Critical Reviews in Microbiology, 30:123–143, 2004

ISSN: 1040-841X print / 1549-7828 online
DOI: 10.1080/10408410490435133

Taylor & Francis
Taylor & Francis Group

The Phylogeny and Signature Sequences Characteristics of *Fibrobacteres, Chlorobi,* and *Bacteroidetes*

Radhey S. Gupta
Department of Biochemistry, McMaster University, Hamilton, Ontario, Canada

Fibrobacteres*, *Chlorobi*, and *Bacteroidetes* (FCB group) comprise three main bacterial phyla recognized on the basis of 16S rRNA trees. Presently, there are no distinctive biochemical or molecular characteristics known that can distinguish these bacteria from other bacterial phyla. The relationship of these bacteria to other phyla is also not known. This review describes many signatures, consisting of defined and conserved inserts in widely distributed proteins, that provide distinctive molecular markers for these groups of bacteria. These signatures serve to clarify the evolutionary relationship between members of the FCB group, and to other bacterial phyla. A 4 aa insert in DNA Gyrase B (GyrB) and a 45 aa insert in the SecA proteins are uniquely shared by various *Bacteroidetes* species. The insert in GyrB is present in all *Bacteroidetes* species (>100) covering different orders and families, indicating that it is a distinctive characteristic of the group. Three signatures consisting of an 18 aa insert in ATPase α-subunit, an 8–9 aa insert in the FtsK protein and a 1 aa insert in the UvrB protein are commonly shared only by the *Bacteroidetes* and *Chlorobi* homologs providing evidence that these two groups are specifically related to each other. Two additional inserts in the RNA polymerase β'-subunit (5–7 aa) and Serine hydroxymethyl-transferase (14–16 aa), which are commonly present in various *Bacteroidetes*, *Chlorobi*, and *Fibrobacteres* homologs, but not any other bacteria, provide evidence that these groups shared a common ancestor exclusive of all other bacteria. The FCB groups of bacteria are indicated to have diverged from this common ancestor in the following order: *Fibrobacteres* → *Chlorobi* → *Bacteriodetes*. The inferences from signature sequences are strongly supported by phylogenetic analyses. These observations suggest that the FCB groups of bacteria should be placed in a single phylum rather than three distinct phyla. Signature sequences in a number of other proteins provide evidence that the FCB group of bacteria diverged at a similar time as the Chlamydiae group, and that the *Spirochetes* and *Aquificales* groups are its closest relatives.

Keywords Bacterial Phylogeny; Cytophaga; Flavobacteria; Green Nonsulfur Bacteria; Signature Sequences; Phylogenetic Trees; Protein Signatures; Branching Order

Received 5 January 2004; accepted 12 January 2004.

Address correspondence to Dr. R.S. Gupta, Department of Biochemistry, McMaster University, Hamilton, Canada L8N 3Z5. E-mail: gupta@mcmaster.ca

INTRODUCTION

Fibrobacteres, *Chlorobi*, and *Bacteroidetes* represent three of the presently recognized main divisions (or phyla) within *Bacteria* (Ludwig & Klenk 2001; Garrity & Holt 2001; Balows, Trüper, Dworkin et al. 1992). Species belonging to these phyla exhibit enormous phenotypic and metabolic diversity (Gherna & Woese 1992). *Fibrobacteres* are major rumen bacteria, which enable the ruminants to breakdown and digest plant cellulose (Hungate 1950; Amann, Lin, Key et al. 1992; Krause & Russell 1996; Lin & Stahl 1995). The *Chlorobi*, also known as Green sulfur bacteria, are obligate photolithotrophs which carry out anaerobic photosynthesis withdrawing electrons from hydrogen sulfide (Truper & Pfennig 1992; Imhoff 2003; Overmann & Tuschak 1997; Frostl & Overmann 2000) They are mainly found in anoxic aquatic environment where sunlight is able to penetrate. The bacteria belonging to *Bacteroidetes* division, previously known as the *Cytophaga-Flavobacteria-Bacteroides* (CFB) group, are widely distributed in different habitats ranging from Antarctic ice to fresh and salt water lakes, to terrestrial soil and hydrothermal Obsidian pool (Shah 1992a, b; Holmes 1992; Reichenbach 1992; Shah 1992; Paster, Dewhirst, Olsen et al. 1994; Hugenholtz, Pitulle, Hershberger et al. 1998; Ventosa, Nieto, & Oren 1998; O'Sullivan, Weightman, & Fry 2002; Ohkuma, Noda, Hongoh, et al. 2002; Humayoun, Bano, & Hollibaugh 2003; Barbieri, Potenza, Rossi et al. 2000; Anton, Rossello-Mora, Rodriguez-Valera et al. 2000). Many of the CFB group bacteria show close association with human and animal hosts, where they perform useful functions and are also the causative agents of a variety of diseases (Murray & Stackebrandt 1995; Duncan 2003; Shah, Gharbia, & Duerden 1998).

The taxonomy and evolutionary relationships both within and among the *Fibrobacteres*, *Chlorobi*, and *Bacteroidetes* (FCB) divisions of bacteria has been revised a number of times but are still not clearly understood. Based on16S rRNA trees, these groups of bacteria were originally indicated to form one of the 10 main phyla within *Bacteria* (Woese, Stackebrandt, Macke et al. 1985; Woese 1987). *Fibrobacter succinogenes*, in the beginning was considered a member of the *Bacteroides* group (Hungate 1950). However, subsequent studies indicated that *F. succinogenes* was less closely related to other *Bacteroides* species, leading to its

