Phylogenetic analysis of vibrios and related species by means of atpA gene sequences

Cristiane C. Thompson,¹,² Fabiano L. Thompson,² Ana Carolina P. Vicente¹ and Jean Swings³

¹Instituto Oswaldo Cruz, Rio de Janeiro, Brazil
²Department of Genetics, Federal University of Rio de Janeiro, Brazil
³Laboratory of Microbiology and BCCM/LMG Bacteria Collection, Ghent University, Belgium

We investigated the use of atpA gene sequences as alternative phylogenetic and identification markers for vibrios. A fragment of 1322 bp (corresponding to approximately 88% of the coding region) was analysed in 151 strains of vibrios. The relationships observed were in agreement with the phylogeny inferred from 16S rRNA gene sequence analysis. For instance, the Vibrio cholerae, Vibrio halioticoli, Vibrio harveyi and Vibrio splendidus species groups appeared in the atpA gene phylogenetic analyses, suggesting that these groups may be considered as separate genera within the current Vibrio genus. Overall, atpA gene sequences appeared to be more discriminatory for species differentiation than 16S rRNA gene sequences. 16S rRNA gene sequence similarities above 97% corresponded to atpA gene sequences similarities above 80%. The intraspecies variation in the atpA gene sequence was about 99% sequence similarity. The results showed clearly that atpA gene sequences are a suitable alternative for the identification and phylogenetic study of vibrios.

INTRODUCTION

Vibrios are ubiquitous in the marine environment and are found in high abundance, particularly in association with animals. Interactions with hosts can range from mutualistic, e.g. Vibrio Fischeri–squid, to pathogenic, e.g. Vibrio harveyi–shrimps. The genus Vibrio is the largest genus within the family Vibrionaceae, comprising 81 species. Thompson et al. (2004) suggested that this genus could be split in several genera to encompass recognizable species groups (Vibrio cholerae-, Vibrio halioticoli-, Vibrio harveyi- and Vibrio splendidus-species clades). The different species groups found within the genus Vibrio may have different ecological niches. For instance, the V. harveyi-species group is frequently associated with diseases in various aquatic organisms worldwide, notably in shrimp. Vibrio coralliilyticus is one of the aetiological agents involved in the phenomenon of coral bleaching (Ben-Haim et al., 2003).

The V. splendidus-species group is frequently involved with diseases in molluscs, while the V. halioticoli-species group comprises species that are potentially mutualistic organisms in abalones and also act as probionts in abalone rearing. V. cholerae, the aetiological agent of cholera, is considered one of the most serious human pathogens within the genus Vibrio.

The identification of vibrios at the family and genus levels can be achieved by comparative analysis of the 16S rRNA, recA, rpoA and pyrH gene sequences (Thompson et al., 2005), whereas identification at the species and strain levels requires the application of several genomic techniques, including DNA–DNA hybridization (DDH), repetitive extragenic palindromic PCR (REP-PCR), amplified fragment length polymorphism (AFLP) analysis and multilocus sequence analysis (MLSA) (Thompson et al., 2004, 2005). Identification of vibrios at the species level is not always trivial because several species of the genus Vibrio, e.g. V. cholerae–Vibrio mimicus, have almost identical gene sequences in several loci and around 70% DDH similarity (Thompson et al., 2005). In order to find new identification markers and to improve our knowledge of the phylogeny and taxonomy of vibrios, it is essential to analyse alternative loci.

The gene that codes for the α-subunit of bacterial ATP synthase (atpA) functions in ATP synthesis coupled to proton transport. In organisms with a respiratory chain,
the primary role of this enzyme is to couple the electrochemical potential difference across the inner membrane with the synthesis of ATP from ADP and phosphate. In bacteria that lack a respiratory chain, the role of this enzyme is to create a proton gradient and this process is then driven by ATP hydrolysis (Kasimoglu et al., 1996). Vibrios are facultative anaerobic organisms that synthesize ATP by aerobic respiration if oxygen is present but are also capable of switching to fermentation. The atpA gene is a housekeeping gene with ubiquitous distribution, functional constancy and is conserved among bacteria. This gene is located on chromosome I of vibrios. Studies based on whole genome sequences and on different loci of various bacterial groups have suggested that the atpA gene may be an alternative phylogenetic and identification marker for bacterial systematics (Zeiger, 2003; Naser et al., 2005).

