Recombination between rRNA operons created most of the ribotype variation observed in the seventh pandemic clone of *Vibrio cholerae*

Ruiting Lan and Peter R. Reeves

Author for correspondence: Peter R. Reeves. Tel: +61 2 9351 6045. Fax: +61 2 9351 4571. e-mail: reeves@angis.org.au

Department of Microbiology, University of Sydney, NSW 2006, Australia

Individual *rrn* operons and their flanking regions have been analysed in a study of the molecular basis of ribotype variation in the seventh pandemic clone of *Vibrio cholerae*. The genome of an early isolate of the seventh pandemic clone had nine *rrn* operons of which two were in tandem with other *rrn* operons. The site for *BglII*, the most discriminatory enzyme used for ribotyping, was found to be present in the 16S sequence of three of the operons of the earliest isolate. This site was observed to be gained or lost in specific operons in many later isolates, presumably by recombination, and this gave most of the ribotype variation. Additional *rrn* recombination events were uncovered by analysis of the 16S-23S intergenic spacers associated with each operon. Spacers of 431, 509, 607 and 711 bp were found. A total of at least eight *rrn* recombination events were detected. Three *rrn* loci were primarily involved in this recombination, with four new forms generated from that in the early strains for operon B and two new forms each for operons C and G. In addition there was variation due to deletion of tandem operons. The frequency of recombination between *rrn* operons was very high as there were nine new ribotypes found among 47 isolates sampled over the 33 year period of study. This means that any variation could undergo precise reversion by the same recombination event within the time frame covered by the study.

Recombination between *rrn* operons may be a factor in ribotype variation in all systems. The recombination observed is thought to be that which results in concerted evolution and the data give an indication of the rate involved.

**Keywords:** ribotyping, *rrn* recombination, *Vibrio cholerae*

**INTRODUCTION**

Cholera is caused by toxigenic strains of *Vibrio cholerae*. Seven pandemics of cholera have been recorded since 1817, the seventh starting in 1961 and continuing today. The seventh pandemic clone is relatively homogeneous: there is a single electrophoretic type in Asia and Africa (Evins et al., 1995; Salles & Momen, 1991) while the South American form, which is a variant of the seventh pandemic clone, differs at one locus; all seventh pandemic isolates have the same *asd* gene sequences (Karaolis et al., 1995) and *ctxB* gene sequences (Olsvik et al., 1993). The seventh pandemic arose about 36 years after the sixth pandemic subsided in 1925 and the new strain differed in the presence of haemolysin, acetoin fermentation (Voges–Proskauer test) and sequence of the toxin coregulated pilus. Note that the sixth pandemic strain still persists with non-pandemic status. The seventh pandemic then appears to be caused by a clone with a well-defined date of origin as a pandemic organism, although there are some interesting related (El Tor) isolates from outbreaks between 1930 and 1960 (Barua, 1992), and also toxigenic isolates from the US Gulf and Australia which appear to be related (Barua, 1992; Blake, 1994).

Ribotyping provided the most useful data for studying variation within the seventh pandemic clone (Evins et al., 1995; Faruque et al., 1994, 1995; Karaolis et al., 1994; Koblavi et al., 1990; Popovic et al., 1993;
Wachsmuth et al., 1993). In our ribotyping study with strains isolated from 1961 (Indonesia and Hong Kong) to 1992, including African and Asian isolates (Karaolis et al., 1994), 11 polymorphic sites and 16 patterns were identified using BglI and SalI. The seven strains isolated up to 1966 were identical and the same pattern was observed throughout the period of study. From 1968 (there were no 1967 samples) there were additional ribotypes and after 1971 considerable variation developed with two major lineages, one confined to Asia and the other present in both Africa and Asia. These data confirmed the belief that the seventh pandemic was caused by a single clone as there was no detectable variation detected with ribotyping data. The resurgence of the pandemic.

Ribotyping is a form of RFLP in which rRNA or DNA is used as a probe. The main advantage of ribotyping is that there are several copies of the rRNA gene cluster in the genome. The typing therefore provides information on several genome regions simultaneously. In this study, we aimed to reveal the specific changes behind ribotype variation. We found to our surprise that all but one ribotype variant is due to changes within the rRNA operons, resulting from recombination between operons. The recombinants are detected because minor sequence variation between operons coincidentally confers the presence or absence of a BglI site in the 16S RNA gene. This also explains the much higher level of variation detected with BglI than with other enzymes.

