Identification of *Vibrio cholerae* and *Vibrio mimicus* by multilocus sequence analysis (MLSA)

Cristiane C. Thompson,^{1,2} Fabiano L. Thompson²
and Ana Carolina P. Vicente¹

Correspondence

Cristiane C. Thompson
thompson@ioc.fiocruz.br

¹Instituto Oswaldo Cruz, Rio de Janeiro, Brazil

²Department of Genetics, Federal University of Rio de Janeiro, Brazil

Vibrio cholerae and *Vibrio mimicus* have similar phenotypes and genomes making rapid differentiation of these two species difficult. The first standard multilocus sequence analysis (MLSA) scheme for the identification of these species is described. A collection of 45 representative isolates from different geographical regions and hosts was examined using segments of the housekeeping genes *pyrH*, *recA* and *rpoA*. Overall, the closest phylogenetic neighbours of these species were *Vibrio furnissii* and *Vibrio fluvialis*. *V. cholerae* and *V. mimicus* formed separate species clusters on the basis of each gene, suggesting that these genes are useful as identification markers. These species clusters arose by the accumulation of point mutations. The *pyrH* gene showed the highest resolution for differentiating *V. cholerae* and *V. mimicus*. The maximum interspecies *pyrH* gene sequence similarity was 91 %. Clearly, *V. mimicus* strains were more heterogeneous than *V. cholerae* strains at the three loci. It is suggested that vibrio species may be defined on the basis of MLSA data. A vibrio species was defined as a group of strains forming a monophyletic group on the basis of these loci and with an intraspecific sequence similarity of at least 95 %. *V. cholerae* and *V. mimicus* isolates can be readily identified through the open database resource 'The Taxonomy of Vibrios' (<http://www.taxvibrio.incc.br/>).

Vibrio cholerae, the causative agent of cholera, is a major pathogen worldwide, particularly in the developing world. Cholera killed at least 5000 people in 2006 in Angola and neighbouring countries (<http://www.who.int/en/>). *V. cholerae* can be classified into more than 200 serogroups, but only serogroups O1 and O139 have been responsible for epidemics of cholera (Hunter, 1997; Kay *et al.*, 1994). The species *Vibrio mimicus* was first proposed to encompass biochemically atypical non-O1 *V. cholerae* isolates that were unable to utilize sucrose (Davis *et al.*, 1981). Phenotypically, most of the features of *V. mimicus* are identical to those found in *V. cholerae* (Davis *et al.*, 1981). A sucrose-negative strain of *V. cholerae* was identified during cholera epidemics (Desmarchelier & Reichelt, 1984; Ramos *et al.*, 1997), making it difficult to differentiate these

two species on this basis. Because *V. cholerae* and *V. mimicus* share nearly 100 % 16S rRNA gene sequence similarity and show around 80 % DNA–DNA relatedness, discriminating between these species remains a difficult task for taxonomists and clinicians (Thompson *et al.*, 2003). In addition, both species may also share important virulence factors, such as *ct* and *tcp* genes (Boyd *et al.*, 2000; O'Shea *et al.*, 2004). *V. mimicus* has only been associated with small outbreaks or isolated cases of diarrhoea (Campos *et al.*, 1996; Uchimura *et al.*, 1993).

Amplified Fragment Length Polymorphism (AFLP) has proved to be the most powerful taxonomic tool for discriminating species and strains of vibrios (Thompson *et al.*, 2003, 2004). However, AFLP is not easily adaptable for use in an online electronic taxonomy (Thompson *et al.*, 2005, 2007; Thompson & Swings, 2006). The online electronic taxonomy of vibrios by means of multilocus sequence analysis (MLSA) may help to track down the widespread clones of *V. cholerae* and its sister species *V. mimicus*. In this study, we analysed the taxonomic resolution of the uridylate kinase (*pyrH*), recombination repair protein (*recA*) and RNA polymerase alpha subunit (*rpoA*) genes for the identification of *V. cholerae* and *V. mimicus* and for the determination of the microevolutionary history of these two species. We showed that *V. cholerae* and *V. mimicus* clearly form two separate

Abbreviations: AFLP, amplified fragment length polymorphism; MLSA, multilocus sequence analysis.

