

Molecular Basis of Ribotype Variation in the Seventh Pandemic Clone and its O139 Variant of *Vibrio cholerae*

Ruiting Lan, Peter R Reeves/⁺

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

*Ribotyping has been widely used to characterise the seventh pandemic clone including South American and O139 variants which appeared in 1991 and 1992 respectively. To reveal the molecular basis of ribotype variation we analysed the *rrn* operons and their flanking regions. All but one variation detected by *Bgl*I, the most discriminatory enzyme, was found to be due to changes within the *rrn* operons, resulting from recombination between operons. The recombinants are detected because of the presence of a *Bgl*I site in the 16S gene in three of the nine *rrn* operons and/or changes of intergenic spacer types of which four variants were identified. As the frequency of *rrn* recombination is high, ribotyping becomes a less useful tool for evolutionary studies and long term monitoring of the pathogenic clones of *Vibrio cholerae* as variation could undergo precise reversion by the same recombination event.*

Key words: ribotyping - *rrn* recombination - seventh pandemic - O139 - *Vibrio cholerae*

Vibrio cholerae is the agent responsible for cholera which was first described in 1854. However, the natural habitat of *V. cholerae* is the aquatic environment. Environmental *V. cholerae* are diverse and most are nontoxicogenic. There are more than 190 O antigens identified (Yamai et al. 1997). The best known forms are O1 and recently O139. Both cause cholera, currently mostly in developing countries. Seven pandemics of cholera are recognised since 1817. The seventh started in 1961 and continues to the present day.

The seventh pandemic clone has been studied by various molecular methods. Ribotyping, a form of restriction fragment length polymorphism analysis using rRNA genes, was shown to be very discriminatory in revealing variation. rRNA sequences are highly conserved and the genes (*rrn*) are present as multiple copies in the genome of many bacteria. The typing therefore provides information on several flanking regions simultaneously. There have been quite a few studies on the epidemiology and molecular typing of *V. cholerae*. Koblavi et al. (1990) were the first to employ ribotyping to fingerprint *V. cholerae* strains and Popovic et al. (1993) proposed a standardised scheme for typing *V. cholera* using *Bgl*I restriction enzyme to allow public laboratories to follow the movement and

identify the origins of *V. cholerae* strains. Ribotyping has since been widely used to characterise the sixth pandemic clone (Faruque et al. 1993) and the seventh pandemic clone (Karaolis et al. 1994, Faruque et al. 1995), the South American and O139 variants (Wachsmuth et al. 1993, Popovic et al. 1995, Dalsgaard et al. 1997) and other O1 outbreaks (Coelho et al. 1995). Ribotype variation has been found in the South American isolates in the four years of its spread from 1991 (Dalsgaard et al. 1997) and in the O139 variant which appeared in 1992 (Popovic et al. 1995, Faruque et al. 1997).

As the *rrn* operons are conserved, variation detected in ribotyping has been generally assumed to be due to variation in flanking regions. We studied the changes behind ribotype variation to help us understand the evolution of pandemic clones (Lan & Reeves 1998).

RIBOTYPE VARIATION IN THE SEVENTH PANDEMIC CLONE

The seventh pandemic clone is an very interesting clone to study because accurate dates of development are known. In our ribotyping study with a total of 47 strains isolated from 1961 (Indonesia and Hong Kong) to 1993 including African and Asian isolates (Karaolis et al. 1994). *Bgl*I detects most of the polymorphisms. The *Bgl*I ribotypes are summarised in Table I. There are 11 ribotypes. Ribotype G was present at the start of the pandemic and 14 type G strains are all from Asia. In the first 10 years of spread there was no detectable variation. In fact after 1966 there was a lull period with relatively little cholera. Ribotype H appeared in 1970 in Asia and spread to Africa;

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

⁺Corresponding author. Fax: +612 9351.4571. E-mail: reeves@angis.usyd.edu.au

Received 15 June 1998

Accepted 30 July 1998

Africa being free from cholera before 1970. The resurgence in 1970 seems to be associated with a substantial increase in genetic variation. There are other minor ribotypes with one or two isolates from Asia and Africa.