transfer into a distinct phlyum (Montgomery, Flesher, & Stahl 1988). *Chlorobium* species generally exhibit close affinity to the CFB group of bacteria in phylogenetic trees (Olsen, Woese, & Overbeek 1994; Olsen & Woese 1993; Gruber, Eisen, Gish et al. 1998; Gupta, Mukhtar, & Singh 1999; Zinder 1998). However, their photosynthetic ability has been an important consideration in their recognition as a distinct phylum. In comparison to *Fibrobacteres* and *Chlorobi*, which contain only a limited number of species (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003), *Bacteroidetes* form a major group within bacteria, exhibiting a potpourri of phenotypes, including gliding behavior and the ability to digest and grow on a variety of complex substrates such as cellulose, chitin, and agar (Shah 1992a; Reichenbach 1992; Paster, Dewhirst, & Olsen 1994; Shah,
Garbia, & Duerden 1998; Holdeman, Kelley, & Moore 1984; Reichenbach & Weeks 1981). Based on their biochemical and physiological properties, DNA homology, and rRNA characteristics, Shah and Collins (Shah & Collins 1988, 1989, 1990) initially proposed division of *Bacteroides* into three main groups, namely *Bacteroides*, *Prevotella*, and *Porphyromonas*. In later, more detailed studies based on 16S rRNA trees, five major subgroups (viz. *Cytophaga*, *Flavobacter*, *Bacteroides*, *Sphingobacter*, and *Saprospira*) were identified within the CFB phylum (Gherna & Woese 1992; Paster, Dewhirst, Olsen et al. 1994). More recently, in the 2nd edition of Bergey's *Manual of Systematic Bacteriology* and the Ribosomal Database Project-II, the *Bacteroidetes* phylum is divided into three main classes viz. *Bacteroides*, *Flavobacteria*, and *Sphingobacteria*, each comprised of a number of different families (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003).

Our current understanding of the taxonomy and evolutionary relationships of FCB divisions is mainly based on 16S rRNA trees (Woese 1987; Olsen, Woese, & Overbeek 1994; Olsen & Woese 1993; Cole, Chai, Marsh et al. 2003). Except for branch patterns in the 16S rRNA trees, there are no distinctive biochemical or molecular characteristics known which can clearly distinguish species belonging to these phyla from other bacteria. The relationships of these phyla to each other and to other main divisions within bacteria are also not resolved at present (Ludwig & Klenk 2001; Ludwig & Schleifer 1999). Recently, a new approach has been described based on conserved inserts or deletions (indels or signature sequences) found in different proteins that is proving very useful in clearly distinguishing the major groups within *Bacteria* and also for deducing the interrelationships among them (Gupta 1998, 2000, 2002, 2003; Gupta & Griffiths 2002). This review examines the evolutionary relationships among the FCB divisions of bacteria using both the signature sequence and traditional phylogenetic approaches. This article describes for the first time a large number of conserved indels in widely distributed proteins that are distinctive characteristics of the FCB divisions of bacteria. A number of these signatures are unique for the *Bacteroidetes* species, whereas several others are commonly shared by either the *Bacteroidetes* and *Chlorobi* groups, or by *Bacteroidetes*, *Chlorobi* and *Fibrobacteres*. These signatures provide molecular means for distinguishing the FCB divisions of bacteria from all other bacteria and they also clarify the evolutionary relationships among them. The distribution patterns of these signatures and phylogenetic analyses based on different proteins, provide strong evidence that the FCB divisions of bacteria shared a common ancestor exclusive of all other bacteria, and that these groups branched off from that ancestor in the order *Fibobacteres* → *Chlorobi* → *Bacteroidetes*. The signature sequences described here also clarify the branching order of the FCB group relative to the other main divisions within *Bacteria*.

THE USEFULNESS OF CONSERVED INDELS FOR EVOLUTIONARY STUDIES

Conserved indels (or signature sequences) that are commonly shared by members of one or more taxa provide a powerful tool for identifying individual taxa in molecular terms and also to understand how these phyla are related to others (Gupta 1998, 2000; Gupta & Griffiths 2002; Meyer, Cusanovich, & Kamen 1986; Gupta & Singh 1992, 1994; Baldauf & Palmer 1993; Rivera & Lake 1992; Karlin & Brocchieri 1998; Griffiths & Gupta 2002; Morse, O'Hanlon, & Collins 2002; Brown, Masuchi, Robb et al. 1994). The evolutionary significant indels are generally of defined size and are present in the same position in a given protein. Such indels are also flanked by conserved regions on both sides to ensure that they are reliable and not resulting from sequence misalignment or other errors. The simplest and most parsimonius explanation for such conserved indels is that they were introduced once in an ancestral species and then passed on to various descendants. The signature sequences that have been identified are of different kinds. Some signatures are "group-specific" i.e., they are commonly present in all species belonging to particular taxa, but not found in the other groups. These signatures were likely introduced at the time when the particular groups of bacteria diverged (Gupta 2000; Griffiths & Gupta 2002). Group-specific signatures that are distinctive of the Chlamydiae, Cyanobacteria, Proteobacteria, *Actinobacteria*, and *Deinococcus-Thermus* groups have been recently described (Gupta 2000; Griffiths & Gupta 2002; Gupta, Pereira, Chandrasekera et al. 2003; and Gupta 2004b) and based on them it is possible to define these groups in clear molecular terms and to identify other species belonging to them. Another type of signatures are so-called "main line" signatures, which are commonly shared by several major groups of bacteria but are absent from other bacterial phyla (Gupta 1998, 2003; Gupta & Griffiths 2002, 2003). Such signatures were likely introduced at critical branch points during the course of bacterial evolution and they are of much value in deducing the branching order and inter-relationships among different groups of bacteria (Gupta 1998, 2001, 2002, 2003; Gupta & Griffiths 2002; Griffiths and Gupta 2004a). The following sections describe a variety of signatures that are

helpful in understanding the taxonomy and evolutionary relationships of the FCB groups of bacteria. These signatures were identified by similar means as described in earlier work (Gupta 1998, 2000; Gupta, Pereira, Chandrasekera et al. 2003). The evolutionary significance of these signatures and other relevant information is also presented.