The GenBank/EMBL/DDBJ accession numbers for the *pyrH*, *recA* and *rpoA* gene sequences of *Vibrio cholerae* and *Vibrio mimicus* are EF990257–EF990363 and EU147787–EU147789.

Tables detailing the sources of the strains used in this study, the primers used for amplification and sequencing of the three loci, features of the genes and signatures of the *V. cholerae* and *V. mimicus* *pyrH*, *recA* and *rpoA* genes are available with the online version of this paper. Information is also available on how the recombination analysis was performed.

species and can be discriminated on the basis of gene sequences.

A total of 29 *V. cholerae* and 16 *V. mimicus* strains from both clinical and environmental sources, from different serogroups and from different geographical regions that had been collected over the last 37 years were examined (see Supplementary Table S1 in IJSEM Online). The strains were grown aerobically on tryptone soy agar [supplemented with 1.5 % NaCl (w/v)] at 37 °C for 24 h. DNA was extracted as described previously (Thompson *et al.*, 2001).

The primers used for amplification and sequencing of the three loci are listed in Supplementary Table S2 (available with the online version of this paper). Amplification and sequencing of these loci were performed as described previously (Thompson *et al.*, 2005). The annealing temperature was 46 °C for *recA* and 55 °C for *pyrH* and *rpoA*. PCR products were purified using Perfect Gel Cleanup (Eppendorf). DNA sequencing was performed in an ABI PRISM 3730 Genetic Analyzer (Applied Biosystems). Raw sequence data were transferred to DNASTAR software (Lasergene) where consensus sequences were determined.

The sequences were aligned by CLUSTAL_X. Phylogenetic analyses were conducted using MEGA version 3.1 (Kumar *et al.*, 2004). The phylogenetic inference was based on the neighbour-joining genetic distance method (Saitou & Nei, 1987). Distance estimations were obtained according to the Tajima–Nei (Tajima & Nei, 1984), Kimura two-parameter (Kimura, 1980) and Jukes–Cantor (Jukes & Cantor, 1969) models for the *pyrH*, *recA* and *rpoA* genes, respectively. The reliability of each tree topology was checked by 1000 bootstrap replications. The DNA G+C content was calculated using MEGA version 3.1 (Kumar *et al.*, 2004).

We report the first standard multilocus sequence analysis (MLSA) scheme based on *pyrH*, *recA* and *rpoA* gene sequences for the identification of *V. cholerae* and *V. mimicus* isolates obtained in the last 37 years from different continents. These strains are thought to represent the known genomic diversity of these species as revealed by AFLP and multilocus enzyme electrophoresis (MLEE) (Thompson *et al.*, 2003; Vieira *et al.*, 2001). Some features of these loci are shown in Supplementary Table S3 (see IJSEM Online). Recombination was detected only in the *recA* gene. The closest neighbours of these species were *Vibrio furnissii* and *Vibrio fluvialis* (Supplementary Table S3). The total number of polymorphic sites for the *pyrH*, *recA* and *rpoA* genes was 77, 83 and 25, respectively. The number of alleles for the *pyrH*, *recA* and *rpoA* genes was 24, 25 and 16, respectively.

