Detection and diversity of pathogenic Vibrio from Fiji

Reema Singh,1 Vinay Narayan,2 Patricia McLenachan,3 Richard C. Winkworth,1,3 Suparna Mitra,4 Peter J. Lockhart,3,5* Lorraine Berry,6 Abdulla M. Hatha,7 William Aalbersberg2 and Dhana Rao1,8
1School of Biological and Chemical Sciences and 2Institute of Applied Sciences, University of the South Pacific, Suva, Fiji.
3Institute of Molecular BioSciences, Massey University, Private Mailbag 11222, Palmerston North, New Zealand.
4Institute of Translational Medicine, University of Liverpool, L69 3GL Liverpool, UK.
5Institute of Fundamental Sciences and 6Massey Genome Service, Massey University, Private Mailbag 11222, Palmerston North, New Zealand.
7Department of Marine Biology, Microbiology & Biochemistry, Cochin University of Science & Technology, Kerala, India.
8Department of Biological Sciences, Marshall University, Huntington, WV 25755, USA.

Introduction

Members of the genus Vibrio are Gram-negative, motile rods ubiquitous in marine environments (Farmer, 1992). Farmer and colleagues (2003) recognize 12 Vibrio species as clinically important human pathogens. Disease-causing vibrios are most commonly associated with gastrointestinal infections. Two well-known examples are V. cholerae and V. parahaemolyticus. Vibrio cholerae is the causative agent of cholera. Seven documented cholera pandemics have claimed millions of lives and the disease continues to affect 3–5 million people annually (Blake, 1994; Wachsmuth et al., 1994; Faruque et al., 1998; Reeves and Lan, 1998; Gaffga et al., 2007). Both toxigenic and non-toxigenic strains of this species occur naturally in aquatic ecosystems and infection is often via contaminated water or food. Vibrio parahaemolyticus is a leading cause of bacterial gastroenteritis in the USA (e.g. Daniels et al., 2000; McLaughlin et al., 2005) and Asia (e.g. Wong et al., 1999; Chiou et al., 2000). In this case infections are commonly the result of consuming contaminated seafood (Chan et al., 1989; Kagiko et al., 2001). In addition to gastrointestinal infections vibrios are also associated with infections of wounds, blood, ears, eyes and skin (Penland et al., 2000; Oliver, 2005; Oliver and Kaper, 2005; Pruzzo et al., 2005).

Several methods are available for the identification of vibrios from clinical and environmental samples. These include phenotypic characterization (e.g. Myhr et al., 1991; Alsina and Blanch, 1994; Noguerola and Blanch, 2008), DNA–DNA hybridization (e.g. Wang et al., 2011), 16S ribosomal sequences (e.g. Dorsch et al., 1992), amplified fragment length polymorphism (e.g. Jiang et al., 2000) and multi-locus sequence analysis (e.g. Thompson et al., 2005; 2007a,b; Sawabe et al., 2007). These approaches have done much to improve our understanding of vibrio diversity. Yet all these methods rely on culturing of the bacteria prior to characterization and this is now widely recognized as biasing the outcomes of subsequent diversity analyses (e.g. Thompson et al., 2004). In contrast, environmental metagenomics provides a tool with which to survey the diversity of microbial communities without the need to culture the bacteria first (Green et al., 2005; Hugenholtz and Tyson, 2008; Husson et al., 2009). This methodology has already been used to investigate the structure of microbial communities in marine water column and soil environments (e.g. Allen...
and Banfield, 2005; Daniel, 2005; Delong, 2005). An extension of describing microbial communities from individual samples is to monitor the structure of such assemblages using samples drawn over time; metagenomics has considerable potential in this context (Lazarevic et al., 2009; Caporaso et al., 2010; Qin et al., 2010; Rodriguez-Brito et al., 2010). To date this approach has not been applied to long-term monitoring (e.g. for water quality management) mainly due to the cost of the high-throughput sequencing technologies on which these analyses are now based. However, rapid methodological developments, in particular the development of cheaper short read sequencing protocols, suggests that it may become cost-effective to use metagenomics for the routine monitoring of microbial communities.

Little is known about the distribution and prevalence of disease-causing *Vibrio* species in the South Pacific. In the present study we use four distinct yet complementary analyses – biochemical testing, phylogenetic analyses, metagenomic analyses and molecular typing – to investigate the diversity of vibrios in this region. In the first part of our study, we combine analyses of fish-associated isolates with those of water column DNA diversity of pathogenic *Vibrio* species present in Fijian marine environments and assess the potential health risks posed by these bacteria. In the second part we use the results of our metagenomic analyses to examine the potential and limitations of short read sequencing for routine monitoring of aquatic bacterial communities.

