5S Ribosomal Ribonucleic Acid Sequences in Bacteroides and Fusobacterium: Evolutionary Relationships within These Genera and among Eubacteria in General

HILDE VAN DEN EYNDE, RAYMOND DE BAERE, HAROUN N. SHAH, SAHEER E. GHARBA, GEORGE E. FOX, JOANNA MICHALIK, YVES VAN DE PEER, AND RUPERT DE WACHTER

Department Biochemie, Universiteit Antwerpen (UIA), Universiteitsplein 1, B-2610 Antwerp, Belgium; Department of Oral Microbiology, The London Hospital Medical College, Whitechapel, London E1 2AD, United Kingdom; and Department of Biochemical and Biophysical Sciences, University of Houston, Houston, Texas 77004

The 5S ribosomal ribonucleic acid (rRNA) sequences were determined for Bacteroides fragilis, Bacteroides thetaiotaomicron, Bacteroides capillosus, Bacteroides veroralis, Porphyromonas gingivalis, Anaerorhabdus furcosus, Fusobacterium nucleatum, Fusobacterium mortiferum, and Fusobacterium varium. A dendrogram constructed by a clustering algorithm from these sequences, which were aligned with all other hitherto known eubacterial 5S rRNA sequences, showed differences as well as similarities with respect to results derived from 16S rRNA analyses. In the 5S rRNA dendrogram, Bacteroides clustered together with Cytophaga and Fusobacterium, as in 16S RNA analyses. Intraphyllum relationships deduced from 5S rRNAs suggested that Bacteroides is specifically related to Cytophaga rather than to Fusobacterium, as was suggested by 16S RNA analyses. Previous taxonomic considerations concerning the genus Bacteroides, based on biochemical and physiological data, were confirmed by the 5S rRNA sequence analysis.

The genus Bacteroides as described by Holdeman et al. (12) consists of more than 40 species which vary considerably in physiological, biochemical, and chemical properties and possess deoxyribonucleic acid (DNA) base compositions within the range from 28 to 61 mol% guanine plus cytosine (G+C). It is now generally accepted that a DNA base composition difference of more than 10 mol% G+C implies that the species concerned are unrelated at the generic level, yet notable exceptions occur (25). On the basis of physiological and biochemical properties, such as peptidoglycan composition, acid end products, lipids, and fatty acids (for a review see reference 30), various proposals have been made to exclude certain species from the genus Bacteroides (3, 30). Many of these proposals are in good agreement with ribosomal ribonucleic acid (rRNA) homology data (17). Thus, many of the species listed in Bergey's Manual of Systematic Bacteriology (12) have now been given separate generic status (1, 2, 4, 23, 28, 29, 31–33). Recently, it has been suggested (3) that only members of the so-called "Bacteroides fragilis group" (viz., Bacteroides fragilis, Bacteroides distasonis, Bacteroides thetaiotaomicron, Bacteroides eggerthii, Bacteroides ovatus, Bacteroides uniformis, Bacteroides vulgatus, Bacteroides coccaceae, Bacteroides merdae, and Bacteroides stercoreis (3, 18)) be retained in the genus Bacteroides, as these species are biochemically and physiologically fairly homogeneous. This group consists of obligately anaerobic, gram-negative, nonsaporeforming rods that produce mainly acetic and succinic acids and contain malate, glutamate, glucose-6-phosphate, and 6-phosphogluconate dehydrogenases. DNA base compositions range between 40 and 48 mol% G+C, and all species possess menaquinones, sphingolipids, and predominantly straight-chain saturated, anteiso- and iso-methyl-branched long-chain fatty acids (3). Therefore, in this study only species of the latter group were considered members of the genus Bacteroides. The genus Bacteroides as defined by Holdeman et al. (12) is referred to below as "Bacteroides," for as mentioned above, many of its species have now been given separate generic status.

Hitherto, 13 oligonucleotide catalogs (26), two partial 16S rRNA sequences (25), and two full 16S rRNA sequences (21, 38) have been determined for members of the genus "Bacteroides." The results have grouped these organisms within the genera Cytophaga, Flavobacterium, Flexibacter, Saprospira, Sporocytophaga, Helicomencebacter, and Fusobacterium (26, 38), thus forming 1 of the 10 major eubacterial phyla proposed by Woese (40). Bacteroides amylophilus, Bacteroides succinogens, Bacteroides termitidis, Bacteroides gracilis, and Bacteroides ureolyticus did not cluster with this phylum on the basis of 16S rRNA analyses (21, 25, 26). These species were shown previously (3, 30) to possess biochemical and physiological properties that are atypical of species of the genus Bacteroides. Since these studies B. termitidis, B. succinogens, and B. amylophilus have been reclassified into new genera (1, 23, 33), while inconsistencies in the present classification of B. gracilis and B. ureolyticus have been pointed out (3, 25).