SIGNATURE SEQUENCES CHARACTERISTICS OF THE *BACTEROIDETES*

The *Bacteroidetes* represents one of the largest groups within *Bacteria*. In early of December 2003, there were 5896 entries for *Bacteroidetes* species in the Ribosomal database project (RDP-II), (Cole, Chai, Marsh et al. 2003) making it the fourth

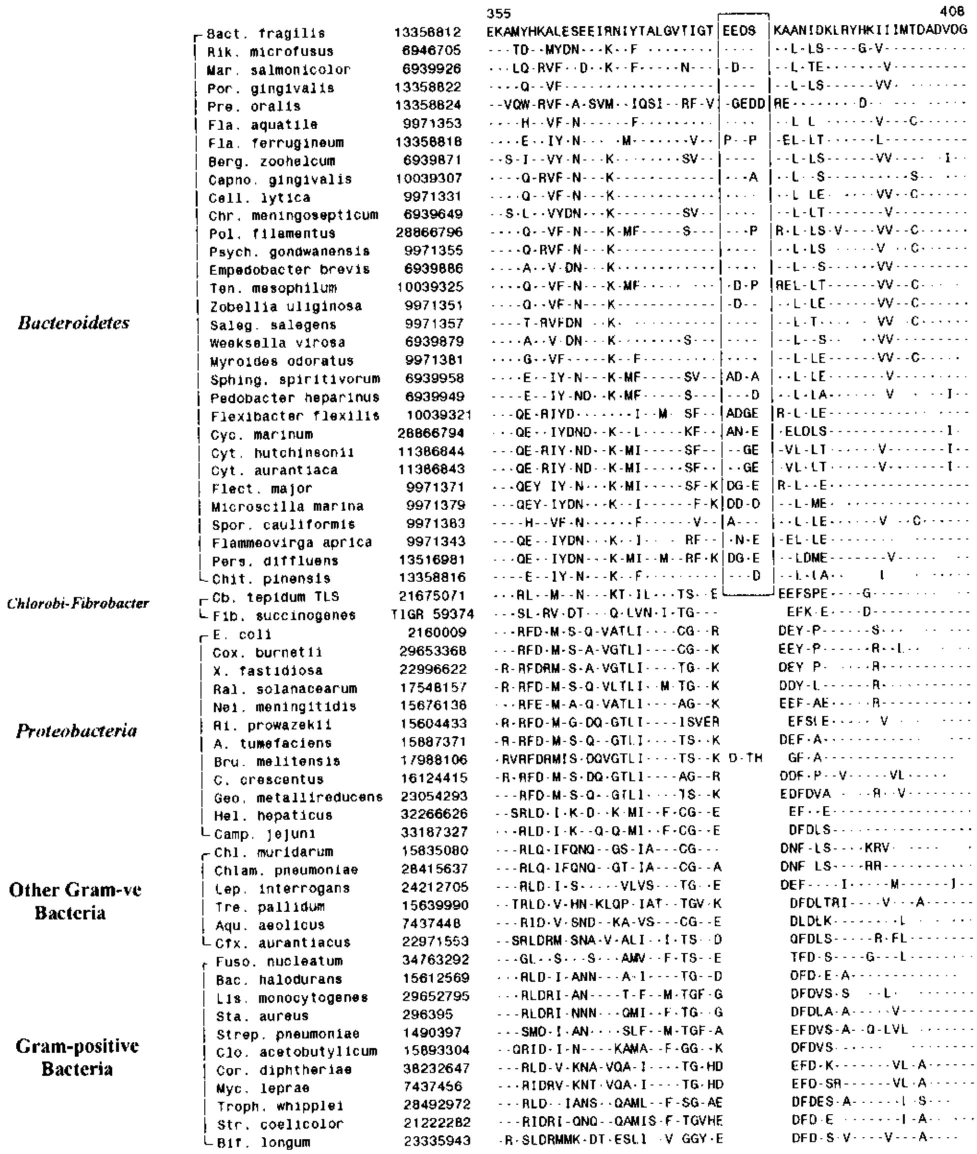

FIG. 1. Partial sequence alignment for Gyrase B showing a 4 aa insert (boxed), which is specific for the *Bacteroidetes* species. Dashes in all sequence alignments indicate identity with the amino acid on the top line. The accession numbers of various sequences are shown in the second column. Only representative sequences from different *Bacteria* are shown. Other *Bacteroidetes* species where this insert is present are listed in Table 1. Abbreviations in the species names are as follows: *A.*, *Agrobacterium*; *Aqu.*, *Aquifex*; *Bac.*, *Bacillus*; *Bact.*, *Bacteriodes*; *Berg.*, *Bergeyella*; *Bif.*, *Bifidobacterium*; *Bru.*, *Brucella*; *C.*, *Caulobacter*; *Camp.*, *Campylobacter*; *Capno.*, *Capnocytophaga*; *Cb.*, *Chlorobium*; *Cell.*, *Cellulophaga*; *Cfx.*, *Chloroflexus*; *Chit.*, *Chitinophaga*; *Chl.*, *Chlamydia*; *Chlam*, *Chlamydophila*; *Chr.* *Chryseo*; *Clo.*, *Clostridium*; *Cor.*, *Corynebacterium*; *Cox.*, *Coxiella*; *Cyc.*, *Cyclobacterium*; *Cyt.*, *Cytophaga*; *E.*, *Escherichia*; *Fib.*, *Fibrobacter*; *Fla.*, *Flavobacterium*; *Flect.*, *Flectobacillus*; *Fuso.*, *Fusobacterium*; *Geo*, *Geobacter*; *Hel.*, *Helicobacter*; *Lep.*, *Leptospira*; *Lis.*, *Listeria*; *Mar.*, *Marinilabilia*; *Myc.*, *Mycobacterium*; *Nei.*, *Neisseria*; *Pers.*, *Persicobacter*; *Pol.*, *Polaribacter*; *Por.*, *Porphyromonas*; *Pre.*, *Prevotella*; *Psych.*, *Psychro.*; *Ral.*, *Ralstonia*; *Ri.*, *Rickettsia*; *Rik.*, *Rikenella*; *Saleg.*, *Salegentibacter*; *Sphing.*, *Sphingobacterium*; *Spor.*, *Sporocytophaga*; *Sta.*, *Staphylococcus*; *Str.*, *Streptomyces*; *Strep.*, *Streptococcus*; *Tan.*, *Tannerella*; *Ten.*, *Tenacibaculum*; *Tre.*, *Treponema*; *Troph.*, *Tropheryma*; *X.*, *Xylella*.

