Systematic and Applied Microbiology 39 (2016) 77–87

Contents lists available at ScienceDirect

Systematic and Applied Microbiology

journal homepage: www.elsevier.de/syapm

The GyrA encoded gene: A pertinent marker for the phylogenetic revision of *Helicobacter* genus☆

Armelle Ménard[a,b,*], Alice Buissonnière[a,b], Valérie Prouzet-Mauléon[a], Elodie Sifré[a,b], Francis Mégraud[a,b]

[a] *Université de Bordeaux, Laboratoire de Bactériologie, Centre National de Référence des Helicobacters et Campylobacters, F33076 Bordeaux, France*[1]
[b] *INSERM U853, F33076 Bordeaux, France*

ARTICLE INFO

Article history:
Received 17 February 2015
Received in revised form 30 July 2015
Accepted 18 September 2015

Keywords:
Helicobacter
Taxonomy
Phylogeny
Gyrase
Ribosomal gene
Chaperonin

ABSTRACT

Phylogeny of *Epsilonproteobacteria* is based on sequencing of the 16S rRNA gene. However, this gene is not sufficiently discriminatory in *Helicobacter* species and alternative markers would be useful. In this study, the 16S rRNA, *gyrA*, *hsp60*, *gyrB*, and *ureA-ureB* gene sequences, as well as GyrA, HSP60 and GyrB protein sequences were analyzed as tools to support *Helicobacter* species phylogeny: 72 *Helicobacter* strains, belonging to 41 species of which 36 are validated species, were included. Results of the phylogenetic reconstructions of the GyrA gene encoded protein (approximately 730 residues) indicated the most stable trees to bootstrap resampling with a good separation of *Helicobacter* taxa, especially between gastric and enterohepatic species. Moreover, the GyrA tree revealed high similarity with that of the *gyrB* and *ureA-ureB* genes (restricted to urease-positive *Helicobacter* species). However, some differences in clustering were observed when compared to the *hsp60* and 23S rRNA gene trees. Altogether, these revised phylogenies (except the 16S rRNA gene for enterohepatic Helicobacters) enabled reliable clustering of *Helicobacter cinaedi* and '*Flexispira*' strains, determined a reliable position for *Helicobacter mustelae* (except the *hsp60* gene) and for novel *Helicobacter* species proposed such as '*Helicobacter sanguini*', '*Helicobacter apodemus*' or '*Helicobacter winghamensis*', and suggest that *Helicobacter* species MIT 09-6949 and MIT 05-5293 isolated from rodents constitute novel species. Although they are not commonly used to study the phylogeny of *Epsilonproteobacteria*, protein sequences and, in particular, the GyrA protein sequence may constitute pertinent phylogenetic markers for *Helicobacter* genus.

Introduction

The *Helicobacteraceae* family consists of six genera, *Helicobacter*, *Sulfuricurvum*, *Sulfurimonas*, *Sulfurovum*, *Thiovulum* and *Wolinella*. This family is dominated by the *Helicobacter* genus, which is currently comprised of 36 validated named species [14]. New *Helicobacter* species are regularly isolated but cannot be easily identified and some proposed species whose names have not been officially validated are awaiting confirmation (e.g. '*Helicobacter winghamensis*', '*Helicobacter sanguini*', etc.).

Advances in the phylogeny of the *Helicobacter* genus are faced with the low discriminatory power of the 16S rRNA gene, the "gold standard" gene used to study bacterial phylogeny. Indeed, the 16S rDNA is not sufficient for the species-level identification of *Helicobacteraceae* [45] since it can lead to misidentification [50], which is consistent with the horizontal transfer of 16S rRNA gene fragments and the creation of mosaic molecules with a loss of phylogenetic information [10]. Consequently, a polyphasic taxonomic analysis was proposed to describe interstrain relationships between *Helicobacteraceae* [45]. Unfortunately the biochemical identification of *Helicobacter* genus strains based on the limited number of commonly used tests remains difficult and relevant phylogenetic markers must be defined. Advances in the phylogeny of the *Helicobacter* genus are also faced with the small number of species included in previous studies to evaluate taxonomic markers [3,4,10,24,28,37,38,51]. Thus, phylogenetic markers other than the 16S rDNA must be identified and tested on a large number of species to confirm and update taxonomic identification in the *Helicobacter* family.