The genus Fusobacterium, as presently constituted (24), also appears to be heterogeneous (10). Apart from its wide range in DNA base composition (25 to 57 mol% G+C) (20, 24), differences have been noted in peptidoglycan composition (19, 22, 36). Members of this genus are distinguishable from members of the B. fragilis group by such properties as the production of major amount of butyric acid, the lack of methyl-branched fatty acids (15), and the absence of malate, glucose-6-phosphate, and 6-phosphogluconate dehydrogenases (10).

The phylogenetic groups proposed on the basis of 16S rRNA analyses have not been put to the test through a comparison with the groups obtained by analyzing the smaller 5S rRNA sequences. We report here the 5S rRNA sequences of B. fragilis, the type species of the genus, B. thetaiotaomicron, Bacteroides capillosus, Bacteroides veroralis of the Bacteroides melaninogenicus-Bacteroides oralis group of species, Porphyromonas gingivalis (formerly Bacteroides gingivalis (32)) of the pigmented asaccharolytic...
FIG. 1. Alignment of 5S rRNA sequences determined in this study. The blank alignment positions are necessary to accommodate nucleotides of 5S rRNAs from other taxonomic groups. Length heterogeneity is indicated by terminal residues printed in lower-case letters. Boxes labeled A and A', B and B', etc., enclose complementary strands of helices A, B, etc., in the secondary-structure model (8). Bulges within the helices are enclosed in nested boxes; bases belonging to odd base pairs (pairs other than A·U, G·C, and G·U) intercalated between Watson-Crick or G·U pairs or both are in parentheses.
Actinomycetes and relatives (23)

Thermobacteroides acetotyphus

Photorhabdus luminescens

Vibrio vulnificus

Kapleniella islae (23)

Actinobacteria (2)

Purple y bacteria (78)

Purple a bacteria (47)

Plastids and Cyanobacteria (23)

Plant mitochondria (4)

Corynebacteria (2)
group, and Anaerorhabdus furcosus (formerly Bacteroides furcosus) [14, 31]. Furthermore, we determined the 5S rRNA sequences of two strains of Fusobacterium mortiferum, Fusobacterium varium, and Fusobacterium nucleatum. As far as we are aware, these species are the first representatives of their phylum for which 5S rRNA sequences are known. We constructed a tree from these sequences along with 270 other eubacterial 5S rRNA sequences, thus covering 8 of the 10 major phyla.

**MATERIALS AND METHODS**

The following strains were used: *B. fragilis* ATCC 25285T (T = type strain), *B. thetaiotaomicron* ATCC 29148T, *B. verorulis* ATCC 33779T, *B. capillosus* ATCC 29799T, *P. gingivalis* ATCC 33277T, *A. fuscosus* ATCC 25662T, *F. varium* VP1 0499A, *F. nucleatum* 85-56A, and *F. mortiferum* ATCC 25557T and VP1 4249. The test strains were maintained by weekly subculturing on 5% horse blood agar plates (blood agar base [Difco Laboratories, Detroit, Mich.]). All strains were grown at 37°C in liquid culture in an atmosphere containing 70% (vol/vol) nitrogen, 20% (vol/vol) hydrogen, and 10% (vol/vol) carbon dioxide. Liquid medium (medium BM) contained (per liter) 10 g of Trypticase [BBL Microbiology Systems, Cockeysville, Md.], 10 g of proteose peptone (Oxoid Ltd., Basingstoke, Hampshire, United Kingdom), 5 g of glucose, 5 g of yeast extract (Difco), and 5 g of sodium chloride. The medium was adjusted to pH 7.4 and autoclaved at 121°C for 15 min. Filter-sterilized hemin (Sigma Chemical Co., St. Louis, Mo.) and menadione (Calbiochem-Behring, La Jolla, Calif.) solutions were added aseptically (12) to give final concentrations of 5 µg of hemin per ml and 0.5 µg of menadione per ml. Broth cultures were checked for purity by plating onto 5% horse blood agar, and identities were confirmed by Gram staining and the API anaerobic system. Cultivated organisms were harvested by centrifugation, washed with distilled water, and freeze-dried.