TABLE 1
Bacteroidetes species containing Gyrase B Insert

Bacteriodales *(Bacteriodaceae)*	
Bact. thetaiotaomicron	29348838
Bacteroides fragilis	13358812
Bacteroides vulgatus	6939896
Bacteriodales *(Rikenellaceae)*	
Rikenella microfusus	6946705
Marinilabilia salmonicolor	6939926
Bacteriodales *(Porphyromondaceae)*	
Porph. asaccharolytica	6939900
Porphyromonas gingivalis	13358822
Porphyromonas levii	13358820
Tannerella forsythensis	*
Bacteriodales *(Prevotellaceae)*	
Prevotella intermedia	*
Prevotella oralis	13358824
Flavobacteriales*(Flavobacteriaceae)*	
Flavobacteria aquatile	9971353
Flavobacteria ferrugineum	19912133
Flavobacteria johnsoniae	9971347
Bergeyella zoohelcum	6939871
Capnocytophaga canimorsus	9971325
Capnocytophaga gingivalis	10039307
Capnocytophaga ochracea	10039309
Cellulophaga baltica	28866792
Cellulophaga fucicola	28866790
Cellulophaga lytica	9971335
Chryseo. meningosepticum	6939649
Chryseobac. indologenes	6946703
Empedobacter brevis	6939886
Polaribacter filamentus	28866796
Polaribacter glomeratus	9971373
Psychro. gondwanensis	9971355
Salegentibacter salegens	9971357
Tenacibac.amylolyticum	10039343
Tenacibaculum maritimum	9971359
Tenacibaculum mesophilum	10039325
Tenacibaculum ovolyticum	9971363
Weeksella virosa	6939879
Zobellia uliginosa	9971351
Flavobacteriales *(Myroidaceae)*	
Myroides odoratus	9971381
Sphingobacteriales *(Sphingobacteriaceae)*	
Sphingobact. spiritivorum	6939958
Pedobacter heparinus	6939949
Sphingobacteriales *(Flexibacteriaceae)*	
Flexibacter aggregans	9971377
Flexibacter canadensis	10039323
Flexibacter elegans	30580459
Flexibacter filiformis	10039317
Flexibacter flexilis	10039321
Flexibacter japonensis	30580460
Flexibacter litoralis	9971369
Flexibacter sancti	10039319
Cyclobacterium marinum	28866794
Cytophaga aurantiaca	11386843
Cytophaga fermentans	6939919
Cytophaga hutchinsonii	23135434
Cytophaga latercula	9971349
Cytophaga marinoflava	9971345
Cytophaga sp. I-545	19912069
Cytophaga sp. I-601	19912059
Cytophaga sp. I-602	19912061
Cytophaga sp. I-762	19912053
Cytophaga sp. I-1147	19912063
Cytophaga sp. I-1856	19912101
Cytophaga sp. I-1858	19912103
Cytophaga sp. I-2029	19912055
Cytophaga sp. I-2031	19912043
Cytophaga sp. MBIC01355	13516993
Cytophaga sp. MBIC01481	19912113
Cytophaga sp. MBIC01482	19912109
Cytophaga sp. MBIC01483	13517059
Cytophaga sp. MBIC01484	13517061
Cytophaga sp. MBIC01485	19912111
Cytophaga sp. MBIC04661	19912077
Cytophaga sp. MBIC04663	19912041
Cytophaga sp. MBIC04664	19912079
Cytophaga sp. MBIC04665	19912139
Cytophaga sp. MBIC04666	19912089
Cytophaga sp. MBIC04669	19912075
Cytophaga sp. MBIC04671	19912033
Cytophaga sp. MBIC04682	19912071
Cytophaga sp. MBIC04693	19912051
Cytophaga sp. T-424	19912091
Cytophaga sp. T-565	19912093
Cytophaga sp. T-588	19912057
Flectobacillus major	9971371
Microscilla furvescens	9971375
Microscilla marina	9971379
Microscilla sericea	10039351
Sporocytophaga cauliformis	9971383
Sphingobacteriales *(Flammeovirgaceae)*	
Flammeovirga aprica	9971343
Persicobacter diffluens	13516981
Sphingobacteriales *(Crenotrichaceae)*	
Chitinophaga pinensis	13358814
Unclassified CFB-Bacteria	
Marine CFB-MBIC1357	9971327
Marine CFB-MBIC01539	13516987
Marine CFB-MBIC01599	13516989
Marine CFB-MBIC04441	13516997
Marine CFB-MBIC04442	13516999

(Continued on next page)

TABLE 1
Bacteroidetes species containing Gyrase B Insert (Continued)