When a gene or an operon such as the rRNA operon is present in more than one copy, it is generally found that all forms from a given species are close to identical but that there are normal levels of sequence difference between species (Li & Graur, 1991). To explain the divergence between the genes in different species but not between copies of the gene in any given species, it has been assumed that recombination between genes within the species leads to any mutation being spread around the several copies and either being fixed in all of them or lost from all. This process allows the several copies of the gene to evolve in concert and has been called 'concerted evolution'. The recombination events involved in concerted evolution most probably occur by gene conversion, a non-reciprocal recombination event in which one copy of the gene is converted to the sequence present in another. The mechanism is now known to involve transfer of one strand of DNA (strand invasion) followed by repair and replication to synthesize both copies of the gene based on the sequence of the donor strand (Leach, 1996). We interpret our data in terms of concerted evolution and the recombination events between rRNA genes observed by us provide information on the dynamics of concerted evolution, and show that it must be taken into account in analysis of ribotyping data.

## METHODS

### Bacterial isolates and DNA preparation.

The isolates used in this study are listed in Table 1. Chromosomal DNA preparation was as described previously by Bastin et al. (1991).

### Southern hybridization.

DNA (2 μg) was digested with an appropriate restriction enzyme. BglI, PvuI and KpnI were from Boehringer Mannheim. I-CeuI was from New England Biolabs. Southern blotting was done as described previously by Karaolis et al. (1994). Probe labelling and hybridization were done using the DIG system according to the manufacturer's instructions (Boehringer Mannheim).

### PCR and sequencing.

The basic programme for PCR was 35 cycles of 15 s at 94 °C, 30 s at 50–60 °C and 1 min at 72 °C with an initial step of 2 min at 94 °C and final extension of 4 min at 72 °C. For further amplification from a PCR product, a 1000–100000-fold dilution of initial PCR product was used as template and 30 cycles were performed. The sequences of the primers are given in Table 2. As many combinations of primers were used for PCR amplification, a schematic