5S rRNA was prepared as described previously (35, 39). On average, 100 U of absorbance at 260 nm of RNA and subsequently 1.4 U of absorbance at 260 nm of 5S rRNA were obtained from 0.1 g of lyophilized cells. 5S rRNA was 3'-terminally labeled with cytidine 3',5'-bisphosphate and purified on 8% polyacrylamide gels (27). The Peattie partial chemical degradation method (27) followed by electrophoresis on 8 to 20% polyacrylamide gels generally made it possible to elucidate the entire sequence. In the case of *B. fragilis* and *B. verorulis*, the identity of the 5'-terminal nucleotide was confirmed by gel electrophoresis of enzymatically degraded (6), 5'-terminally labeled (34) 5S rRNA, whereas in the case of *A. furcosus* ligation of (Ap)A to the 5' terminus (6) was followed by 3'-terminal labeling and partial chemical degradation. The 5S rRNA of *F. mortiferum* ATCC 25557T was shown to be heterogeneous in length. Components having different chain lengths were sequenced separately.

Dendrograms were constructed by using a clustering algorithm (7, 13). In short, this algorithm calculated a dissimilarity matrix for all of the sequences involved, while corrections were used to meet the effects of multiple substitutions and unequal evolutionary rates. A phenogram was then constructed by using weighted pairwise grouping. The standard deviations on the substitution terms of the computed dissimilarities were indicated by error flags on all of the branching points on the resulting dendrograms.

**RESULTS AND DISCUSSION**

Figure 1 shows part of an alignment (available from us) in which the newly determined sequences were aligned with more than 500 5S rRNA sequences which were eubacterial, archaebacterial, organelle, and eucaryotic in nature. For an update of all previously published 5S rRNA sequences, see reference 9. Boxes enclosing putative double-stranded areas were superimposed on the alignment, conforming to a sec-

---

**TABLE 1. Major biochemical and chemical characteristics of the “Bacteroides” species tested**

<table>
<thead>
<tr>
<th>Species</th>
<th>G+C content (mol%)</th>
<th>Dehydrogenases</th>
<th>Metabolism</th>
<th>Major metabolic end products</th>
<th>Major long-chain fatty acids(s)</th>
<th>Predominant menaquinone(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>MDH GDH G6PDH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>B. fragilis</em></td>
<td>42</td>
<td>+ + + + +</td>
<td>F</td>
<td>Acetic, succinic acids</td>
<td>Anteiso-C15:0 MK10, MK11</td>
<td></td>
</tr>
<tr>
<td><em>B. thetaiotaomicron</em></td>
<td>42</td>
<td>+ + + + +</td>
<td>F</td>
<td>Acetic, succinic acids</td>
<td>Anteiso-C15:0 MK10, MK11</td>
<td></td>
</tr>
<tr>
<td><em>B. verorulis</em></td>
<td>42</td>
<td>+ + + + +</td>
<td>NF</td>
<td>Butyric, acetic acids</td>
<td>Iso-C15:0 MK9</td>
<td></td>
</tr>
<tr>
<td><em>P. gingivalis</em></td>
<td>46-48</td>
<td>+ + + + +</td>
<td>MF</td>
<td>Acetic, succinic acids</td>
<td>Anteiso-C15:0 MK10</td>
<td></td>
</tr>
<tr>
<td><em>B. capillosus</em></td>
<td>60</td>
<td>+ + + + +</td>
<td>NF</td>
<td>Acetic, succinic acids</td>
<td>C14:0, C16:0</td>
<td>Absent</td>
</tr>
<tr>
<td><em>A. fuscosus</em></td>
<td>34</td>
<td>+ + + + +</td>
<td>WF</td>
<td>Lactic acid</td>
<td>C18:1, C18:0</td>
<td>Absent</td>
</tr>
</tbody>
</table>

a Data from references 3, 12, 16, 18, 30, 32, and 37.
b MDH, Malate dehydrogenase; GDH, glutamate dehydrogenase; G6PDH, glucose-6-phosphate dehydrogenase; 6PGDH, 6-phosphogluconate dehydrogenase.
+ Present; - absent.
c F, Fermentative; MF, moderately fermentative; WF, weakly fermentative; NF, nonfermentative.
d C14:0, Tetradecanoic acid; C16:0, hexadecanoic acid; C18:0, octadecanoic acid; C18:1, octadecenoic acid; iso-C15:0, 13-methyltetradecanoic acid; anteiso-C15:0, 12-methyltetradecanoic acid.
* Low levels of both acids.