**Results**

**Phenotypic identification of Vibrio species**

Colony growth on selective TCBS agar suggests that *Vibrio* species are common on fish available from retail outlets in Suva, Fiji. In the present study we detected vibrios on 88.9% (160/180) of the sampled fish; these bacteria were cultured from all retail locations (roadside fish stalls, local fish markets and fish shops) and all three regions of the fish tested (skin, gills and the gut cavity).

Biochemical tests suggest a number of *Vibrio* species were present on sampled fish. Tests provided unequivocal identifications for just a small portion of isolates – *V. cholerae* (2 isolates), *V. metschnikovii* (9 isolates), *V. mimicus* (12 isolates) and *V. navarrensis* (46 isolates). For the remaining 256 isolates our tests are consistent with several candidate species (see Supporting information for a complete list of the biochemical test results). Often the unequivocally identified species were not among the candidate species thus suggesting additional diversity is associated with these fish. While we were unable to describe the fish-associated *Vibrio* community in full, our results indicate the presence of potentially pathogenic species. All four of the unequivocally identified species are known to infect humans (Farmer et al., 2003) while several pathogens are among the candidate species for the unidentified isolates (e.g. *V. parahaemolyticus* and *V. vulnificus*).

**Phylogenetic analyses of V. parahaemolyticus isolates**

We used phylogenetic analyses of 16S, recA and pyrH sequences to further examine the diversity of presumptive *V. parahaemolyticus* isolated from fish (see Supporting information). Partial 16S rDNA sequences from 22 accessions share 100% identity with previously published *V. alginolyticus*, *V. diabolicus*, *V. natriegens* and *V. parahaemolyticus* sequences. These accessions belong to a clade containing additional representatives of *V. alginolyticus* and *V. parahaemolyticus* with the remaining core group species (i.e. *V. campbellii*, *V. harveyi*, *V. mytillus* and *V. rotiferianus*) more distantly related (Fig. 1; Tables S3 and S4). Whereas our 16S analysis did not distinguish between *V. parahaemolyticus* and its close relatives, analyses of *recA* and *pyrH* sequences recovered well-supported clades corresponding to each of the included species (Figs 2 and 3). Fijian *recA* and *pyrH* sequences are a combination of novel forms (e.g. Fiji³ and Fiji¹⁰ for *recA*) and those encountered in previous analyses (e.g. Fiji² is equivalent to *V. parahaemolyticus* SG259).

Although most Fijian accessions were associated with the *V. parahaemolyticus* clade, Fiji¹² and Fiji¹¹ are more closely related to sequences from *V. alginolyticus* in *recA* and *pyrH* analyses (Figs 2 and 3). The *recA* sequence for one non-Fijian *V. parahaemolyticus* accession (strain 28; Gonzalez-Escalona et al., 2008) also grouped with the *V. alginolyticus* clade (Fig. 2).

**Metagenomic analysis of seawater samples**

Metagenomic analyses of total community DNA were used to investigate *Vibrio* diversity in the water column close to Suva (see Supporting information). For single end analyses a random sample of $2 \times 10^5$ 75 bp reads was analysed using MEGAN (Huson et al., 2011). MEGAN assigned 23 561 (11.8%) of reads to the NCBI taxonomic hierarchy. *Vibrio* and the related genus *Photobacterium* were both identified in this sample. Reads were assigned to five *Vibrio* species – *V. cholerae*, *V. harveyi*, *V. parahaemolyticus*, *V. shilonii* and *V. vulnificus* – and to strains within three of these (Fig. 4). The number of reads assigned to nodes and the number of phylogenetic lineages identified increased substantially when reads were treated as paired ends (i.e. sequences from each end of a DNA fragment are considered jointly). When the same sample of $2 \times 10^5$ reads was analysed in this way, the number of assigned fragments increased to 42 354
(21.2%) with reads assigned to three genera of Vibrionaceae (i.e. *Aliivibrio*, *Photobacterium* and *Vibrio*), six *Vibrio* species (i.e. those identified in the single read analysis plus *V. splendidus*) and four *Vibrio* strains (Fig. 4).

The vibrio community identified by metagenomic analyses contained three important human pathogens, *V. cholerae*, *V. parahaemolyticus* and *V. vulnificus*. In addition, *V. harveyi* has been found to sporadically infect humans. The two remaining species, *V. shilonii* and *V. splendidus*, are non-pathogenic (Farmer et al., 2003).