### Table 1. Bacterial strains used in this study

<table>
<thead>
<tr>
<th>Strain</th>
<th>Year isolated</th>
<th>Country of isolation</th>
<th>Source</th>
<th>Original laboratory identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>M803</td>
<td>1961</td>
<td>Hong Kong</td>
<td>Institut Pasteur, Paris, France</td>
<td>HK1</td>
</tr>
<tr>
<td>M807</td>
<td>1966</td>
<td>Vietnam</td>
<td>Institut Pasteur, Paris, France</td>
<td>601</td>
</tr>
<tr>
<td>M811</td>
<td>1971</td>
<td>Burma</td>
<td>Institut Pasteur, Paris, France</td>
<td>930029</td>
</tr>
<tr>
<td>M812</td>
<td>1971</td>
<td>Chad</td>
<td>Institut Pasteur, Paris, France</td>
<td>930046</td>
</tr>
<tr>
<td>M813</td>
<td>1972</td>
<td>Senegal</td>
<td>Institut Pasteur, Paris, France</td>
<td>9292</td>
</tr>
<tr>
<td>M820</td>
<td>1978</td>
<td>Malaysia</td>
<td>Institut Pasteur, Paris, France</td>
<td>EB 251/1MR</td>
</tr>
<tr>
<td>M825</td>
<td>1988</td>
<td>Zaire</td>
<td>Institut Pasteur, Paris, France</td>
<td>Zaire1</td>
</tr>
<tr>
<td>M799</td>
<td>1989</td>
<td>Hong Kong</td>
<td>University of Hong Kong</td>
<td>In21</td>
</tr>
<tr>
<td>M826</td>
<td>1990</td>
<td>Malawi</td>
<td>Institut Pasteur, Paris, France</td>
<td>Bakala Malenge</td>
</tr>
<tr>
<td>M654</td>
<td>1991</td>
<td>India</td>
<td>National Institute of Cholera and Enteric Diseases, Calcutta, India</td>
<td>413/91</td>
</tr>
<tr>
<td>M662</td>
<td>1993</td>
<td>Indonesia (Bali)</td>
<td>State Health Laboratory, Perth, Australia</td>
<td>7340</td>
</tr>
</tbody>
</table>
**Table 2. Primers used**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Gene or target</th>
<th>Sequence*</th>
<th>GenBank accession no. or source</th>
</tr>
</thead>
<tbody>
<tr>
<td>602</td>
<td>Suppression PCR adaptor</td>
<td>gagatgaacacctgtagttgccggccgcccggcaggt</td>
<td></td>
</tr>
<tr>
<td>604</td>
<td>Suppression PCR outer primer</td>
<td>gagatgaacacctgtagttgccggcaggt</td>
<td>X74696, X74697</td>
</tr>
<tr>
<td>605</td>
<td>Suppression PCR inner primer</td>
<td>M13R-ctaatgcggtcgagcggcaggtcggcaggt</td>
<td></td>
</tr>
<tr>
<td>524</td>
<td>16S</td>
<td>tcgactttgctaggttggcaggtcggcaggt</td>
<td>X74696, X74697</td>
</tr>
<tr>
<td>588</td>
<td>16S</td>
<td>M13F-gacagggatcaaaaccttgcggagtcgctctttgctaggttggcaggtcggcaggt</td>
<td>X74696, X74697</td>
</tr>
<tr>
<td>666</td>
<td>16S</td>
<td>M13F-tggcagccagagctgacgtagcgaattccttg</td>
<td>X74696, X74697</td>
</tr>
<tr>
<td>561</td>
<td>16S</td>
<td>aacacacgacagcgaacagtccatttctggcaggtcggcaggt</td>
<td></td>
</tr>
<tr>
<td>562</td>
<td>16S</td>
<td>M13R-tctgactgcccaggcatcc</td>
<td></td>
</tr>
<tr>
<td>542</td>
<td>23S</td>
<td>tagcgaataattccttg</td>
<td>X67300, X70371, X06485, Z35330, X81578, X63425, K00637, U09611</td>
</tr>
<tr>
<td>543</td>
<td>23S</td>
<td>ctttagatgcttccgctttcagc</td>
<td>As primer 542</td>
</tr>
<tr>
<td>634</td>
<td>5S</td>
<td>agtggagatagacactgcctttcagcaggttggcaggtcggcaggt</td>
<td>X02239, X02240</td>
</tr>
<tr>
<td>673</td>
<td>Operon-C specific</td>
<td>agttactttcctggcaggttggcaggtcggcaggt</td>
<td></td>
</tr>
<tr>
<td>674</td>
<td>Operon-B specific</td>
<td>aacacacgacagcgaacagtccatttctggcaggtcggcaggt</td>
<td></td>
</tr>
<tr>
<td>740</td>
<td>Operon-G specific</td>
<td>gagaacacgacagcgaacagtccatttctggcaggtcggcaggt</td>
<td></td>
</tr>
<tr>
<td>736</td>
<td>23S</td>
<td>M13R-tgcagctggccaggctcctttcagcaggtcggcaggt</td>
<td>U10956, VVU10951</td>
</tr>
<tr>
<td>841</td>
<td>Operon-D specific</td>
<td>cagctctcttcatgtagcgaacagtccatttctggcaggtcggcaggt</td>
<td>This study</td>
</tr>
<tr>
<td>842</td>
<td>Operon-E specific</td>
<td>ctctctcttcatgtagcgaacagtccatttctggcaggtcggcaggt</td>
<td></td>
</tr>
</tbody>
</table>

* M13F and M13R indicate that M13 forward or M13 reverse sequences were attached to those primers to facilitate sequencing.

**Fig. 1.** Schematic representation of primer positions in the operon region. Arrows indicate the orientation of primers. The intergenic spacer between 16S and 23S genes is drawn with a dashed line indicating that sizes vary with the four spacers found as reported in this study.

representation of primer position and orientation is depicted in Fig. 1 to assist in interpretation of results.

To obtain DNA of the regions flanking *rrn* genes, we used suppression PCR, which is a method whereby an adaptor is first attached to the end of a restriction fragment, enabling one to walk from known sequence to its flanking regions using the adaptor sequence for one of the PCR primers in conjunction with, in this case, a 16S RNA gene specific primer. This was done essentially as by Siebert *et al.* (1995). Adaptors and adaptor primers are shown in Table 2. Adaptor-ligated DNA was prepared as follows. Chromosomal DNA (2 pg) was digested with 20 U restriction enzyme for a minimum of 5 h at 37 °C. For enzymes generating overhanging ends, 1 μl 10 mM dNTPs and 0.5 μl Klenow fragment were added after digestion to fill the ends. The DNA was then extracted by phenol/ chloroform/isoamyl alcohol, precipitated and dissolved in 10 μl TE. One microlitre was ligated with an excess of adaptor overnight at 25 °C under the following conditions: 1 μl adaptor, 15% PEG, 50 mM Tris/HCl (pH 7.6), 10 mM MgCl₂, 0.5 mM ATP, 10 mM DTT and 2 U T4 DNA ligase in a total volume of 20 μl. The ligation was terminated by heat inactivation at 70 °C for 5 min. The ligated DNA was then diluted 10-fold. One microlitre was used for PCR in a reaction volume of 50 μl.