---
ondary-structure proposal (8) applicable to SS rRNAs of all phylogenetic lineages.

We constructed a dendrogram (Fig. 2a), starting from an alignment containing the 276 eubacterial SS rRNA sequences known to us at present. The great majority of these sequences are found in reference 9. Included in these sequences, were 13 SS rRNA sequences of species and strains belonging to the genera “Bacteroides” sensu Holdeman et al. (12) (this paper), Fusobacterium (unpublished data), and Cytophaga (unpublished data), genera that were allotted to the Bacteroides-Cytophaga-Flavobacterium cluster on the basis of 16S rRNA analyses (26, 38). In the case of “Bacteroides,” care was taken to choose representatives from both the B. fragilis group and other non-B. fragilis species. Furthermore, 259 sequences of eubacteria belonging to seven different phyla and sequences of four plant mitochondria were taken into account. The dendrogram shows a number of clusters, one of which comprises all but two (A. fucosus and Cytophaga johnsonae [see below]) of the species expected to constitute the “Bacteroides”-Cytophaga-Flavobacterium cluster; the detailed structure of this cluster is shown in Fig. 2b. Except for effects due to the numerous sequences added, in particular the purple subd-"vion, 16S rRNA data set, and the entry for the first time of sequences from members of the phylum of Planctomycyes and relatives, the composition and topology of the other eubacterial clusters has not changed much in comparison with previous trees (5, 7, 13, 35). However, the overall topology of the dendrogram differs substantially from the topologies of dendrograms published previously, in that the position of the clusters with respect to one another tends to be altered with the addition or removal of sequences as well as with the nature of the reference group chosen (data not shown). Tree topology not being stable is not really unexpected in view of the error flags indicated on the branching points of the dendrogram (Fig. 2a).

When the pattern deduced from the 5S rRNA sequence information was compared with the pattern based on 16S rRNA analyses (40), certain similarities, as well as differences, became apparent. In the 5S rRNA tree (Fig. 2a), cyano bacteria and plastids are found as a monophyletic group, although they tend to branch off earlier in time than the branching deduced on the basis of 16S rRNA analyses. From Fig. 2a, the existence of subdivisions α, β, and γ of the phylum of purple photosynthetic bacteria and relatives is evident. However, this cluster also includes the phylum of radioresistant bacteria and relatives, which contains the genera Deinococcus and Thermus. From these studies there is no evidence for a fourth subdivision of purple bacteria and relatives, subdivision δ (40). There is no relationship between the 5S rRNA of plant mitochondria and subdivision α. Campylobacter species, suggested to be weakly related to the phylum of purple bacteria and relatives, show no relationship to any of the eubacterial clusters on the basis of 5S rRNA sequences. The phylum of gram-positive bacteria and relatives, a monophyletic group with four subdivisions according to Woese (40), shows considerable aberrations on the basis of 5S rRNA analysis. A substantial number of 5S rRNA sequences is known for the two most important subdivisions, namely, the gram-positive bacteria with low G+C contents and the gram-positive bacteria with high G+C contents. In the 5S rRNA tree, representatives of the low-G+C-content subdivision are split into two clearly distinct groups. One consists mainly of bacilli and clostridia, while the other encompasses predominantly mycoplasmata and lactobacilli. In the 5S rRNA tree, these two clusters are connected to the cluster containing gram-positive bacteria having high G+C contents and to phyla such as the Planctomycyes and their relatives and the “Bacteroides”-Cytophaga-Fusobacterium cluster. A number of members of underrepresented phyla are associated with some of these groups. In general, with few exceptions, the subdivisions proposed on the basis of 16S rRNA analyses compare well with the subdivisions derived from 5S rRNA sequences. However, it is difficult to predict whether the differences or similarities in branching order at the phylum and subdivision levels observed when 16S and 5S rRNA trees are compared will fade or will be confirmed as additional 5S rRNA sequences become available, making the data set more equivalent to the 16S rRNA data set.