Previous studies on the evolutionary history of the atpA gene suggested that an ancient gene duplication gave rise to three types of proton pumping ATPase genes: a vacuolar type in Eukarya and Archaea and Bacteria types. The duplicated genes were used to root the universal tree of life. However, the archaean type was occasionally observed in Bacteria and vice versa (Hilario & Gogarten, 1993). These authors suggested that horizontal gene transfer events could explain this grouping. More recently, Olendeznski et al. (2000) and Lapierre et al. (2006) showed that members of the family Deinococcaceae (e.g. the genera Thermus, Meiothermus and Deinococcus) contain ATPase genes typically found in Archaea or Eukarya which have probably been acquired by horizontal gene transfer. A Na⁺-ATPase ω-subunit of the Eukarya type was found in Enterococcus hirae (Kakinuma et al., 1991), but a detailed study on the identification of species of the genus Enterococcus based on atpA gene sequences (Naser et al., 2005) proved that this gene can be used for identification. The aim of this study was to analyse the usefulness of atpA gene sequences for reliable species identification of vibrios and for inferring their evolutionary history.

**METHODS**

The strains used in this study are listed in Supplementary Table S1 (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online). Strains were grown aerobically on tryptic soy agar (Difco) with 1.5 % (w/v) NaCl at 28°C (available in IJSEM Online).

**Phylogenetic analysis of vibrios**

Alignment of gene sequences was performed in Clustal W (Thompson et al., 1994). Phylogenetic analysis was conducted using MEGA version 3.1 (Kumar et al., 2004). The phylogenetic inference was based on the neighbour-joining (NJ) distance method (Saitou & Nei, 1987) and the maximum-parsimony (MP) character method. The NJ tree was constructed using the Kimura two-parameter model (Kimura, 1980). Correlations and a regression curve between atpA and 16 tRNA gene sequences were generated by Pearson’s product-moment correlation coefficient using BioNumerics software (Applied Maths). Recombination tests and the ratio of mean synonymous substitutions per synonymous site to mean non-synonymous substitutions per non-synonymous site (dS/dN) were calculated using the START software package obtained from http://pubmlst.org/software/analysis/start (Jolley et al., 2001). Split tree decomposition analysis was performed with publically available software (http://bibiserv.techfak.uni-bielefeld.de/splits/).

**RESULTS AND DISCUSSION**

The atpA gene sequences (1322 bp fragments corresponding to approximately 88% of the coding region) of 151 strains of the family Vibrionaceae were analysed. In order to

**Table 1. Amplification and sequencing primers**

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5’→3’)</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>atpA-01-F</td>
<td>CTDATTCHACNGAATYAGYG</td>
<td>37</td>
</tr>
<tr>
<td>atpA-02-F</td>
<td>GCNATGGGBGAATAYTCCG</td>
<td>760</td>
</tr>
<tr>
<td>atpA-03-R</td>
<td>CGGAAATTCVCXCCATNGC</td>
<td>760</td>
</tr>
<tr>
<td>atpA-04-R</td>
<td>TTACARGWYTTGGTGTC</td>
<td>1554</td>
</tr>
<tr>
<td>atpA-05-R</td>
<td>GCHAGHGCDGTACGRATACC</td>
<td>1204</td>
</tr>
<tr>
<td>atpA-06-F</td>
<td>AGCGGWCTRAYAARCGRG</td>
<td>55</td>
</tr>
</tbody>
</table>