PCR product for sequencing was purified by using the Wizard PCR purification system (Promega) to remove excess PCR primers, and eluted in 30 μl H₂O. Sequencing was by the Dye-labelling primer technique using a thermal cycler (Perkin-Elmer Cetus) and an automated 377 DNA sequencer (Applied Biosystems) through the Sydney University and Prince Alfred Hospital Macromolecular Analysis Centre.

**Computer analysis.** DNA sequences were analysed by use of programs from the Australian National Genomic Information Service at the University of Sydney. Phylogenetic trees were constructed using the computer program PAUP (D. L. Swofford, phylogenetic analysis using parsimony, version 3.0).

**RESULTS**

**Number of *rrn* operons**

The number of *rrn* operons in bacteria varies. There are seven operons in *Escherichia coli* and *Salmonella enterica*. Recently, Majumder *et al.* (1996) mapped
R. LAN and P. R. REEVES

seven operons to the genome of the sixth pandemic V. cholerae strain 569B. However, ribotyping in V. cholerae detects 10 or more bands in BglII digests in seventh pandemic isolates (Karaolis et al., 1994) and it was not clear how to assemble seven operons from the patterns. To determine the number of operons, two probes were made specific to the 5'-region of the 16s RNA gene and the 3'-region of the 23s RNA gene corresponding to base positions 21-1079 and 1931-2759, by PCR amplification using primer pairs 561-562 and 542-543, respectively. Apart from differences in band intensity our 16s-specific DNA probe produced the same BglII pattern (Fig. 2a) with strain M803 as did ribotyping using 16s plus 23s rRNA as probe (Karaolis et al., 1994). All bands were of similar intensity in 16s DNA probing while signal intensities for 7.2, 4.6 and 2.3 kb fragments were much lower in the rRNA probing. Note that we recalibrated the size of each fragment reported in our previous study (Karaolis et al., 1994). The new and old sizes (kb) are as follows with old sizes in parentheses: 19.1 (19.8); 12.3 (13.1); 10.7 (11.3); 7.2 (7.9); 6.9 (7.6); 6.3 (6.9); 6.2 (6.8); 5.9 (6.5); 5.8 (6.3); 4.6 (4.9); 4.3 (4.5); 4.1 (4.3); 4.0 (4.1); 2.4 (2.1); 2.3 (2.0).

23s-specific DNA probing gave a similar pattern to that found previously with 16s DNA probing but the 7.2, 4.6 and 2.3 kb bands were not present (Fig. 2b). We concluded that these fragments hybridize only to 16s probe DNA and that there is an internal BglII site in three rRNA clusters. The BglII site was located in the region covered by the 16s probe as shown by BglII digestion of 16s PCR product (oligonucleotides 561 and 562, data not shown). Hybridization using oligonucleotide 561 (5'-end of the 16s RNA gene) revealed that the 4.0 and 4.1 kb bands did not hybridize to the oligonucleotide but did hybridize to the 16s probe (Fig. 2b, c). Thus these fragments contain the downstream part of the three operons. Tentatively it seems that there are nine operons because nine fragments hybridized to oligonucleotide 561.

As there is a unique I-CeuI site in the 23s RNA gene in bacteria (Liu & Sanderson, 1995a), we also used this to examine how many 23s RNA genes there are using a 1 kb probe extending from the I-CeuI site almost to the 3'-end of the 23s RNA gene (Fig. 1). Chromosomal DNA of strain M803 was digested with I-CeuI, then with BglII, KpnI, NsiI or PvuI, and probed by the 23s-specific probe (Fig. 2d). I-CeuI and BglII double digests showed only a 1 kb band, indicating there is a BglII site in all rRNA operons probably near the end of the 23s RNA gene. The presence of the BglII site was confirmed by sequencing and was found to be at 20 bp from the end of the 23s RNA gene. The KpnI, NsiI and PvuI double digests showed nine distinctive bands in each case. Each of these is larger than 1 kb, the size of the 23s probe, which indicates that there are no internal sites for KpnI, NsiI or PvuI in the last kilobase of the 23s RNA gene. This confirms the suggestion that there are nine rRNA operons.