Marine CFB-MBIC04443	13517001
Marine CFB-MBIC04445	13517005
Marine CFB-MBIC04446	13517007
Marine CFB-MBIC04450	13517011
Marine CFB-MBIC04451	13517013
Marine CFB-MBIC04452	13517015
Marine CFB-MBIC04453	13517017
Marine CFB-MBIC04455	13517021
Marine CFB-MBIC04458	13517025
Marine CFB-MBIC04466	13517029
Marine CFB-MBIC04467	13517031
Marine CFB-MBIC04468	13517033
Marine CFB-MBIC04470	13517037
Marine CFB-MBIC04471	13517039
Marine CFB-MBIC04473	13517043
Marine CFB-MBIC04474	13517045
Marine CFB-MBIC04475	13517047
Marine CFB-MBIC04476	13517049
Marine CFB-MBIC04477	13517051
Marine CFB-MBIC04487	13517065
Marine CFB-MBIC05204	13516991

Information for other *Bacteroidetes* species which contain the GyrB insert, including accession numbers of the sequences, and the orders and families to which they belong is provided. The sequences marked with astericks (*) are from the NCBI unpublished microbial genome database.

largest phylum within *Bacteria* after the Proteobacteria, *Firmicutes* and *Actinobacteria* (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003). Based on 16S rRNA data, the *Bacteroidetes* group has been divided into three main classes or orders (viz. *Bacteroidales*, *Flavobacteriales*, and *Sphingobacteriales*) and a fourth class or order of unclassified species (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003). The order *Bacteroidales* consists of four main families *Bacteroidaceae*, *Rikenellaceae*, *Porphyromonadaceae*, and *Prevotellaceae*, and a fifth family of unclassified bacteria. The first four families are comprised of 7, 2, 4, and 1 genera, respectively. The order *Flavobacteriales* is made up of three families viz. *Flavobacteriaceae*, *Myroidaceae*, and *Blattabacteriaceae*. Of these, *Flavobacteriaceae* contain 20 different genera, whereas the remaining two families each have a single genus (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003). The last of the *Bacteroidetes* orders (i.e. *Sphingobacteriales*) which contains most of the gliding bacteria is again quite large and it has been divided into five families viz. *Sphingobacteriaceae*, *Saprospiraceae*, *Flexibacteriaceae*, *Flammovirgaceae*, and *Crenotrichaceae*. These families contain 2, 3, 12, 4, and 6 genera, respectively, (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003). As indicated above, the division of *Bacteroidetes* into these orders and families is mainly based on 16S rRNA sequences and presently no distinctive molecular or phenotypic characteristics are known that can distinguish either this entire phylum or various orders/families within it. Gherna and Woese (1992) have previously proposed the division of CFB phylum into five subgroups (viz. Cytophaga, Flavobacter, Bacteroides, Sphingobacter, and Saprospira) and they also described specific substitutions in 16S rRNA that were distinctive of these subgroups. However, since this work was published, the phylogeny of the CFB group has been revised extensively as noted above and the sequence information for this group has also increased by more than 100-fold (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003). Hence, the usefulness of the described signatures in rRNA (Gherna and Woese 1992) as markers for any of the orders or families within the *Bacteroidetes* is presently unclear.

Recently, the complete genomes of two species belonging to the CFB group viz. *Bacteroides thetaiotamicron* and *Porphyromonas gingivalis* Strain W83 have been published (Xu, Bjursell, Himrod et al. 2003; Nelson, Fleischmann, DeBoy et al. 2003), and genomes of several other species are in the process of being completed. Partial sequences for a number of these species viz. *Bacteroides fragilis*, *Cytophaga hutchinsonii*, *Prevotella intermedia*, and *Tannerella forsythensis* are available in the NCBI microbial genomes database (NCBI 2001). This information, in conjunction with that available for *Chlorobium tepidum* (Eisen, Nelson, Paulsen et al. 2002) and the partial genome sequence for *Fibrobacter succinogenes* have proven instrumental in the identification of a number of signatures that appear distinctive for the *Bacteroidetes*, or those showing relationships among these phyla. The signatures described in this review were identified by visual examination of multiple sequence alignment of different proteins, which included all available sequences from *Fibrobacteres*, *Chlorobi*, and *Bacteroidetes*. The indels that were unique to only particular species were not further investigated. However, the potential usefulness of various other indels in conserved regions which were shared by more than one species from these groups of bacteria was examined by additional BLAST searches on short sequence segments (usually between 60–100 aa) containing the indels and their flanking regions. The sequence information from these searches were compiled into signature files. These signature sequences and their evolutionary significances are described below.

DNA gyrase, a type II topoisomerase, which is responsible for introducing negative supercoils into DNA, plays an essential role in DNA replication (Wang 1996; Wigley 1995). The enzyme is made up of two subunits, A and B, and is present in different bacteria. Because of their ubiquity and high degree of sequence conservation, protein sequences for DNA gyrase subunits have been extensively used for phylogenetic studies. (Kasai, Watanabe, Gasteiger et al. 1998; Suzuki, Nakagawa, Harayama et al. 2001; Sawada, Suzuki, Matsuda et al. 1999; Seo & Yokota 2003; Yamamoto, Kasai, Arnold et al. 2000). The resolution obtained from such studies is generally comparable, and in many cases has been noted to be superior, to that observed in the 16S rRNA trees (Kasai, Watanabe, Gasteiger et al. 1998; Suzuki, Nakagawa, Harayama et al. 1999). We have identified

