Gene Reports 4 (2016) 16–25

Contents lists available at ScienceDirect

Gene Reports

journal homepage: www.elsevier.com/locate/genrep

The Use of Recombined Ribosomal RNA Operon (rrn) Type-Specific Flanking Genes to Investigate rrn Differences Between *Vibrio parahaemolyticus* Environmental and Clinical Strains

Volker Gurtler [a,*], Danilla Grando [a], Ballamoole Krishna Kumar [b,c], Biswajit Maiti [b,c], Iddya Karunasagar [d], Indrani Karunasagar [b,c]

[a] School of Applied Sciences, Building 223, Level 1, Bundoora Campus, RMIT University, PO Box 71, Bundoora 3083, Australia
[b] UNESCO-MIRCEN for Marine Biotechnology, College of Fisheries, Karnataka Veterinary, Animal and Fisheries Sciences University, Mangalore, Karnataka, India
[c] Faculty of Biomedical Science, Nitte University Centre for Science Education and Research, University Enclave, Medical Sciences Complex, Deralakatte, Mangalore, Karnataka, India
[d] Subba-Meena, Jayanagar, Mangalore, India

ARTICLE INFO

Article history:
Received 3 February 2016
Received in revised form 5 February 2016
Accepted 5 February 2016
Available online 27 February 2016

Keywords:
Recombination
Fragmentation
Vibrio parahaemolyticus
16S rRNA gene
23 rRNA gene
Internal transcribed spacer

ABSTRACT

In order to interrogate the published genomes of clinical and environmental isolates of *Vibrio parahaemolyticus* an exploratory study was performed on the ribosomal RNA (*rrn*) type-specific conservation of the genes directly flanking the *rrn* operons of each *rrn* type. WGS was sourced from databases where available or performed *de novo* and resulted in 124 *rrn* operon sets for 5 environmental and 7 clinical isolates. Analyses showed that the number of operons varied between clinical (11–12 operons) and environmental (9–10) isolates. Recombination of operons was evident when using flanking gene homology ($p < 0.0001$) as a criterion for identifying native structural arrangements. This was confirmed by and correlated with ($p < 0.0001$) Mauve whole genome analysis, demonstrating considerable genome plasticity in environmental and clinical isolates. In environmental isolates, *rrn* types *rrn*D and *rrn*I had 23S and 16S ITS regions deleted. In comparison clinical isolates had the equivalent operons *rrn*EF and *rrn*GH arranged in tandem with no ORF between *rrn*E and *rrn*F or *rrn*G and *rrn*H. This suggested that operon components 16S, ITS and 23S respectively have been duplicated in the environmental isolates to produce 2 operons in tandem in the clinical isolates. These exploratory findings suggest that genome plasticity and recombination account for significant differences between environmental and clinical strains of *V. parahaemolyticus*. This information can be used to develop new typing methods that would be useful for studying the evolution of new pathogenic strains. Additionally, these types of analyses are useful for studying large genomic *rrn* operon recombination events over the entire genomes of *V. parahaemolyticus* strains.

Introduction

Vibrio parahaemolyticus is a gram negative halophilic bacterium that occupies diverse niches of marine environments and has been responsible for some cases of seafood-borne gastroenteritis (Nishibuchi and Kaper, 1995; Deepanjali et al., 2005; Nair et al., 2007; Karunasagar et al., 2013). Emergence of a unique pandemic clone belonging to O3:K6 serotype in Kolkata, India, in 1996 has abruptly changed the epidemiology of this organism which now accounts for large outbreaks and hospitalisations. At present, more than 20 serovariants belonging to the pandemic clone (Chen et al., 2011) have been responsible for human cases in different parts of the world (Nair et al., 2007; Karunasagar et al., 2013). Strains of *V. parahaemolyticus* recently involved in mortalities in brackish water shrimp that have led to production losses, however were not found to possess virulence characteristics known to be associated with human pathogenic strains (FAO, 2013; Tran et al., 2013; Kongrueng et al., 2014; Kumar et al., 2014c; Nunan et al., 2014). Virulence determinants vary from strain to strain and if there is genome flexibility, this may affect their environmental fitness and their infection capabilities.