We conclude that there are nine operons in strain M803. We assign operons A–G to the seven operons with BglII fragments of 107, 7.2, 6.9, 6.3, 6.2, 4.6 and 2.3 kb and assign H and I to operons giving 5.9 and 5.8 kb fragments, respectively, using letters to denote an operon as in E. coli, S. enterica and Bacillus subtilis. Operons B, F and G have a BglII site in their 16s RNA gene. We
were cloned and sequenced. The sequences were aligned
needed sequences of DNA upstream of 16s RNA to
PanBL7, which hybridized to every operon with a very
genes. Chromosomal DNA was digested separately with
524 from the 5'-end of the 16s RNA gene was used in
to each digest. For the first round amplification, primer
were obtained from the BgZI digest. The five fragments
detected by ribotyping and each of the four
fragments of convenient size and separation
flanking sequences could be related to one of
fragments reported in M803 (kb) flanking region variation in
mM803 (kb) flanking region variation in
Operon I
Operon E
Operon D
Operon C
Operon B
Operon A

PanBL7, PanKL2
PanBL7, PanKL1
PanBL4
PanBL3, PA50
PAF13
Not obtained

Ribotype
G
I
J
M
Q

Fig. 3. Schematic representation of each BglI site variation in
the seventh pandemic clone. The operons are indicated to the
left for each band carrying the 5' 16s gene. Alternative bands
same operon are indicated by a dotted line for operons
B, C and G. The size of each band is indicated on the left.
The main objective of this study was to reveal the
changes behind ribotype variation. Due to the presence of a BglI site at the 3'-end of the 23s RNA genes all BglI
ribotype variation involves the rrn genes or the upstream
flanking regions as changes in the downstream flanking
DNA are not detected in ribotyping. Therefore we
concentrated our effort on the upstream DNA.
To characterize the variation detected by ribotyping, it
is necessary to study individual operons. To do this we
needed sequences of DNA upstream of 16S RNA to
facilitate operon-specific PCR. This would also enable
us to walk out to various restriction sites in the flanking
DNA. Suppression PCR (Siebert et al., 1995) was used to
obtain sequence immediately upstream of the 16S RNA
genes. Chromosomal DNA was digested separately with
several 6 base restriction enzymes (AflIII, BglI, KpnI,
NcoI, PvuII, SalI and SnaBI) and adaptors were ligated
to each digest. For the first round amplification, primer
524 from the 5'-end of the 16S RNA gene was used in
combination with suppression PCR primer 604 corresponding to the outermost segment of each adaptor.
Amplified fragments were isolated from agarose gels and reamplified using nested primer pair 588–605. Five
distinct fragments of convenient size and separation
were obtained from the BglI digest. The five fragments
were cloned and sequenced. The sequences were aligned
and showed considerable conservation for 300 bp up-
stream of the 16S RNA gene. Primers were synthesized for
each flanking region using sequences upstream of the
300 bp common DNA, and used with primer 588 to
amplify five PCR products which were used to probe
BglI digests: four of them hybridized with one of the
fragments detected by ribotyping and each of the four
flanking sequences could be related to one of rrn operons
C, D, E and G (Table 3). The remaining one was
PanBL7, which hybridized to every operon with a very
faint signal: we subsequently isolated a longer fragment,
PanKL1, from a KpnI enzyme digest and sequenced it.
Surprisingly the upstream sequence of PanKL1 matches the 3'-end of the 23S RNA gene. Apparently PanKL1 is
derived from two operons in tandem. However, from
this information alone, it was not possible to work out
to which operon PanBL7 belonged. As described later,
PanBL7 was subsequently deduced to be the segment
upstream of operon H or I and there are either two loci
with two operons in tandem or one locus with three
operons in tandem. The flanking region of operon B
(Table 3) was obtained from an AflIII digest by a similar
procedure. Flanking DNA of operons A and F was not
obtained and was not needed for this study as no ribotype variation was detected in these two
operons.

### Analysis of BglI ribotyping variation

With relevant flanking sequences available, BglI ribotype
variation could be analysed in detail. Operons D and E
do not vary in the seventh pandemic and thus were not
investigated further. The schematic representation of
each BglI change is shown in Fig. 3 for ease of
interpretation of the following results. Note that in the
schematic diagram of our previous paper the 4-0 kb
fragment reported for M811 (ribotype I) and M812
(ribotype J) (Karaozis et al., 1994) was in fact absent.

#### (i) Variation in operon B

In our previous study (Karaozis et al., 1994), 7-2 and 12-3 kb bands in M803 and M811,
Table 4. Spacer types and status of BglI site in the 16S RNA gene for operons B, C and G and status of operons H and I in type strains of each ribotype

The spacer types are small (S), 431 bp; medium (M1 and M2), 509 and 607 bp; and large (L), 711 bp. A status identical to ribotype G is indicated by a dot. Del, deletion of an operon.