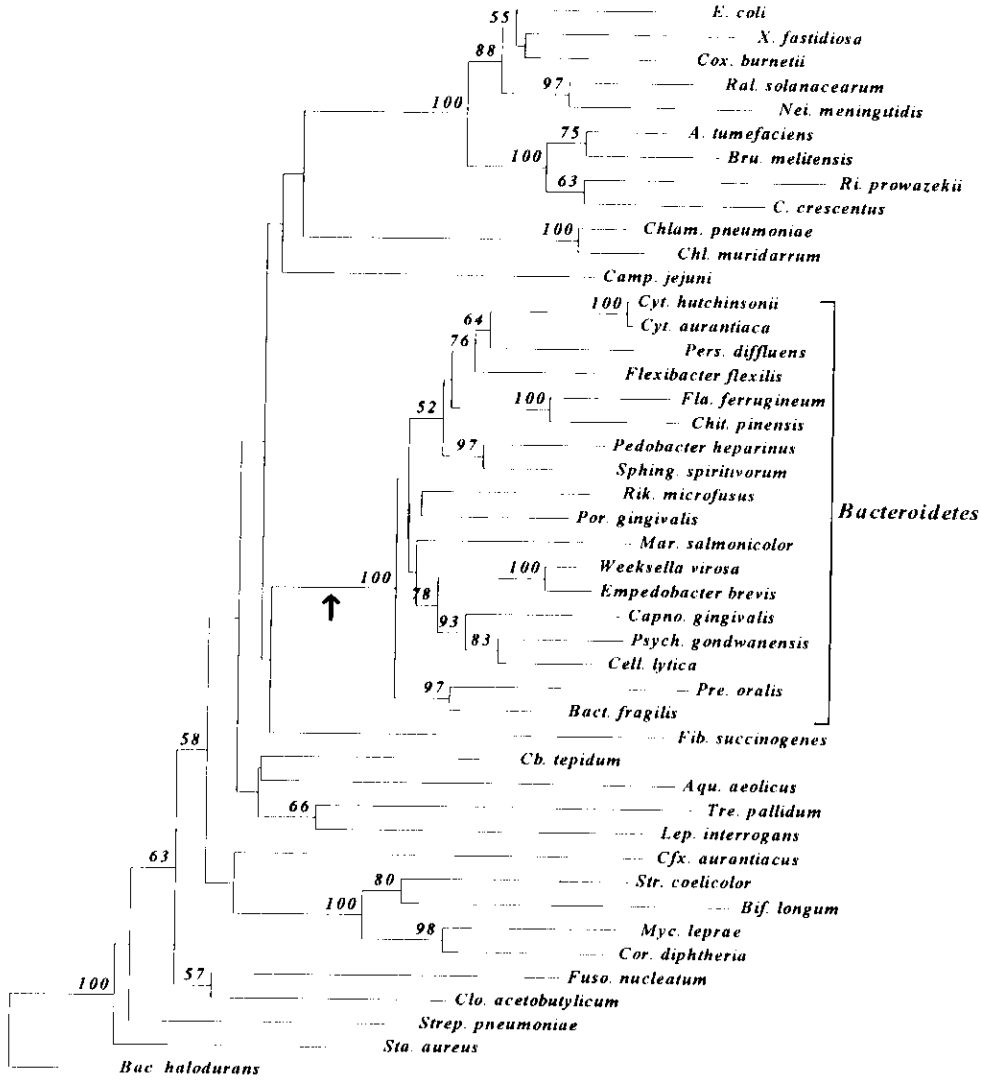

FIG. 2. A neighbor-joining distance tree based on GyrB sequences. The tree is based on 354 aa, excluding all indels, which could be aligned without ambiguity in the different sequences. The tree was constructed using the TREECON phylogenetic program (Van de & Wachter 1994). The root of the tree has been arbitrarily placed between *Firmicute* species, which based on our work on signature sequences is indicated to be one of the earliest branching phyla within *Bacteria* (Gupta 1998, 2000, 2003; Gupta & Griffiths 2002). The bootstrap scores for various nodes which were >50 are indicated. The abbreviations in the species names are as in Figure 1. The arrow marks the evolutionary stage where the identified signature in GyrB was likely introduced.

a 4 aa insert in a conserved region of DNA gyrase B (Gyb B) that is commonly shared by various *Bacteroidetes* species, but is not found in other groups of bacteria (Figure 1). The sequence information for GyrB is available from a large number of bacteria, and only representative results from the main groups are shown in Figure 1. Suzuki et al. (1999, 2001), have carried out extensive work on GyrB sequences from the CFB division. As a result of their work, sequence information for this region of GyrB is available from a large number of *Bacteroidetes* species listed in Table 1. The species for which information is available include all of the different orders as well as the main families within the *Bacteroidetes* phylum. The insert in GyrB (Figure 1) is present in all of these species but is not found in other groups of bacteria, indicating that it is a distinctive characteristic of the entire phylum. Besides *Bacteroidetes*, a 4 aa insert is also present in this region in the *Brucella* species (α-Proteobacteria). Since this insert is not present in any other proteobacteria, it has either occurred independently or acquired by means of lateral gene transfer (LGT).

GyrB sequences have been used by other investigators for phylogenetic analyses (Kasai, Watanabe, Gasteiger et al. 1998; Suzuki, Nakagawa, Harayama et al. 1999, 2001; Seo & Yokota 2003; Yamamoto, Kasai, Arnold et al. 200; Yanez, Catalan, Apraiz et al. 2003; Fukushima, Kakinuma, Kawaguchi 2002). Suzuki et al. have reported detailed phylogenetic analysis based on GyrB sequences for the CFB group of bacteria, which led to the identification of a new genus *Tenacibaculum* (Suzuki, Nakagawa, Harayama et al. 1999, 2001). However, most of the