and Shewanella oneidensis (AE014299). All reference sequences were acquired from publicly available databases from whole genome sequence projects.
Fig. 1. Phylogenetic tree based on the neighbour-joining method, using atpA gene sequences (1322 bp) of 151 representative strains of the family Vibrionaceae. Distance estimations were obtained by Kimura two-parameter model. Bootstrap percentages after 1000 simulations are shown. Bar, 2% estimated sequence divergence. *Shewanella oneidensis* LMG 19005 was included as an outgroup.
determine the intraspecies variation within this locus, 36 species were represented by more than one strain. The DNA G+C content of the atpA gene (47.9 mol%) was consistent with the mean for the total genomes of vibrios (Thompson et al., 2004). The $d_\text{s}$:$d_\text{n}$ ratio for the atpA gene for all strains analysed was 25.5, suggesting that this locus is under negative selective pressure. Correlations and a regression curve between pairwise similarities of atpA and 16S rRNA gene sequences were generated by Pearson’s product-moment correlation coefficient. A significant correlation was obtained between the two genetic loci ($R^2=0.85$). The atpA and 16S rRNA gene sequence datasets fit a polynomial regression of the second degree. The atpA gene sequences are more discriminatory than the 16S rRNA gene sequences. 16S rRNA sequence similarities above 97 % were clearly separated from the genus Vibrio. 16S rRNA sequence similarities above 97 % corresponded to atpA gene sequence similarities above 80 % (see Supplementary Fig. S1 in IJSEM Online).

The genera Enterovibrio, Grimontia and Photobacterium were clearly separated from the genus Vibrio in the NJ tree (Fig. 1). In addition, species groups were recognized within the genus Vibrio. These species groups may represent different genera awaiting formal taxonomic description. The V. harveyi-, Vibrio tubishii-, V. splendidus- and V. halioticoli-species groups formed separate cohesive groups in accordance with previous polyphasic systematic studies (Thompson et al., 2004, 2005). V. cholerae and V. mimicus appeared at the outskirts of the genus Vibrio in the NJ tree. Several Vibrio species formed deep branches, e.g. Vibrio cincinnatiensis, Vibrio aestuarianus, Vibrio rumonensis, Vibrio gallicus and Vibrio gazogenes, probably because these species may have accumulated a greater number of substitutions in their evolutionary history when compared with other species of vibrios. The MP tree (Supplementary Fig. S2 in IJSEM Online) based on the atpA gene sequences did not separate the genus Vibrio from the other genera of the family Vibrionaceae.

V. harveyi-, V. scopthalmi-, V. coralliilyticus-, V. splendidius-, V. mediterranei-, V. diazotrophicus-, V. furnissii-, V. anguillarum-, V. halioticoli-, V. fischeri- and V. cholerae-species groups were detected in both NJ and MP trees (Fig. 1 and Supplementary Fig. S2). The V. harveyi-species group comprised V. campbellii, V. rotiferianus, V. brasiiliensis, V. parahaemolyticus, V. mythili, V. natriegens, V. diabolicus and V. alginolyticus. These species shared at least 95 % atpA gene sequence similarity. The closest neighbour of V. harveyi was V. campbellii with 99 % sequence similarity, followed by V. rotiferianus (98 %) and V. parahaemolyticus (97 %). V. harveyi strains LMG 20370 and R-14913 clustered separately from the other V. harveyi strains. These two strains had about 70 % DNA–DNA similarity to V. harveyi LMG 4044 and were also found in AFLP clusters A30 and A31 in a previous study (Thompson et al., 2001). V. campbellii R-14902 clustered together with V. harveyi strains LMG 20370 and R-14913. In a former study, V. campbellii R-14902 showed a similar Biolog pattern to V. harveyi, suggesting that this strain belongs to V. harveyi (Vandenbergh et al., 2003). V. parahaemolyticus and V. alginolyticus were clearly separated, having around 97 % atpA gene sequence similarity.

V. scopthalmi and V. ichthyomenteri formed a group in both NJ and MP trees, sharing around 97 % atpA gene sequence similarity. The pair V. coralliilyticus and V. neptunius had 95 % atpA gene sequence similarity, while V. pacini strains formed a separate group in both NJ and MP trees. Previous analyses of 16S rRNA gene sequences of this species suggested that it was related to V. kanalaoae, V. pomeroyi and V. furnissii (Thompson et al., 2003). The group comprising V. tubishii, V. hepatarius and V. orientalis had around 97 % atpA gene sequence similarity and appeared scattered in both NJ and MP trees. For instance, V. brasilienis strains clustered with the V. harveyi-species clade, sharing around 96 % sequence similarity with this group. V. splendidus, V. tasmaniensis, V. kanalaoae, V. pomeroyi, V. cyclostrophicus, V. chagasii, V. pelagius and V. fortis formed a group and had atpA gene sequence similarities of around 97.5 %, except for V. chagasii which had 93.7 %. V. splendidus strains LMG 16748 and LMG 16752 clustered with V. kanalaoae, suggesting that these strains may belong to the latter species. V. mediterranei formed a group along with V. shilonii, sharing almost identical atpA gene sequences. V. diazotrophicus and V. hispanicus grouped together, having 96 % atpA gene sequence similarity. V. furnissii and V. fluvialis also formed a group. These two species have 94 % atpA gene sequence similarity. The group composed of V. anguillarum and V. ordalii had 98 % atpA gene sequence similarity.