<table>
<thead>
<tr>
<th>Ribotype</th>
<th>Strain</th>
<th>Operon B</th>
<th>Operon C</th>
<th>Operon G</th>
<th>Operon H</th>
<th>Operon I</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>M803</td>
<td>S</td>
<td>+</td>
<td>M1</td>
<td>+</td>
<td>Present</td>
</tr>
<tr>
<td>H</td>
<td>M807</td>
<td>.</td>
<td>.</td>
<td>M1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>M811</td>
<td>L</td>
<td>.</td>
<td>M1</td>
<td>+</td>
<td>Present</td>
</tr>
<tr>
<td>J</td>
<td>M812</td>
<td>L</td>
<td>.</td>
<td>M1</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>K</td>
<td>M813</td>
<td>.</td>
<td>S</td>
<td>M1</td>
<td>+</td>
<td>Present</td>
</tr>
<tr>
<td>L</td>
<td>M820</td>
<td>L</td>
<td>.</td>
<td>M1</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>M825</td>
<td>L</td>
<td>M1</td>
<td>+</td>
<td>S</td>
<td>Del</td>
</tr>
<tr>
<td>N</td>
<td>M799</td>
<td>M2</td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>O</td>
<td>M826</td>
<td>L</td>
<td>.</td>
<td></td>
<td>.</td>
<td>Del</td>
</tr>
<tr>
<td>P</td>
<td>M654</td>
<td>.</td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>M662</td>
<td>M2</td>
<td>.</td>
<td></td>
<td>.</td>
<td></td>
</tr>
</tbody>
</table>

* + / −, Presence or absence of BglI site at base 838 of the 16S RNA gene.
† + / −, Presence or absence of a BglI site in the 16S RNA gene proximal flanking region of operon G.

respectively, were deduced to be alternative fragments of the same operon. Use of the operon B probe confirmed this and also showed that the 19-1 kb band in M662 represents another variant. These alternative forms result from loss of the BglI site in the 16S RNA gene. We sequenced the region around the BglI site in M803, M811 and M662. A single base substitution of T for C at position 849 (E. coli K-12 numbering) altered the 11 bp BglI recognition sequence GCCCTAGAGCT in M803 to GCCCTAGAGCT in M811 and M662. However, the change to 19-1 kb in M662 from 7-2 kb in M803 or 12-3 kb in M811 also involves a change at about 7 kb upstream of the operon. We are yet to walk to that BglI site.

(ii) Variation in operon C. Ribotyping showed that the 6-9 and 24 kb bands were alternative fragments of the operon previously identified by BglI site 2 (Karaolis et al., 1994). The 24 kb variant only occurs in two ribotypes. Use of the operon-C-specific probe confirmed that the 2-4 kb band in M825 is indeed a replacement of the 6-9 kb fragment. To create such a small restriction fragment a new site must be created within the 16S RNA gene. This was confirmed by BglI digestion of the PCR product of the 16S RNA gene of operon C which was amplified using primer pair 673-562. Primer 673 is based on the PanBL3 sequence and is unique to the flanking region of operon C.

(iii) Variation in operon G. The 2-3 kb band, loss of which defined BglI site 7 (Karaolis et al., 1994), is absent only in M662 but there is no replacement band. Use of the operon-G-specific probe showed that there was a variant band in M662 similar in size to the 6-9 kb operon C fragment. The newly created band in M662 was masked by the 6-9 kb operon C band in ribotyping. Digestion of the PCR product amplified using primer pair 740-562 showed that the BglI site in the operon C 16S RNA gene was absent in M662.

(iv) Spacer between 16S RNA and 23S RNA genes. The 4-0, 4-1 and 4-3 kb BglI fragments are internal rrr bands containing part of the 16S RNA gene and most of the 23S RNA gene of the three rrr operons. It was not possible to consistently assign the three fragments to specific operons in the 11 ribotypes. As 16S and 23S RNA genes do not vary in length within a species, the variation in size must be in the intergenic spacer between the 16S and 23S RNA genes. The spacer region was amplified by nested PCR using operon-specific primers (673, 674 or 740) with 23S primer 736, then primer pair 666 and 736. Three types of spacers were initially found with lengths of 431, 509 and 711 bp. The three internal bands (4-3, 4-1 and 4-0 kb) correspond to spacers of the three different sizes. We later identified another spacer variant of 607 bp present in operon B of M662 and M799. The spacers are referred to as small (S; 431 bp), medium (M1 and M2; 509 or 607 bp) and large (L; 711 bp).

