year period showing a similar banding pattern on Southern hybridization. However, perhaps this region of the chromosome has evolved by tandem duplications mediated by VCR, in a way similar to that by which the much larger RS1 elements are involved in toxin gene amplification (44).

An alternative possible role for VCR is to provide a region where foreign genes can integrate into the chromosome. Of 500 V. cholerae O1 isolates screened by Takeda et al. (53), one that was ST+ (heat-stable toxin) was identified. DNA sequencing revealed that the gene is flanked by copies of VCR, suggesting that these sequences are somehow involved in the acquisition of the ST gene (53). Elements for gene integration, termed integrons, have been identified in gram-negative bacteria and exhibit conserved sequences (8), but these do not show similarity to VCR, so if VCR does fulfill a similar function, then it represents a different class of integron than those described previously (8). However, as sequencing has revealed that ST genes in virulent non-O1 V. cholerae are also flanked by copies of VCR (53), it is also possible to explain the acquisition of the ST gene by this strain as occurring by a conjugation or transduction event from a non-O1 strain with-

FIG. 8. Hybridization of probes derived from the ends of pPM471 to V. cholerae 569B and C31 DNAs. (A) DNA hybridized with a probe from the proximal end of pPM471; (B) DNA hybridized with a probe from the distal end of pPM471; (C) DNA hybridized with an ORF2-specific probe. Lanes in all panels: 1, 569B, BamHI digested; 2, 569B, HindIII digested; 3, C31, BamHI digested; 4, C31, HindIII digested. The position of a 6.3-kb band that hybridizes with the mhrB probe in BamHI-digested DNA is indicated by an arrowhead.
out invoking a novel integration mechanism. Another alternative is that the extreme clustering of VCR could be due to the entire region, including VCR sequences, originating in a foreign host, but the presence of genes such as mrrh, exclusively in O1 V. cholerae (55), (not shown), and ST toxin genes, predominantly in non-O1 V. cholerae (53), makes this possibility seem less likely. Similarly, the suggestion that a single copy of VCR has integrated and is slowly spreading throughout the chromosome requires an explanation of why the spread is restricted to only one region of the chromosome and is not occurring randomly. Nevertheless, there are insufficient data to discount any of these possibilities.

Although the biological function of VCR is not known, the extreme clustering and maintenance of these elements suggest that each fulfills the same or a similar role. An earlier proposal that REP sequences may be involved in bacterial speciation within the family Enterobacteriaceae (21) has been questioned, as REP-related sequences have recently been identified in diverse bacterial taxons (57). We have not detected VCR-related sequences in bacteria other than V. cholerae in the limited survey reported here but are loath to make a similar suggestion for VCR, as Takeda et al. (53) have identified cross-reacting sequences in Vibrio mimicus by colony blot. Identification of proteins (if any) that interact specifically with VCR should provide some clues as to their function and hopefully help explain how they have persisted in what is potentially a deleterogenic arrangement without being lost from the V. cholerae genome.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of the National Health and Medical Research Council of Australia, the Diarrhoeal Diseases Programme of the World Health Organization, and the Clive and Vera Ramaciotti Foundations. A.B. is the recipient of an Australian Postgraduate Research Award priority scholarship.

We are grateful to Sue Williams for comments on the manuscript, Chris Cursaro for invaluable assistance with oligonucleotide synthesis and photography, and Steve Attridge for suggesting the acronym VCR. We are also grateful to M. W. Heunenroeder for allowing us access to a Biorad CHEF II apparatus and to S. Bowden for allowing access to a densitometer.

REFERENCES

2. Barker, A., S. G. Williams, and P. A. Manning. Submitted for publication.