V. cholerae and *V. mimicus* clearly formed separated monophyletic clusters within the genus *Vibrio* on the basis of the three loci. *V. mimicus* strains were more heterogeneous than *V. cholerae* strains for the three genetic loci (Fig. 1). The *pyrH* gene showed the highest resolution for the differentiation of *V. cholerae* and *V. mimicus* (12 %

gene sequence divergence). Clearly, the *pyrH* gene will be very useful for further taxonomic environmental surveys of vibrios. The *recA* gene had 9 % sequence divergence and the *rpoA* gene had only 2 % sequence divergence, but even so the latter did separate the two species. *V. cholerae* strains had 1–5 % *recA* gene sequence variation and at least 1 % *pyrH* and *rpoA* gene sequence variation. The *pyrH* gene sequences were more heterogeneous than the *recA* and *rpoA* gene sequences in *V. mimicus*. Strains of *V. mimicus* had 1, 2 and 5 % *rpoA*, *recA* and *pyrH* gene sequence variation, respectively. Most of the *V. mimicus* strains originated from Brazil, but these strains did not form a tight group. The pairs of strains VM 535 and VM 542, and VM 603 and VM 605 had identical sequences for all loci. The latter pair belong to the same zymovar (Vieira *et al.*, 2001). In contrast, the pairs of strains LMG 7896^T and VM 461, R-20568 and VM 606, and VM 337 and VM 343 showed different relationships for both loci.

All *V. cholerae* O139 strains (R-20544, R-20545, R-20546 and R-20548) and most of the *V. cholerae* O1 strains (LMG 21698^T, R-18307, R-18308, R18314, VM 33, VM 34, VM 37, VM 79, VM 96, VM 111 and VM 508) were found to belong to the same clone on the basis of the three loci. Interestingly, strain VM 508 is a sucrose non-fermenting O1 strain that emerged in French Guiana during the South American pandemic between 1994 and 1995, displacing the regular *V. cholerae* O1 strain (Ramos *et al.*, 1997). *V. cholerae* O139 emerged from *V. cholerae* O1 El Tor possibly by one or more horizontal gene transfer events that resulted in deletions and replacements of the gene cluster encoding enzymes involved in lipopolysaccharide O-side chain synthesis (Bik *et al.*, 1995; Mooi & Bik, 1997). Our data clearly demonstrate the close genetic relationship between the *V. cholerae* O1 El Tor and the O139 serogroups, a finding that is in accordance with previous studies on multilocus sequence typing (MLST) of *V. cholerae* (O'Shea *et al.*, 2004). Interestingly, the African *V. cholerae* O1 strains (VM 97 and VM 104) were identical to the Amazonian variant O1 strains R-18332 and R-18356 based on *pyrH* and *recA* gene sequences. *V. cholerae* strain VM 104 was an exception to this observation as it grouped together with the other O1 and O139 strains in the *recA* gene analysis. Originally, the Amazonian variant O1 strains had genotypic traits completely distinct from other clinical epidemic *V. cholerae* O1 strains (Coelho *et al.*, 1995). *V. cholerae* non-O1/non-O139 strains were separated from the serogroups O1 and O139 by *pyrH* and *recA* gene sequences. *V. cholerae* non-O1/non-O139 strains are known as causative agents of sporadic and localized outbreaks of diarrhoea (Blake *et al.*, 1980). Strain R-18256 (non-O1/non-O139) clustered in-between *V. cholerae* and *V. mimicus*. This strain appeared to be at the outskirts of the species *V. cholerae* as shown by *pyrH* gene sequences. Non-O1/non-O139 strains were found in different branches, indicating their diverse evolutionary origins. Analysis of the *rpoA* gene sequences grouped together *V. cholerae* O1, O139 and non-O1/non-O139 strains. The

Fig. 1. Phylogenetic trees based on the neighbour-joining method. (a) *pyrH* (460 bp). (b) *recA* (583 bp). (c) *rpoA* (733 bp). Distance estimations were obtained using the Tajima–Nei, Kimura two-parameter and Jukes–Cantor models, respectively. Bootstrap percentages after 1000 replications are shown. *Photobacterium leiognathi* LMG 4228^T, *Photobacterium phosphoreum* LMG 4233^T, *Photobacterium angustum* LMG 8455^T and *Photobacterium iliopiscarium* LMG 19543^T were included as an outgroup. Bars, 2% estimated sequence divergence.

ct-positive *V. cholerae* strains (R-18307, R-18308, VM 33, VM 34, VM 96 and VM 111) were mixed with *ct*-negative strains for all loci. The signatures of *V. cholerae* and *V. mimicus* for each locus are shown in Supplementary Table S4 (available in IJSEM Online).