**Molecular typing of pathogenicity loci**

Both biochemical testing and metagenomic analyses identified potentially pathogenic *Vibrio* species but neither approach specifically evaluates whether these bacteria carry the genetic determinants of pathogenicity. For one species, *V. parahaemolyticus*, we examined the pathogenicity of Fijian strains using molecular typing of loci with known or putative roles in pathogenicity (see Supporting information).

We first used a pair of *V. parahaemolyticus*-specific PCR markers to confirm the identity of presumptive *V. parahaemolyticus* isolates. Results for the *tlh* and *toxR* markers suggested all but one isolate (i.e. Fiji³¹) was *V. parahaemolyticus* (Table S5). Pathogenicity typing used markers for two key pathogenicity loci (i.e. *tdh* and *trh*) plus six loci (i.e. *ure*, VPA1321, VPA1339, VPA1346, VPA1376 and MTase) with known or putative roles in pathogenicity. The thermostable direct haemolysin (*tdh*) locus was detected for all the confirmed *V. parahaemolyticus* isolates; all the isolates lacked the *tdh*-related haemolysin (*trh*) locus (Table S5). Clinical strains of *V. parahaemolyticus* commonly express one or both of these proteins with *tdh*-producing strains identified in > 90% of cases (Caburlotto et al., 2009). For the remaining pathogenicity loci results varied; up to 22 isolates testing positive for any given locus (Table S5).

In contrast, molecular typing of DNA isolated from our seawater samples did not produce positive amplifications for the two *V. parahaemolyticus*-specific nor for any of the eight pathogenicity loci (see Supporting information). Since metagenomic analyses indicate *V. parahaemolyticus* is present in the water column we suspect the issue here is sensitivity of the staining protocols used to visualize PCR products. It appears that amplification of loci did not increase copy number sufficiently for taxa-specific and

© 2012 Society for Applied Microbiology and Blackwell Publishing Ltd, *Environmental Microbiology Reports*, 4, 403–411
pathogenicity markers to be visualized using ethidium bromide and SBYR green staining protocols.

Discussion

Diversity of pathogenic vibrios in Fijian marine environments

Our analyses of isolates from commercially available fish and DNA extracted from the water column confirmed the presence of nine *Vibrio* species in marine environments close to Suva. However, further studies of vibrio diversity in the Pacific are warranted, as we cannot rule out the presence of other taxa. Among those identified from Fiji are three of the most important disease-causing vibrio species. Specifically, we detected *V. cholerae*, *V. parahaemolyticus* and *V. vulnificus* in the water column and two of these, *V. cholerae* and *V. parahaemolyticus*, on commercially available fish. Species that more sporadically infect humans were also found (e.g. *V. metschnikovii* and *V. mimicus*).

Phylogenetic analyses and molecular typing of *V. parahaemolyticus* isolates further suggest that this species is represented by several variants in the region. For *recA* and *pyrH* we identified both widespread and novel sequence types. Given the importance of mobile genetic elements and horizontal gene transfer in the evolution of *Vibrio* (e.g. Hazen *et al.*, 2010; Pascual *et al.*, 2010) these may be best interpreted as suggesting a diverse genetic background for this species in Fijian marine environments. Indeed our results for two fish-associated isolates (e.g. Fiji21 and Fiji21) suggest horizontal genetic exchange. Based on biochemical testing both these isolates are presumptively *V. parahaemolyticus* yet phylogenetic analyses and molecular typing are contradictory. These tests suggest genetic elements have been acquired from *V. alginolyticus*; Pascual and colleagues (2010) have recently reported a similar result for the *rctB* gene. Interestingly, we did not unequivocally identify *V. alginolyticus* in our samples suggesting these transfers may not be specific to Fijian strains.
In general our results suggest a number of pathogenic vibrios occur naturally in Fijian marine environments. In the case of *V. parahaemolyticus* we also show that toxigenic forms of this species are also present. That we find toxigenic vibrios associated with commercially available fish in Fiji is consistent with findings from various other studies (e.g. Chan *et al.*, 1989; El hadi *et al.*, 2004; Gopal *et al.*, 2005). While there are currently no data directly linking gastrointestinal illness to *Vibrio* infection in Fiji such illnesses are common (Fiji Centre for Communicable Disease Control, 2009a,b; 2010). Given that raw fish is frequently eaten in Fiji our results strongly suggest *Vibrio* is a potential health risk.