TABLE I

Summary of *Bgl*I ribotypes detected in the seventh pandemic clone (Karaolis et al. 1994)

Ribotype	No. isolates	Year of isolation	Location of isolation
G	14	1961-1991	Asian region only
H	22	1970-1993	Asia and Africa
I	1	1971	Burma
J	2	1971, 1974	Chad
K	1	1972	Senegal
L	1	1978	Malaysia
M	1	1988	Zaire
N	2	1989, 1992	HK & Indonesia
O	1	1990	Malawi
P	1	1991	India
Q	1	1993	Indonesia

rrn OPERON FEATURES IN THE EARLY SEVENTH PANDEMIC ISOLATE M803

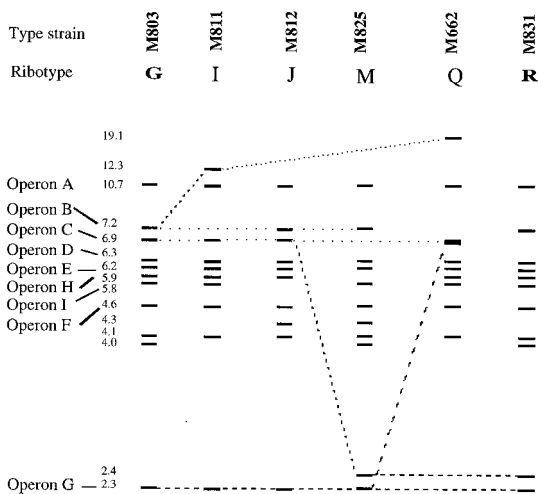
Majumder et al. (1996) mapped seven operons to the genome of *V. cholerae* strain 569B. However, ribotyping in *V. cholerae* detects 10 or more bands in *Bgl*I digests in the seventh pandemic isolates (Karaolis et al. 1994) and it was not clear how to assemble seven operons from the patterns. We know that in general a *rrn* operon has three genes in the order 16S-23S-5S and there is a unique *I-Ceu* I site in the 23S gene (Liu & Sanderson 1995). We used 2 DNA probes, one specific to the 5' region of the 16S gene and the other specific to the 3' region of the 23S gene, and an oligo probe from 16S 5' end to probe Southern blots of *Bgl*I digests. From the probings we concluded that there are nine operons in the strain of M803, an early seventh pandemic isolate. Three operons have a *Bgl*I in their 16S gene. We also used the 23S probe to probe strain M803 DNA digested with *I-Ceu* I and *Bgl*I, which showed only a strong 1 kb band rather nine bands as expected and thereby discovered that there is a *Bgl*I site near the end of the 23S gene. However, the number of operons is different from the seven operons detected by Majumder et al. (1996). Later we discovered that two of the nine operons are in tandem with another operon. Thus there are only seven loci on the chromosome. A recent study using pulsed-field gel electrophoresis of *I-ceu* I digest from the Majumder group (Nandi et al. 1997) found that there are nine fragments, one of which is 6 kb, the equivalent of a

tandem operon. However, we could not reconcile our data with their conclusion of eight independent *rrn* loci.

IDENTIFICATION OF ALTERNATIVE FRAGMENTS FOR EACH RIBOTYPE CHANGES BY OPERON SPECIFIC PROBING

Due to the presence of a *Bgl*I site at the end of the 23S genes, changes in the 23S proximal DNA are not detected by ribotyping and all *Bgl*I variation involves the 16S genes or their flanking regions. PCR walking was used to obtain sequence immediately upstream of the 16S genes. We then made operon specific probes and probed *Bgl*I digests of different ribotypes to analyse the variation. The probing experiments are summarised in the Fig. with schematic representation of each *Bgl*I change.

For operons B, C, and G, alternative bands were identified through probing. For operon B, the alternative bands are 7.2 kb, 12.3 kb and 19.1 kb. We found that these alternative bands result from loss of the *Bgl*I site in the 16S gene. The 7.2 kb band has only 5' end of the 16S gene because of the *Bgl*I site in the 16S gene. We sequenced the region around the *Bgl*I site in M803, M811, and M662 of operon B. A single base difference from C to T at base 849 determines the presence or absence of the elevenmer *Bgl*I recognition sequence. However the change to 19.1 kb in M662 from 7.2 kb in M803 or 12.3 kb in M811 also involves