A bacterial species is defined as a group of strains that share phenotypic, genotypic and phylogenetic similarities that are unique and useful for differentiating it from its closest neighbours (Gevers *et al.*, 2005). Phenotypic identification of vibrios is problematic, largely because of the large variability in biochemical characteristics and the close genomic similarity of several species (Thompson & Swings, 2006). In this study, we showed that *V. mimicus* and *V. cholerae* can be differentiated on the basis of *pyrH*, *recA* and *rpoA* gene sequences. Species of vibrios can be defined as groups of strains forming a monophyletic group on the basis of these loci and possessing an intraspecific sequence similarity of at least 95 %. We suggest that species may be defined on the basis of MLSA data using the online electronic taxonomy resource 'The Taxonomy of the Vibrios' (<http://www.taxvibrio.lncc.br>).

Acknowledgements

C. C. T. is a PhD student of the graduate program on Genetics of the Federal University of Rio de Janeiro (UFRJ). C. C. T. acknowledges a PhD scholarship from CNPq. F. L. T. acknowledges grants from IFS, FAPERJ and CNPq. We thank Fernanda S. Freitas and Koko Otsuki for their excellent technical assistance and the PDTIS/FIOCRUZ genomic platform for DNA sequencing.

References

- Bik, E. M., Bunschoten, A. E., Gouw, R. D. & Mooi, F. R. (1995). Genesis of the novel epidemic *Vibrio cholerae* O139 strain: evidence for horizontal transfer of genes involved in polysaccharide synthesis. *EMBO J* **14**, 209–216.
- Blake, P. A., Weaver, R. W. & Hollins, D. G. (1980). Diseases of humans (other than cholera) caused by vibrios. *Annu Rev Microbiol* **34**, 341–367.
- Boyd, E. F., Moyer, K. E., Shi, L. & Waldor, M. K. (2000). Infectious CTXΦ and the vibrio pathogenicity island prophage in *Vibrio mimicus*: evidence for recent horizontal transfer between *V. mimicus* and *V. cholerae*. *Infect Immun* **68**, 1507–1513.
- Campos, E., Bolanos, H., Acuna, M. T., Diaz, G., Matamoros, M. C., Raventos, H., Sanchez, L. M., Sanchez, O. & Barquero, C. (1996). *Vibrio mimicus* diarrhea following ingestion of raw turtle eggs. *Appl Environ Microbiol* **62**, 1141–1144.
- Coelho, A., Andrade, J. R., Vicente, A. C. & Salles, C. A. (1995). New variant of *Vibrio cholerae* O1 from clinical isolates in Amazonia. *J Clin Microbiol* **33**, 114–118.
- Davis, B. R., Fanning, G. R., Madden, J. M., Steigerwalt, A. G., Bradford, H. B., Jr, Smith, H. L., Jr & Brenner, D. J. (1981). Characterization of biochemically atypical *Vibrio cholerae* strains and designation of a new pathogenic species, *Vibrio mimicus*. *J Clin Microbiol* **14**, 631–639.
- Desmarchelier, P. & Reichelt, J. L. (1984). A phenotypic and genetic study of sucrose nonfermenting strains of *Vibrio mimicus* and *Vibrio cholerae*. *Curr Microbiol* **10**, 41–47.
- Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., Stackebrandt, E., Van de Peer, Y., Vandamme, P. & other authors (2005). Opinion: re-evaluating prokaryotic species. *Nat Rev Microbiol* **3**, 733–739.
- Hunter, P. R. (1997). *Waterborne Disease: Epidemiology and Ecology*. England: John Wiley and Sons.
- Jukes, T. H. & Cantor, C. R. (1969). Evolution of protein molecules. In *Mammalian Protein Metabolism*, pp. 21–132. Edited by H. H. Munro. London: Academic Press.