Within the “Bacteroides”-Cytophaga-Fusobacterium cluster, two main groups are distinguishable (Fig. 2b). One contains the 5S rRNAs of F. nucleatum, F. mortiferum ATCC 25557T and VPI 4249, and F. varium and, atypically, the 5S rRNA of Streptococcus lactis subsp. cremoris, which one would expect to cluster with the mycoplasmata and lactobacilli. 5S rRNA sequences (5, 40). Unexpectedly, the 5S rRNA of Mycoplasma hyopneumoniae was also recovered close to the “Bacteroides”-Cytophaga-Fusobacterium cluster (Fig. 2a). The second group contains 5S rRNAs from B. fragilis, B. veroralis, B. thetaiotaomicron, B. capillosus, P. gingivalis (this paper), Cytophaga aquatilis, Cytophaga hectorina (unpublished data), and, atypically, Chlorobium limicola and another Chlorobium strain (N. R. Pace, personal communication). This is a clear discrepancy with the picture obtained from 16S rRNA analyses, for based on cataloging results from four species (11) and information from a full 16S rRNA sequence (40), the genera Chlorobium and Chlororherpeton are considered a tight phylogenetic unit, which is assigned the rank of phylum. Information from 16S rRNA catalogs (26) has grouped species of the genus “Bacteroides” with species of the genus Fusobacterium, thus forming one of two major phylum subgroups; the other consists of the genera Cytophaga, Flavobacterium, Flexibacter, Synechococcus, Halococcus, and Synechocystis. In this paper we present a somewhat different view since 5S rRNA analysis groups “Bacteroides” with Cytophaga, exclusive of Fusobacterium.

The data for the “Bacteroides” species are in excellent agreement with the recently refined proposal (3, 30) suggested for the reclassification of the genus. On the basis of 5S rRNA analyses (Fig. 2b) the two representatives of the B. fragilis group (viz., B. fragilis and B. thetaiotaomicron) are more related to one another than to B. veroralis and P. gingivalis. B. capillosus is only distantly related to the members of the B. fragilis group, while A. fucosus is totally unrelated to the “Bacteroides”-Cytophaga-Fusobacterium cluster (Fig. 2a). The latter is grouped with the mycoplasmata and lactobacilli cluster. Therefore, these results reinforce the chemical, biochemical, and physiological data of the species previously reported (3, 30). These data are summarized in Table 1.

Previous studies in which rRNA cataloging was used have highlighted the heterogeneity of the genus “Bacteroides” (26). According to these studies, B. fragilis, B. thetaiotaomicron, B. ovatus, B. vulgatus, B. uniformis, and B. distasonis of the B. fragilis group group together with Porphyromonas asaccharolyticus (formerly Bacteroides asaccharolyticus [32]) of the pigmented asaccharolytic group, to the exclusion of Bacteroides ruminicola and Bacteroides melaninogenicus of the Bacteroides melaninogenicus-Bacteroides oralis group of species, to whom they appear to be somewhat more
distantly related. The results of our study, which involved more species from outside the *B. fragilis* group, are in good agreement with the results of this study (26) and other chemotaxonomic data (for a review see reference 3).

The narrow range of DNA base compositions (28 to 32 mol% G+C) for most *Fusobacterium* species (10, 20, 24), together with the cellular fatty acid compositions (15) and metabolic end products (24), suggests that this genus forms a coherent group of species. However, studies on the peptidoglycan compositions of *Fusobacterium* species have revealed significant differences within the genus. Some species, such as *F. nucleatum* and *Fusobacterium necrophorum*, possess a directly cross-linked peptidoglycan based upon the amino acid lantionine (A18) (19, 22, 36), whereas other species, such as *F. varium* and *F. mortiferum*, contain diaminopimelic acid as their dabisacemic acid (A1γ) (22). These two groups of species are also discernible by enzyme electrophoretic patterns (10). Our SS rRNA data, although limited to a few species, further support the peptidoglycan and enzyme data in delineating the same two groups of species (Fig. 2b). Thus, *F. nucleatum*, the type species, is phylogenetically more distant to related to other species, which must cast doubt on its selection as a good centrospecies. Furthermore, we recently reported considerable heterogeneity within this species (10), and DNA-DNA hybridization studies are now in progress to clarify the intraspecies relatedness of strains labeled *F. nucleatum*. The high affinity between some strains of *F. mortiferum* and *F. varium* is supported by the present SS rRNA data (Fig. 2b). The division of *F. mortiferum* into two groups is in accord with the two peptidoglycan chemotypes reported for this species (22).

The results of our study, although limited to a few representative but well-characterized species of the *Bacteroidae*, essentially confirm the results of previous taxonomic studies (for a review see reference 3) and provide further support for several of the newly established genera.

ACKNOWLEDGMENTS

We thank Norman R. Pace for providing unpublished data. Our research was subsidized in part by a grant from the Fonds voor Kollektief Fundamenteel Onderzoek.

LITERATURE CITED


25. Paster, B. J., and F. E. Dewhirst. 1988. Phylogeny of campylo-