Schematic representation of *Bgl*I ribotype data of *Vibrio cholerae* seventh pandemic isolates to indicate variation in specific operons among ribotypes G, I, J, M and Q and the O139 variant R. The operons are indicated to the left for each band representing the 5' 16S gene. Alternative bands of the same operon are indicated by a dotted line for operons B, C, and G. The size of each band is indicated on the left. After Karaolis et al. (1994) and Lan and Reeves (1998).

change in the flanking region of the operon. We have yet to walk to that *BglII* site. For operon C operon-specific probing showed that the 2.4 kb band in M825 is a replacement of the 6.9 kb fragment which is due to the creation of a *BglII* site in the 16S gene. For operon G the alternative band for the 2.3 kb band is similar in size to the 6.9 kb operon C fragment. Apparently this resulted in the newly created band being masked. This change is due to the loss of *BglII* site in the 16S gene.

We found that operons H and I are tandem operons and absence of a band means deletion of the operon. We did Southern probing of I-*ceu I* digests to show that H and I are immediately downstream of other *rrn* operons. For tandem operons I-*Ceu I* digestion, which cuts uniquely in 23S gene, would release a fragment from the I-*Ceu I* site of the upstream operon to the I-*Ceu I* site of the downstream operon.

SIZE VARIATION IN INTERGENIC SPACERS

The bands of 4.0, 4.1 and 4.3 kb in the Fig. are fragments having part of 16S gene and the whole of 23S gene and variation in size of these bands is due to size differences of the intergenic spacer. We analysed the intergenic spacer of the three operons B, C and G. The spacer region was amplified by nested PCR. Three types of spacers were initially found: small (S) 411 bp, large (L) 692 bp, and medium (M1) 488 bp. We later identified another spacer variant (M2) 587 bp.

REASSORTMENT OF SPACER TYPES AND *rrn* RECOMBINATION

The nature of variation detected by *BglII* ribotyping is summarised in Table II, including eleven ribotypes of the seventh pandemic and its O139 variant. It is evident that there was extensive reassortment of spacers. For operon B, there are four forms: S, *BglII*-; L, *BglII*+; L, *BglII*-; and M2, *BglII*- generated from the S, *BglII*+ form of the early strains. Both operons C and G have two additional forms. So there are in total eight types of changes to *rrn* operons which we attribute to homologous recombination between operons. We don't know why the most frequent changes are in operon B. The other frequent change is deletion of operon I. Deletions might be due to single or multiple events. The majority of the ribotypes seems to have arisen from ribotype G by a single recombination event though M, N and Q require two or more recombinational changes.

RIBOTYPE VARIATION IN THE O139 VARIANT

It is well recognised from multilocus enzyme electrophoresis (MLEE) and other data that the O139 Bengal strain is derived from the seventh pandemic strain. A single ribotype, R, (represented by strain M831) was identified in our previous study (Karaolis et al. 1994). The O139 clone is very similar to the Asia-dominant ribotype G. The only difference in the *BglII* digests of ribotypes R and G is that in ribotype R the 6.9 kb operon C fragment

TABLE II

Spacer types and status of *BglII* site in the 16S RNA gene for operons B, C and G and status of operons H and I in the ribotypes (G-Q) of the seventh pandemic clone and ribotype R of the O139 variant

Ribotype	Strain	Operon B		Operon C		Operon G		Operon H	Operon I
		Spacer ^a	<i>BglII</i> site ^b	Spacer	<i>BglII</i> site	Spacer	<i>BglII</i> site		
G	M803	S	+	M1	-	M1	+	Present	
Present									
H	M807	. ^c	Del
I	M811	L	-
J	M812	L	Del
K	M813	.	.	S	+	.	.	.	del
L	M820	L	-	Del
M	M825	L	.	.	+	S	.	Del ^d	.
N	M799	M2	-	.	.	.	-	.	.
O	M826	L	Del	Del
P	M654	.	-	Del
Q	M662	M2	-	.	.	.	-	.	.
R	M831	.	.	.	+	.	.	.	Del

^a: the spacer types are small (S) 431bp, large (L) 711bp, medium (M1) 509bp and a variant of size between M1 and L (M2) 607bp; ^b: +/- indicates presence or absence of *BglII* site at base 838 of the 16S RNA gene; ^c: status identical to ribotype G is indicated by a dot; ^d: deletion of an operon.