- Kay, B. A., Boop, C. A. & Wells, J. G. (1994). Isolation and identification of *Vibrio cholerae* O1 from fecal specimens. In *Vibrio cholerae and Cholera. Molecular to Global Perspectives*, pp. 3–25. Edited by K. Wachsmuth, P. A. Blake & O. Olsvik. Washington: American Society for Microbiology.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.
- Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.
- Mooi, F. R. & Bik, E. M. (1997). The evolution of epidemic *Vibrio cholerae* strains. *Trends Microbiol* **5**, 161–165.
- O'Shea, Y. A., Reen, F. J., Quirke, A. M. & Boyd, E. F. (2004). Evolutionary genetic analysis of the emergence of epidemic *Vibrio cholerae* isolates on the basis of comparative nucleotide sequence analysis and multilocus virulence gene profiles. *J Clin Microbiol* **42**, 4657–4671.
- Ramos, F. L., Lins-Laison, Z. C., da Silva, E. L., Proietti, A. A., Jr, Mareco, M. L. & Lamarao, M. L. (1997). Cholera in north Brazil: on the occurrence of strains of *Vibrio cholerae* O1, which fail to ferment sucrose during routine plating on thiosulphate-citrate-bile salt-sucrose agar (TCBS). A new problem in diagnosis and control? *Rev Latinoam Microbiol* **39**, 141–144.
- Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
- Tajima, F. & Nei, M. (1984). Estimation of evolutionary distance between nucleotide sequences. *Mol Biol Evol* **1**, 269–285.
- Thompson, F. L. & Swings, J. (2006). Taxonomy of the vibrios. In *The Biology of Vibrios*, pp. 29–43. Edited by F. L. Thompson, B. Austin & J. Swings. Washington: American Society for Microbiology.
- Thompson, F. L., Hoste, B., Vandemeulebroecke, K. & Swings, J. (2001). Genomic diversity amongst *Vibrio* isolates from different sources determined by fluorescent amplified fragment length polymorphism. *Syst Appl Microbiol* **24**, 520–538.
- Thompson, F. L., Thompson, C. C., Vicente, A. C. P., Theophilo, G. N. D., Hofer, E. & Swings, J. (2003). Genomic diversity of clinical and environmental *Vibrio cholerae* strains isolated in Brazil between 1991 and 2001 as revealed by FAFLP analysis. *J Clin Microbiol* **41**, 1946–1950.
- Thompson, F. L., Iida, T. & Swings, J. (2004). Biodiversity of vibrios. *Microbiol Mol Biol Rev* **68**, 403–431.
- Thompson, F. L., Gevers, D., Thompson, C. C., Dawyndt, P., Naser, S., Hoste, B., Munn, C. B. & Swings, J. (2005). Phylogeny and molecular identification of vibrios on the basis of multilocus sequence analysis. *Appl Environ Microbiol* **71**, 5107–5115.
- Thompson, F. L., Gomez-Gil, B., Ribeiro Vasconcelos, A. T. & Sawabe, T. (2007). Multilocus sequence analysis reveals that *Vibrio harveyi* and *Vibrio campbellii* form distinct species. *Appl Environ Microbiol* **73**, 4279–4285.

Uchimura, M., Koiwai, K., Tsuruoka, Y. & Tanaka, H. (1993). High prevalence of thermostable direct hemolysin (TDH)-like toxin in *Vibrio mimicus* strains isolated from diarrhoeal patients. *Epidemiol Infect* **111**, 49–53.

Vieira, V. V., Teixeira, L. F. M., Vicente, A. C. P., Momen, H. & Salles, C. A. (2001). Differentiation of environmental and clinical isolates of *Vibrio mimicus* from *Vibrio cholerae* by multi locus enzyme electrophoresis. *Appl Environ Microbiol* **67**, 2360–2364.