The presence or absence of the 4-3 and 4-0 kb bands varies from strain to strain and the variation must be generated by change of spacers associated with the position 838 BglI site. For operons B and C, internal bands could not be consistently assigned to a given operon from Southern probing results. To further characterize the distribution of spacer forms, the spacer for each of the varying operons B, C and G was amplified by nested PCR. The first PCR was using operon-specific primers 674, 673 and 740, respectively, with primer 736, with the second PCR using primer pair 666-736. The results are shown in Table 4. It is evident that there was extensive reassortment of spacers in
which is identical to the larger spacer of M803, and one spacer in M825 similar in size to the smaller spacer of M803, while no fragment was amplified from M826.

**DISCUSSION**

Several studies on *V. cholerae* using ribotyping have produced useful data on the evolution of the seventh pandemic clone (Evins et al., 1993; Faruque et al., 1994, 1995; Karaolis et al., 1994; Koblavi et al., 1990; Popovic et al., 1993; Wachsmuth et al., 1993). We have analysed the changes which created the ribotype variation in the seventh pandemic clone revealed by *BglII* restriction and found that it was largely due to changes at a *BglII* site within the 16S RNA gene or changes in the spacer region. These changes are consistent with an origin by recombination between different *rrn* operons on the same chromosome. Replacement of the spacer must be due to recombination between *rrn* clusters, and that is also the best explanation for the gain or loss of a *BglII* site in the 16S RNA gene, which entails a single base change, as it would account for the much higher level of change in this *BglII* site than in *BglII* sites in the flanking regions. This also explains why *BglII* is the most discriminatory enzyme among the seven enzymes used by Popovic et al. (1993) and the four enzymes used in our previous study (Karaolis et al., 1994). In our previous study, both *HindIII* and *Hpal* showed a single ribotype among the seventh pandemic isolates. It was not reported why *BglII* was chosen in the first ribotyping study of *V. cholerae*, which was done by Koblavi et al. (1990).

As can be seen from Table 4, there are eight situations where the combination of spacer types, S, M1, M2 or L, and the presence or absence, indicated by + or −, of the *BglII* site in the 16S RNA gene differs from that in our earliest isolates. There are four combinations in operon B (S−, L+, L− and M2−) generated from the S+ form of the earliest strains. Operons C and G each have two additional forms. That indicates a minimum of eight recombination events, and more if changes in spacer and *BglII* status are due to independent events. The losses of operons H and I are also presumably due to recombination within an operon pair. The simplest explanation for the origin of each of these 10 new forms is recombination between two of the *rrn* operons on the *V. cholerae* chromosome. We cannot exclude mutation for the gain or loss of the *BglII* site within the 16S RNA gene or recombination involving two cells but consider these to be much less probable. The overall level of change is extensive considering that the seventh pandemic began in 1961 and the strains were isolated from 1961 to 1993. In the 1961 isolate there are three operons having the *BglII* site and six (including operons H and I) having none. Only recombination between operons which differ in this regard causes change of restriction pattern. Thus 18 of the 36 possible recombination events can alter the *BglII* restriction pattern. Typing of the spacers revealed several additional *rrn* recombination events which do not change the *BglII* restriction pattern. Further exploration of minor sequence variation be-

---

**Fig. 4.** (a) Probing of I-Ceul digests by a 235 probe. Lanes 1-4 are M803, M812, M825 and M826, respectively. The size of each band is indicated on the left. (b) PCR amplification of the intergenic spacer of operons H and I using primers 634-736 and 666-736. Lanes 2-5 are M803, M812, M825 and M826, respectively. Lane 1 is the SPP1 EcoRI molecular mass marker.
between operons and spacer sizes in all operons may reveal more events and 10 is a minimum number of events in the ancestry of the strains studied. It is not known why operons B, C and G are involved more frequently in rrrn recombination while no recombination events were observed in operons A, D, E and F. It is possible that operons H and I are associated with the former group and stimulate the recombination.

Although rrrn recombination and deletion have been observed to occur under laboratory conditions in S. enterica Typhimurium and E. coli (Hill et al., 1990), only recently has rrrn recombination been detected in natural isolates in S. enterica Typhi and Paratyphi by macrorestriction mapping using the I-CeuI enzyme (Liu & Sanderson, 1995a, b, c, 1996). They found variants which have arisen by recombination between rrrn loci resulting in rearrangement of the seven I-CeuI DNA fragments. Our study shows that the level of rrrn variation in the seventh pandemic clone is comparable to that observed in S. enterica Typhi (Liu & Sanderson, 1996), but it is important to note that all the variation within the seventh pandemic V. cholerae must have arisen since 1961, whereas there is no time frame available for generation of the variation observed in S. enterica Typhi.