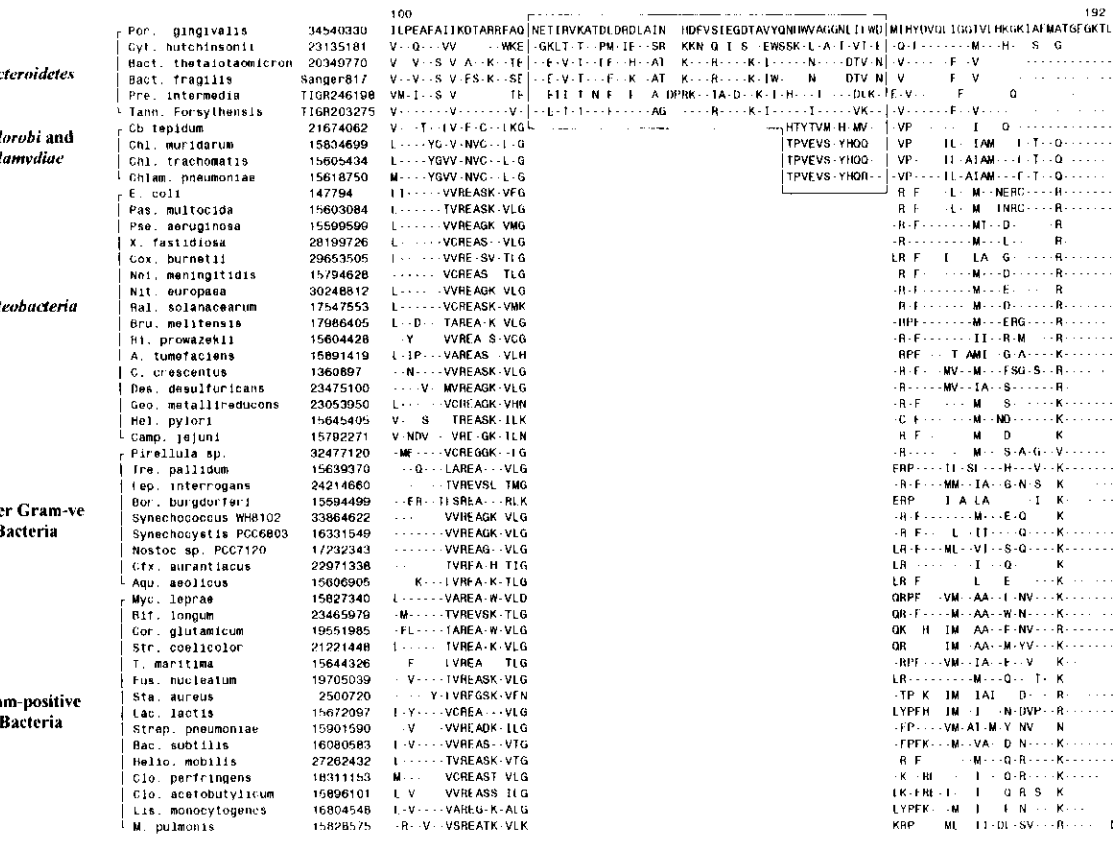

FIG. 3. Excerpt from SecA sequence alignment showing a large insert (~45 aa) that is distinctive of the *Bacteroidetes* species. A smaller 13 aa insert is also present in this region in various Chlamydiae and *Chorobium tepidum*, which is likely of independent origin. Additional species abbreviations are: *Bor., Borrelia*; *Des., Desulfovibrio*; *Helio., Heliobacillus*; *L., Lactococcus*; *M., Mycoplasma*; *Nit., Nitrosomonas*; *Pas., Pasteurella*; *Pse., Pseudomonas*; *T., Thermotoga*.

earlier studies on GyrB have been limited to examining the relationship between members of particular phylum and they do not provide information regarding relationships among different main groups of bacteria. Figure 2 shows a phylogenetic tree based on GyrB sequences that includes representative species from the different main phyla. Such analysis was carried out on sequence regions for which information was available from different species. Any sequence region where the alignment was deemed unreliable due to poor sequence conservation was omitted from the analysis. The signature region was also not included in the phylogenetic analysis to prevent any bias due to this indel on the resulting tree topology. The tree shown is a neighbor-joining distance tree (Saitou & Nei 1987) obtained after bootstrapping 100 times (Felsenstein 1985), and was constructed using the TREECON phylogeny program (Van de & De Wachter 1994). The bootstrap scores for different nodes which were >50 are shown. As seen from the tree based on GyrB in Figure 2, all of the *Bacteroidetes* species formed a well-defined clade grouping together 100% of the time. The observed insert in the GyrB sequences was thus likely introduced in a common ancestor of this group at the stage marked by the arrow in Figure 1. Furthermore, species belonging to various other bacterial groups viz. *Proteobacteria*, *Chlamydiae*, *Actinobacteria*, *Firmicutes*, and *Spirochetes*, were found to form well-defined clades grouping together with high affinity in this tree, as represented by high bootstrap scores. However, the relative branching order of these phyla was not resolved in the GyrB tree, which is similar to that seen in the rRNA and other trees (Olsen, Woese, & Overbeek 1994; Olsen & Woese 1993; Gupta, Eisen, Gish et al. 1998; Gupta, Pereira, Chandrasekera et al. 2003; Gupta 1995; Eisen 1995; Gupta, Bustard, Falah et al. 1997; Brown & Doolittle 1997). The inability of the phylogenetic trees to resolve the branching order of higher phyla is related to their dependence upon a large number of variables, as discussed in earlier work (Gupta 1998). It should be noted that in the GyrB tree, *Brucella* species branched within the α-Proteobacteria clade, which is as expected (Cole, Chai, Marsh et al. 2003; Gupta 2000; Boone, Castenholz & Garrity 2001). This provides evidence against the LGT of *gyrB* gene from *Bacteroidetes* to *Brucella*, because if the insert in *Brucella* was due to LGT, than one expect this sequence to be most homologous to *Bacteroidetes* and it should branch with this group in the phylogenetic tree. However, the branching of *Brucella* with the α-Proteobacteria, as is normally expected indicates that the 4 aa insert in this genus has originated independently.