☆ The accession numbers of nucleotide and protein sequence data are available in Supplementary Table S1.

* Corresponding author at: INSERM U853, Université de Bordeaux, Laboratoire de Bactériologie, Bat. 2B RDC Zone Nord, 33076 Bordeaux cedex, France. Tel.: +33 5 57 57 12 88; fax: +33 5 56 51 41 82.
E-mail address: armelle.menard@u-bordeaux.fr (A. Ménard).

[1] http://www.cnrch.u-bordeaux2.fr/

http://dx.doi.org/10.1016/j.syapm.2015.09.008
0723-2020/

DNA gyrase is a quadrimeric protein composed of two subunits, GyrA and GyrB, which introduce negative supercoils into DNA. This essential bacterial enzyme is found in all bacteria and is conserved among species, making it an attractive target for phylogenetic studies. The *gyrB* gene has been used extensively to determine bacterial phylogeny, more so than the *gyrA* gene which has nevertheless been reported to be sufficiently conserved over a long period of evolution [2]. Consequently, the *gyrA* gene was employed for the phylogenetic identification of *Campylobacter* and *Arcobacter* species [1,30,32,35].

In the present study, the phylogeny of the *Helicobacter* genus was investigated using the *gyrA* gene from 72 *Helicobacter* strains belonging to 41 species, including all 36 currently validated species, 5 *Candidatus* strains and 7 unclassified *Helicobacter* strains, and compared to that of 16S rRNA, *hsp60* (heat-shock protein) and *gyrB* genes, as well as *ureA-ureB* genes (mainly for gastric *Helicobacter* species). The GyrA, HSP60 and GyrB protein encoded genes were also investigated.

Materials and methods

Bacterial strains and gene sequences

Campylobacter jejuni, *Mycobacterium tuberculosis*, *Wolinella succinogenes* and 72 *Helicobacter* strains (60 strains with validated names, 5 *Candidatus* strains and 7 unclassified *Helicobacter* strains) were included in the study. Their origin and associated pathologies are described when available and are referenced in Supplementary Table S1.

Gene sequence downloading and sequencing are described in Supplementary Materials. GenBank accession numbers are available in Supplementary Table S1.

Phylogenetic analyses

The gene and protein sequences were aligned using the multiple alignment options on ClustalW. Phylogenetic tree topologies of nucleotide and predicted amino acid alignments were constructed using the MEGA (Molecular Evolutionary Genetics Analysis) software version 6.06 [47]. 23S rRNA gene analyses were performed using the "Quick BioInformatic Phylogeny of Prokaryotes" (leBIBI) software version 1.0.1 available on-line [8].

Development of a PCR for Helicobacter species identification

The *gyrA* gene sequences of *Helicobacter species* ($n = 72$) were aligned [6] and analyzed to identify conserved regions for the design of 3 PCR primer sets (Table 1). Amplification parameters consisted of 1 cycle at 95 °C for 5 min, followed by 40 cycles at 95 °C for 30 s, 54 °C for 30 s, and 72 °C for 1 min 30 s, and finally 1 cycle at 72 °C for 5 min. The PCR products were sequenced using PCR primers.

Table 1
Primers designed for the *gyrA* gene sequence amplification of *Helicobacter* species.