The potential of high-throughput sequencing approaches for biomonitoring

Our MEGAN analyses indicate that a relatively small sample of short DNA sequences (2 × 10^5 75 bp reads) is sufficient to distinguish closely related *Vibrio* species and identify strain level lineages. Further, as predicted by Mitra and colleagues (2010), when we treat reads as paired ends the number of DNA fragments that can be unambiguously assigned to the NCBI taxonomic hierarchy is substantially increased. Our results are encouraging since it had been thought that only analyses of long read sequences would provide the resolution necessary to describe the composition of naturally occurring microbial communities.

If, as our analyses suggest, it is possible use short read sequencing to determine the structure of naturally occurring microbial communities this would increase the cost-effectiveness of monitoring such assemblages over time using metagenomics (c.f. long read-based protocols). Further, if it can be shown that a small sample of short reads is sufficient to consistently describe community structure this would result in further cost reductions. For example, the current capacity of Illumina’s MiSeq sequencer is 5 × 10^6 150 bp paired-end reads per run. Assuming 2 × 10^5 reads are sufficient to describe a community up to 25 indexed samples could be analysed on a single flow cell lane. These observations suggest that it may become cost-effective to use metagenomics for the routine monitoring of microbial communities in the near future.

Limitations to overcome with high-throughput sequencing approaches

One limitation illustrated by the present study involves the detection of toxigenic forms. While we identified potentially pathogenic *Vibrio* species using MEGAN our analyses did not detect genetic loci specifically associated with pathogenicity. Clearly the ability to identify such loci is important in the context of monitoring programmes. The result is, however, not unexpected given the relatively small sample of reads analysed; identifying specific loci is likely to require a larger number of reads than does identifying taxa. While one could increase the number of reads analysed (e.g. from 2 × 10^5 to 3 × 10^7 reads, approximately the number of reads generated our Illumina GAII runs) there is no guarantee that this approach would result in detection of individual genes from an environmental sample. We think it is more effective to combine...
metagenomic analyses that describe the overall structure of the microbial community with molecular typing of specific loci to determine whether toxigenic forms are present. However, the detection of PCR products with standard staining protocols is a problem. A much more sensitive approach is to use the high-throughput sequencing apparatus to detect PCR products. This could easily be done by spiking the environmental DNA sample prior to sequencing with aliquots of PCR amplifications for pathogenicity loci; in each case the template for PCR amplifications would be the same environmental DNA sample.

A second potential issue is completeness of the reference database used by MEGAN. If target organisms are poorly represented then both the number and phylogenetic resolution of assignments are likely to be limited (Huson et al., 2007; Morgan et al., 2010). Here we used a database (NCBI non-redundant protein, February 2010 version) that contained 57 complete Vibrio genomes. All the strain-level lineages identified by MEGAN are represented by complete genome sequences (e.g. V. harveyi ATCC BAA-1116 and V. parahaemolyticus 16) suggesting that the database needs to contain complete, or nearly so, genomes in order to differentiate very closely related taxa. At face value this is a potential weakness of the methodology since not all lineages will be represented in this way. However, we suspect that the clinical importance of pathogens, which are often the targets of monitoring programmes, will result in them being over-represented in databases and thus readily identified. More generally, the continued growth of reference databases should reduce the problem of under-representation over time.

Conclusion

Our study provides a first insight into the diversity of Vibrio species in Fijian marine environment. Understanding both the distribution and diversity of vibrios in the Pacific Islands is an important step towards managing the health risks posed by these bacteria. Metagenomic analyses based upon high-throughput sequencing protocols will likely have an important role to play in helping assess the risks posed by vibrios and other disease-causing microbes in naturally occurring communities.

Acknowledgements

We thank the Oceania Research Network, NZaid, the Allan Wilson Center for Molecular Evolution and Ecology (Massey University) and Faculty of Science, Technology and Environment (University of the South Pacific) for supporting this work. Peter Lockhart thanks the New Zealand Royal Society for financial support.

References


© 2012 Society for Applied Microbiology and Blackwell Publishing Ltd, Environmental Microbiology Reports, 4, 403–411


Supporting information

Additional Supporting Information may be found in the online version of this article:

Experimental procedures.

Table S1. Biochemical test results for colonies that were green or green/blue on selective TCBS agar.

Table S2. Biochemical test results for colonies that were yellow on selective TCBS agar.

Table S3. GenBank accession details for DNA sequences from non-Fijian *Vibrio* species/strains included in phylogenetic analyses.

Table S4. GenBank accession details for DNA sequences from presumed *Vibrio parahaemolyticus* isolates included in phylogenetic analyses.

Table S5. Results of molecular typing for isolates presumed to be *Vibrio parahaemolyticus*.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.