Various virulence factors (TDH, TRH and T3SS-1&2) have been found to be useful for typing virulent and non-virulent strains of *V. parahaemolyticus* (Makino et al., 2003; Hurley et al., 2006). However the chromosomal mobility of these factors and their absence in some strains make the use of these genetic markers problematic as general typing tools (Kumar et al., 2014b). This report investigates the organisation of ribosomal genes to querry whether these motifs may identify outbreak strains. The typical bacterial *rrn* organisation is the 16S rRNA gene, 23S rRNA gene and 5S rRNA gene (Gurtler and Stanisich, 1996; Gurtler, 1999; Deutscher, 2009; Gürtler et al., 2014), which are each encoded from a continuous rRNA gene. Ribotyping using the 16S–23S

Abbreviations: ITS1, 16S–23S rRNA internal transcribed spacer; WGS, Whole genome sequence; LCB, Locally Collinear Blocks.

* Corresponding author.
E-mail address: volker.gurtler@gmail.com (V. Gurtler).

http://dx.doi.org/10.1016/j.genrep.2016.02.006
2452-0144/

Table 1
Genomic, genetic and phenotypic properties of *Vibrio parahaemolyticus* strains studied.

Strain name	GenBank accession no. (Chr I/Chr II)	Source	*rrn* no	Type III ChrI	Type III ChrII	*tdh*2	*trh*1	References
CDC_K4557	CP006008/CP006007	Clinical (stool)	11	Present	Absent	Absent	Absent	Jones et al. (2012)
RIMD2210633	NC_004603/NC_004605	Clinical	11	Present	Present T3SS2α	Present	Absent	Makino et al. (2003)
VP250	NZ_AVOK00000000	Clinical	NA	Present	T3SS2α	Present	Absent	
3259	NZ_AVOL00000000	Clinical	NA	Present	T3SS2α	tdhA	Absent	Turner et al. (2013)
NIHCB0603	NZ_AVOM00000000	Clinical	NA	Present	T3SS2α	tdhA	Absent	
949	NZ_AVPV00000000	Clinical	NA	Present	T3SS2α	tdhA	Absent	Turner et al. (2013)
NIHCB0757	NZ_AVPX00000000	Clinical	NA	Present	T3SS2β	tdhS	Present	
BB220P	NC_019955/NC_019971	Environmental	11	Present	Present T3SS2α	Present	Absent	Jensen et al. (2013)
FDA_R31	CP006004/CP006005	Environmental (oyster)	12	Present	Absent	Absent	Present	Jones et al. (2012)
UCM-V493	CP007004/CP007005	Environmental (sediment)	9	Present	Absent	Absent	Absent	Karunasagar et al. (2013); Gürtler et al. (2014); Kalburge et al. (2014)
VP49	JEMS00000000	Environmental (oyster)	9	Present	Present T3SS2β	Absent	Present	Kumar et al. (2014a)
VPA67	JPIP00000000	Environmental (shrimp haemolymph)	9	Present	Absent	Absent	Absent	Kumar et al. (2014b)
Refer to Figure:			1a&b					

NA, not available; sequencing data is incomplete.

rDNA internal transcribed spacer (ITS) has been used to type *Clostridium difficile* in clinical laboratories (Gurtler and Grando, 2013) and the use of such a tool may prove useful for typing *V. parahaemolyticus*. A number of studies (Maeda et al., 2000; Gonzalez-Escalona et al., 2006; Hoffmann et al., 2010) have investigated and shown sequence variation in and between the various operons of *V. parahaemolyticus*. These studies have used the 16S–23S rRNA gene internal transcribed spacer region (ITS) of the RNA operon (*rrn*) for typing and identification of several pathogenic *V. parahaemolyticus* strains. However these studies failed to take into account the complexity of the 8–12 *rrn* copies per genome. This paper will examine this complexity to determine if this variation can be utilised to better type these organisms.