is replaced by a 2.4 kb band due to the gain of a *Bgl*I site in the 16S gene. It is very interesting that Popovic *et al* (1995) detected two ribotypes, 3a and 5a, in O139 isolates. Ribotype 3a is identical to ribotype R. Ribotype 5a differs from ribotype 3a by the absence of operon I. Popovic *et al* (1995) noted that ribotypes 3a and 5a are very similar to their ribotypes 3 and 5 of seventh pandemic clone respectively, which implied that ribotype 3a may be derived from ribotype 3 and 5a derived from 5. Most likely ribotype 5a is derived from ribotype 3a through an independent operon I deletion rather than from ribotype 5. If 3a and 5a were derived separately, it would require independent transfer of a new O antigen gene cluster. This also illustrates the weakness of ribotyping for defining relationships between strains.

RIBOTYPE VARIATION IN THE SOUTH AMERICAN VARIANT

In our ribotyping study (Karaolis *et al.* 1994), only one South American isolate was included which belongs to ribotype H. Dalsgaard *et al.* (1997) analysed 50 South American isolates isolated in Peru from 1991 to 1995 and found four ribotypes: R1, R2, R3 and R4. The publication of good quality Southern blot by Dalsgaard *et al.* (1997) and use of probes identical to our study allow us to interpret their results in terms of the basis of ribotype changes. R1 is identical to our ribotype H. R2 differs from R1 by the appearance of a 4.3 kb band and disappearance of a 4.0 kb band representing the change in spacer type from small to large in operon C by *rrn* recombination. R3 differs from R1 in the presence of a 5.6kb band which is equivalent to the position of operon I and the pattern is identical to our ribotype G, the frequent ribotype isolated in Asian region only. It seems that a R1 (ribotype H pattern) strain reversed back to ribotype G pattern by recreation of the tandem operon I, presumably also due to *rrn* recombination. This is an good example where identical ribotypes may not mean genetic similarity. It is possible that the R3 strain is actually a ribotype G strain separately introduced into South America and R3 is not derived from R1. The early South American isolates differ at the locus of leucine aminopeptidase in MLEE from other seventh pandemic isolates. Therefore the allele profile could be used to determine whether R3 is derived from R1. R4 is very different from the other three ribotypes. In our opinion R4 is not developed from R1.

GENERAL CONCLUSIONS

It has always been assumed that ribotype variation is due to changes outside the *rrn* operons. This study shows that ribotyping detects two types of changes of very different nature: within operon

changes produced by *rrn* recombination or mutational changes outside the operons. For the seventh pandemic clone the variation observed is predominantly *rrn* recombination. The frequency of *rrn* recombination is high. Ten new ribotypes were found in 46 isolates over 33 year span in our previous study (Karaolis *et al.* 1994). Nine of the ten ribotypes were generated by *rrn* recombination. There are new ribotypes found in the South American isolates in its first four years spread (Dalsgaard *et al.* 1997). Similarly new ribotypes were found in the O139 variant (Popovic *et al.* 1995, Faruque *et al.* 1997). Therefore, the level of parallel and reversal changes will also be high. And similar ribotype does not necessarily reflect genetic similarity. Thus ribotyping is not suitable for long term monitoring of the seventh pandemic clone or any *V. cholerae* clones. Other species need to be studied to see whether this is a general phenomenon.

However, some laboratories may wish to continue to use *Bgl*I ribotyping for typing seventh pandemic isolates as an accessory tool as it has been well established technically. In such case we recommend using a fragment from the 16S gene encompassing the *Bgl*I site as a probe, eg the 1 kb fragment from base 21 to base 1097 of the 16S gene in this study, which will produce bands of uniform intensity and identical patterns to probing using a mixture of 16S and 23S rRNA or rDNA (Lan & Reeves 1998). Using a 16S and 23S rRNA as probe the hybridisation signal for those fragments having only part of the 16S gene is much weaker than for other fragments which may lead to misidentification of patterns. The use of probes including 5S gene or any flanking sequences in addition to 16S and 23S genes is not recommended as it produces ribotyping patterns not comparable to those of other laboratories. Separation of the four bands between 6.3 kb to 5.8 kb, and the three bands between 4.0 kb to 4.3 kb is usually poor. Reference ribotypes representing each band variation should be included for comparison to determine the present or absent of a band to increase accuracy.