Primer names	Primer sequences (5′ → 3′)	Expected PCR product(s)
PCR no. 1		
F1-QRDR-Hspe	CATAGGCGTATHYTDTATGCVATG	
F2-QRDR-Hspe	TGGGTGATGTRATYGGTAAAT	1,105 bp or/and 939 bp
R1-QRDR-Hspe	TGATTAAGCCCTCYAARATATG	
PCR no. 2		
F1-2-gyrA-Hspe	CAACAGGTGGDATYATHTATGG	
R1-2-gyrA-Hspe	GGTYGCCATRATYTTTTCAY	1,142 bp
R1-2B-gyrA-Hspe	GGTYGCCATRATYTTTTCAT	
PCR no. 3		
F3-4-gyrA-Hspe	GAGGGACCWGAYTTYCCMAC	
R3-4-gyrA-Hspe	GGATTTTATAMACYTTKAGCCAAT	1,036 bp
R3-4B-gyrA-Hspe	GGATTTTATAMACYTTKAGCCART	

These sets of primers were designed using the web Primer3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi).

Results

In the present study, 16S rRNA, *gyrA*, *hsp60* (*groEl* or *cpn60*), and *gyrB* gene sequences were used to analyze evolutionary relationships among 72 *Helicobacter* strains (presented in Supplementary Table S1). The *ureA-ureB* genes were also used for Helicobacters harboring these genes (essentially gastric *Helicobacter* species). In addition, the GyrA, HSP60 and GyrB protein encoded genes were also analyzed.

Some of these markers as well as 23S rDNA, *rpoB-rpoC* and *atpA* (ATP synthase α subunit) genes have already been used to discriminate between *Helicobacter* species [3,4,10,24,28,37,38,51]. However, most of these studies, except the ones based on ribosomal genes [10], only included a few *Helicobacter* species. Thus the phylogenetic reconstructions obtained in the present study were primarily compared to those obtained by Dewhirst et al. [10]. Similar parameters were used: trees were not rooted and they were constructed using the neighbor-joining method and bootstrapping.

Phylogenetic reconstruction – general observations

Low bootstrap values were present at the main nodes of the 16S rRNA gene tree (indicated by slashes in Fig. 1), demonstrating that the tree was not stable to bootstrap resampling and confirming that this gene is not informative enough to be used for *Helicobacter* taxonomy [10,50]. No cluster was reliably individualized for unsheathed and sheathed enterohepatic Helicobacters (see cluster no. 2 and *Helicobacter bilis* strains). Another illustration of the low discriminatory power of the 16S rRNA gene for *Helicobacter* is the position of *Helicobacter equorum* (a Helicobacter with a single sheathed flagellum) in cluster no. 2 (comprised of unsheathed Helicobacters); this position was nonetheless consistent with a previous 16S rRNA gene sequence analysis [39].

The *gyrA* gene sequence analysis (Supplementary Fig. S1) did not yield trees as robust as those obtained with the corresponding protein (Fig. 2, approximately 730 residues, i.e. 87–90% of the GyrA protein's total length). On the other hand, the phylogenetic analyses of the *hsp60*, *gyrB*, and *ureA-ureB* genes (Figs. 3–5) led to more robust trees than those obtained with their corresponding proteins (Supplementary Figs. S2 and S3).

GyrA protein, *gyrB* and *hsp60* gene sequence analyses showed clusters different than those obtained from the 16S rRNA gene sequence analysis (Figs. 1–4). Overall the GyrA- and *gyrB*-resulting trees were similar to that of the 23S rRNA gene [10], i.e. they show 2 main branches separating Helicobacters of gastric origin from those of enterohepatic origin (Figs. 2–4) while *hsp60* tree topology did not: cluster no. 5 is out of the main branch of enterohepatic Helicobacters (Fig. 3).

Phylogenetic reconstruction of gastric Helicobacter *species (cluster no. 1)*

Gastric *Helicobacter* species are usually divided into 2 main phylogenetic branches with subclusters [3,4,10,24,38]. As expected, all 17 gastric *Helicobacter* strains (belonging to 11 species) included in the present study were grouped into a cluster composed of 2 branches (Figs. 1–5), i.e. cluster no. 1 according to the numbering of previously reported clusters [10]. Clustering of these gastric