Recombination has been proposed to occur within and between *rrn* operons (Gurtler, 1999; Maslunka et al., 2014), leading to gene incongruence (Knowles, 2009; Ane, 2011). Furthermore, a study of *V. parahaemolyticus rrn* operons has suggested that there are very high levels of intergenomic 16S rRNA gene recombination between the 8–12 *rrn* copies on chromosomes I & II (Harth et al., 2007). The program Mauve (Darling et al., 2010) may be used to determine the mosaic pattern of homology between sequence regions of highly related strains, which are created when these regions are rearranged, duplicated, gained or lost from a genome (Doerks et al., 2002; Darling et al., 2010). Mauve has been used to differentiate strains of *Aggregatibacter actinomycetemcomitans* with markedly different genome arrangements, classifying them into serotype "a" strains compared to strains with serotypes "b" or "c" (Kittichotirat et al., 2010). However no such analyses using Mauve have been published for *V. parahaemolyticus* strains, nor have studies been done on the effect of these genome rearrangements on the organisation of *rrn* operons in any bacterial species. Two studies have used the genes directly flanking either side of the *rrn* operons to demonstrate rearrangements in *Salmonella* serovars (Helm and Maloy, 2001) and *Actinobacillus actinomycetemcomitans* strains by PCR amplification (Eriksen et al., 2005) of each *rrn* type as defined above. The genomes of the genus *Vibrio* contain many more *rrn* types with much greater variability than most other bacterial species.

Notes to Table 2:
Table 2 Abbreviations, symbols and colour codes:
Yellow shading corresponds to genes flanking rrn operons.
All other coloured shades correspond to *rrn* operons with the same colour/alphabet coding as in Fig. 1.
[a]The percentage of full length 23S rRNA gene where "100" is the full length of 2900 bp.
[b]The percentage of full length 16S rRNA gene where "100" is the full length of 1550 bp.
[c]The *rrn* operons have been colour coded according to Fig. 1.
d—This contig does not contain the thioredoxin gene which could be explained if this region is truncated by oriC.
g—This contig does not contain the Na/H gene which could be explained if this region is truncated by oriC. This would make this region into a quadruple rrn operon like rrnBAKJ in FDA_R31.
Further information for abbreviations* for flanking genes from strain BB220P can be directly looked up in the NCBI database using the code given below in parenthesis immediately following the gene abbreviation.
aroE shikimate 5-dehydrogenase.
CadA (VPBB_2740), lysine decarboxylase.
CDS, coding sequence.
ClpB (VPBB_0534; M636_19035), ClpB protein or protein disaggregation chaperone.
crcB (VPBB_2859), camphor resistance protein.
fre (VPBB_2829), NAD(P)H-flavin reductase CDS.
gabD (VPBB_2450), succinate-semialdehyde dehydrogenase [NADH+] CDS.
GGDEF (VPBB_2736), diguanylate cyclase CDS.
HD-GYP (VPBB_2451), HD-GYP family phosphohydrolase CDS.
HemG (VPBB_0032), protoporphyrinogen IX oxidase, oxygen-independent, HemG CDS.
hemY (VPBB_2828), HemY-like protein CDS.
IlvY (VPBB_0034), HTH-type transcriptional regulator, CDS. Also known as lysR.
LdhA (VPBB_A0133), D-lactate dehydrogenaase.
MDR (VPBB_A0131), multi-drug resistance protein.
mepA (VPBB_0538), Murein endopeptidase/peptidase/penicillin-insensitive murein endopeptidase family protein CDS.
Na/H (VPBB_2538), Na+/H+ antiporter, putative CDS.
PaaY (VPBB_2863), carbonic anhydrase, family 3.
pssA (VPBB_2772), CDP-diacylglycerol–serine O-phosphatidyltransferase.
RRM (VPBB_2774), RNA-binding protein.
Thioredoxin (VPBB_2539), phosphoadenylyl-sulphate reductase [thioredoxin] CDS.