REFERENCES

- Coelho A, Andrade JRC, Vincente ACP, Salles CA. 1995. New variant of *Vibrio cholerae* O1 from clinical isolates in Amazonia. *J Clin Microbiol* 33: 114-118.
- Dalsgaard A, Skov MN, Serichantalergs O, Echeverria P, Meza R, Taylor DN 1997. Molecular evolution of *Vibrio cholerae* O1 strains isolated in Lima, Peru, from 1991 to 1995. *J Clin Microbiol* 35: 1151-1156.
- Faruque SM, Ahmed KM, Siddique AK, Zaman K, Alim ARMA, Albert MJ 1997. Molecular analysis of toxigenic *Vibrio cholerae* O139 bengal strains isolated in Bangladesh between 1993 and 1996: evidence for emergence of a new clone of the bengal Vibrios. *J Clin Microbiol* 35: 2299-2306.

- Faruque SM, Alim ARMA, Rahman MM, Siddique AK, Sack RB, Albert MJ 1993. Clonal relationships among Classical *Vibrio cholerae* O1 strains isolated between 1961 and 1992 in Bangladesh. *J Clin Microbiol* 31: 2513-2516.
- Faruque SM, Roy SK, Alim ARMA, Siddique AK, Albert MJ 1995. Molecular epidemiology of toxigenic *Vibrio cholerae* in Bangladesh studies by numerical analysis of rRNA gene restriction patterns. *J Clin Microbiol* 33: 2833-2838.
- Karaolis DKR, Lan R, Reeves PR 1994. Molecular evolution of the 7th pandemic clone of *Vibrio cholerae* and its relationship to other pandemic and epidemic *V. cholerae* isolates. *J Bacteriol* 176: 6199-6206.
- Koblavi S, Grimont F, Grimont PAD 1990. Clonal diversity of *Vibrio cholerae* O1 evidenced by rRNA gene restriction patterns. *Res Microbiol* 141: 645-657.
- Lan R, Reeves PR 1998. Recombination between rRNA operons created most of the ribotype variation observed in the seventh pandemic clone of *Vibrio cholerae*. *Microbiology* 144: 1213-1221.
- Liu S-L, Sanderson KE 1995. Rearrangements in the genome of the bacterium *Salmonella typhi*. *Proc Natl Acad Sci USA* 92: 1018-1022.
- Majumder R, Sengupta S, Khetawat G, Bhadra R, K, Roychoudhury S, Das J 1996. Physical map of the genome of *Vibrio cholerae* 569B and localization of genetic markers. *J Bacteriol* 178: 1105-1112.
- Nandi S, Khetawat G, Sengupta S, Majumder R, Kar S, Bhadra RK, Roychoudhury S, Das J 1997. Rearrangements in the genome of *Vibrio cholerae* strains belonging to different serovars and biovars. *Int J Syst Bacteriol* 47: 858-862.
- Popovic T, Bopp CA, Olsvik O, Wachsmuth K 1993. Epidemiologic application of a standardized ribotype scheme for *V. cholerae* O1. *J Clin Microbiol* 31: 2474-2482.
- Popovic T, Fields PI, Olsvik O, Wells JG, Evins GM, Cameron DN, Farmer III JJ, Bopp CA, Wachsmuth K, Sack RB, Albert MJ, Nair GB, Shimada T, Feeley JC 1995. Molecular subtyping of toxigenic *Vibrio cholerae* O139 causing epidemic cholera in India and Bangladesh, 1992-1993. *J Infect Dis* 171: 122-127.
- Wachsmuth IK, Evins GM, Fields PI, Olsvik O, Popovic T, Bopp CA, Wells JG, Carrillo C, Blake PA 1993. The molecular epidemiology of cholera in Latin America. *J Infect Dis* 167: 621-626.
- Yamai S, Okitsu T, Shimada T, Katsube Y 1997. Distribution of serogroups of *Vibrio cholerae* non-O1 non-O139 with specific reference to their ability to produce cholera toxin, and addition of novel serogroups. *J Japan Assoc Infect Dis* 71: 1037-1045.