The V. halioticoli-species group, i.e. V. neonatus, V. ezurae, V. superstes and V. gallicus had atpA gene sequence similarities of between 89 and 99.5 %. V. neonatus and V. ezurae were the closest neighbours of V. halioticoli (99.5 % and 98 %, respectively), but were separated in different clusters, suggesting that this locus can serve as an alternative means to identify this group. The V. halioticoli-species group had greater than 98 % 16S rRNA gene sequence similarity. The V. fischeri group, e.g. V. logei and V. wodanis, had between 89 and 93 % atpA gene sequence similarity and was detected in both the NJ and MP trees. V. cholerae and V. mimicus were clearly separated in two groups with a high (100 %) bootstrap value. These two species formed monophyletic groups within the genus Vibrio. The atpA gene sequence similarity values between V. cholerae and V. mimicus were at most 96 %, suggesting that this locus is an alternative marker for the identification of these species.

The majority of the species groups based on the atpA gene sequences appeared in both the NJ and MP trees and were in agreement with 16S rRNA gene sequence analysis. Overall, the NJ trees constructed on the basis of the 16S rRNA and atpA gene sequences showed similar topologies. Most species groups and genera were congruent for both markers, but some differences were observed. For instance, V. brasilienis appeared within the V. harveyi group in the atpA gene sequence tree and within the V. tubishii group.
in the 16S rRNA gene sequence tree (Thompson & Swings, 2006).

In order to check for possible recombination events in the atpA gene, we simultaneously analysed the 151 atpA gene sequences. Only Sawyer’s test, based on the sum of the squares of condensed fragments (P<0.01) detected possible recombination events in several strains, including V. wodanis strains K16 and K26, V. fischeri ATCC 700601, V. neonatus LMG 19976 and V. scopphalimi strains LMG 19158T and R-15029, V. alginolyticus strains LMG 4407 and LMG 4409T, V. cholerae R-18303 and V. mimicus R-20565, Enterovibrio corallii LMG 22228T and V. cholerae R-18303, Enterovibrio corallii LMG 22228T and V. mimicus R-20565 and V. campbellii LMG 20369 and V. wodanis K16. We also generated a split tree decomposition analysis for all strains. The split tree showed massive networks that are hallmarks of recombination, but the Phi test did not provide statistical support for this conclusion. Several different detection methods should provide evidence that recombination did occur within the atpA gene sequences. As the Sawyer’s test only detected recombination events, we can assume that recombination is not a common evolutionary process in the atpA gene of vibrios.

The formal delineation of species of vibrios still relies on DDH similarity, but we have encouraged the use of gene sequences as a reliable alternative towards the establishment of an online electronic taxonomy (Thompson & Swings, 2006). This study shows the usefulness of atpA gene sequences for discriminating species groups and species and should lead to a refinement of the online electronic taxonomy of vibrios (http://www.taxvibrio.lnc.br/). Work is under way to analyse other genetic loci that will add even more discriminatory power in order to distinguish closely related species. Eventually widespread vibrio clones will be identified through the internet.

ACKNOWLEDGEMENTS

C. C. T. acknowledges a PhD scholarship from Conselho Nacional de Desenvolvimento Cientifico e Tecnologico (CNPq, Brazil) and the financial support of Instituto Oswaldo Cruz (FIOCRUZ, Rio de Janeiro, Brazil). F. L. T. acknowledges grants from FAPESP, FUJB, IFS and CNPq.

REFERENCES