The method used by Liu & Sanderson (1996) on S. enterica Typhi detects rearrangements due to recombination between rrrn operons but would not reveal rrrn sequence variation of the type we have observed, and likewise the work we have done would not reveal inversion of large segments of the chromosome due to reciprocal recombination between two rrrn operons. Reciprocal recombination events of the type observed by Liu & Sanderson (1996) could also generate non-reciprocal events in the region of reciprocal recombination, as reciprocal recombination involves single-strand invasion and opportunity for local gene conversion (Kowalczykowski et al., 1994) in addition to reciprocal recombination between distant loci. However, each of the changes we observed in the BglI site was non-reciprocal with reference to M803 (ribotype G) and the simplest explanation is that we are observing non-reciprocal changes of the type normally envisaged for concerted evolution. This explanation is supported by the observation that in a model system for Salmonella (Segall & Roth, 1994), reciprocal recombination involving a 5 kb duplicated segment is only rarely accompanied by gene conversion. We suggest that our observations are an example of concerted evolution in action. The recombination events were detected by ribotyping because minor sequence variation in the 16S RNA gene gives the presence or absence of a BglI recognition site.

The seventh pandemic is caused by a single clone which emerged in 1961, and the first date of isolation for each new ribotype reflects the latest time for emergence of newly developed variation created by rrrn recombination. This information provided us with a unique opportunity to estimate the dynamics of concerted evolution in the rrrn gene clusters. Nine new ribotypes (H–P) were found among 47 isolates sampled over the 33 year period of study. The isolates are quite evenly distributed in time, ranging from 5 to 10 isolates per 5 year period with a mean of 6.7. This means 0.0058 recombinants per isolate per year. It is difficult to accurately translate this to recombination rate as we do not have a good estimate of the generation time for V. cholerae. For the commensal E. coli which lives relatively stably in the intestines of humans and animals, an estimate of 200 generations per year was used for population genetic studies (Milman & Bridges, 1990). If we apply the same generation time to V. cholerae we get a minimum rrrn recombination rate of 2.9 x 10^{-5} per cell per generation. This rate is 1000 times higher than the rate of gene conversion estimated for the duplicated tuf genes in S. enterica Typhimurium, which is in the order of 2 x 10^{-8} per cell per generation under laboratory conditions (Abdulkarim & Hughes, 1996). However, the generation time for V. cholerae must be much higher than 200 per year during infection, although possibly lower when in the environment. In any case, the gene conversion rate in rrrn genes of V. cholerae could not be as low as the rate for the tuf genes as the generation time could not be 1000 times higher than the above assumption. Nevertheless, for the first time to our knowledge we have provided an indication of the rate of concerted evolution in bacteria.

The presence of tandem repeat operons and occasional deletion of one of the pair is similar to the reported situation in B. subtilis (Widom et al., 1988). The genome of the widely used B. subtilis strain 168 has 10 rrrn operons in seven loci with one set each of operon doublet and triplet. Deletion of an rrrn operon is observed to occur quite often, presumably by intrachromosomal recombination within the tandemly repeated sets. It remains to be determined for V. cholerae whether operons H and I are present separately as doublets with two other operons or together as a triplet with one other operon.

In conclusion, rrrn recombination between loci has generated tremendous variation in the seventh pandemic clone of V. cholerae and all but one of the ribotype variants observed are thought to be due to such recombination. It is apparent that ribotyping detects two distinctive types of change: flanking region changes and changes within rrrn operons arising from recombination between operons within a single chromosome. Ribotype variation within the seventh pandemic clone is predominantly due to recombination. The frequency of recombination between rrrn operons is very high, in the order of 10^{-5} per cell per generation. This means that any variation could undergo precise reversion by the same recombination event within the time frame covered by the study and there will also be high levels of parallel changes due to recombination. Hence a similar ribotype does not necessarily reflect similarity by descent from a common ancestor. Ribotype variation in other species may also be due to rrrn recombination, which must be taken into account in the use of ribotyping for evol-
utionary studies and monitoring of pathogenic or other clones. Alternative methods such as AFLP (Janssen et al., 1996; Vos et al., 1995) may be more appropriate for evolutionary studies or even strain identification where ribotype variation appears to be due to recombination within rrr.

ACKNOWLEDGEMENTS

We thank all the institutions listed in Table 1 that kindly supplied us with isolates.

This work was supported by a grant from the Australian National Health and Medical Research Council.

REFERENCES


Received 30 September 1997; revised 23 January 1998; accepted 28 January 1998.