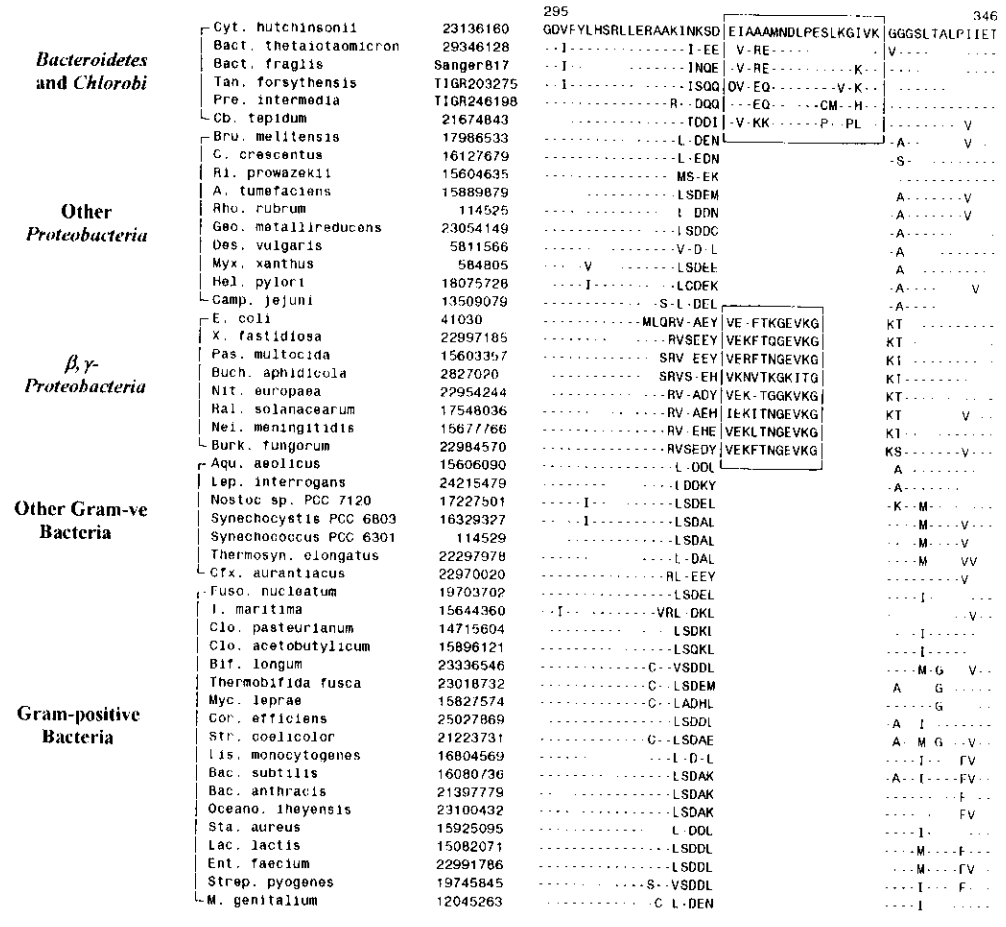

FIG. 4. Partial sequence alignment for ATPase α-subunit showing an 18 aa insert that is commonly shared by various *Bacteroidetes* species and *Cb. tepidum*. An unrelated insert of 11 aa is also present in this region in various β- and γ-Proteobacteria. Additional abbreviations in species names are: *Buch.*, *Buchnera*; *Burk.*, *Burkholderia*; *Ent.*, *Enterococcus*; *Meso.*, *Mesorhizobium*; *Myx.*, *Myxococcus*; *Oce.*, *Oceanobacillus*; *Rho.*, *Rhodospirillum*; *Tan.*, *Tannerella*; *Thermosyn.*, *Thermosynechococcus*; *Y.*, *Yersinia*.

Another prominent signature that appears to be distinctive of the *Bacteroidetes* phylum is present in the SecA protein, which is involved in the export of proteins to the periplasmic compartment (Valentin 1997; Schmidt & Kiser 1999). Similar to GyrB, SecA homologs are also present in different bacterial genomes. The identified signature in this case consists of a 45 aa insert in a conserved region that is unique to various *Bacteroidetes* species for which sequence information is available (Figure 3), but not found in any other bacteria. Although sequence information for SecA is limited for *Bacteroidetes*, it includes representative species from a number of different families within *Bacteroidales* (viz. *Bacteroides, Porphyromonas, Tannerella, Prevotella*) as well as *Sphingobacteriales*. Thus, it is likely that this signature will also be present in other *Bacteroidetes* species. Although the 45 aa insert is unique to the *Bacteroidetes* species, a smaller insert of 13 aa is also present in the same region in various Chlamydiae as well as in *Cb. tepidum*. This latter indel differs from that found in *Bacteroidetes*, both in terms of its length and the sequence, and therefore, one expects this insert to have originated independently.

SIGNATURES CHARACTERISTICS OF THE *BACTEROIDETES* AND *CHLOROBI*

The *Chlorobia* phylum is comprised of a single class (*Chlorobia*), Order (*Chlorobiales*), and family (*Chlorobiaceae*), which is further divided into five different genera (viz. *Chlorobium, Ancalochloris, Chloroherpeton, Pelodictyon*, and *Prosthecochloris*) and another group consisting of unclassified species (Garrity & Holt 2001; Frostl & Overmann 2000; Cole, Chai, Marsh et al. 2003; Alexander, Andersen, Cox et al. 2002). One distinctive cytological feature of this group is the presence of unique light harvesting complex chlorosomes, which harbor visible light absorbing pigments such as the bacteriochlorophylls and carotenoids (Truper & Pfenning 1992; Alexander, Andersen, Cox et al. 2002; Blankenship 1992, 1994). Besides *Chlorobi*, chlorosomes have only been found in the Green nonsulfur group of bacteria (viz. *Chloroflexi*) (Pierson & Castenholz 1992), and it is possible that some photosynthetic genes have been transferred between these two groups of bacteria via LGT (Gupta 2003; Raymond, Zhaxybayeva, Gogarten et al. 2002; Igarashi, Harada, Nagashima et al. 2001). Recently, the complete genome

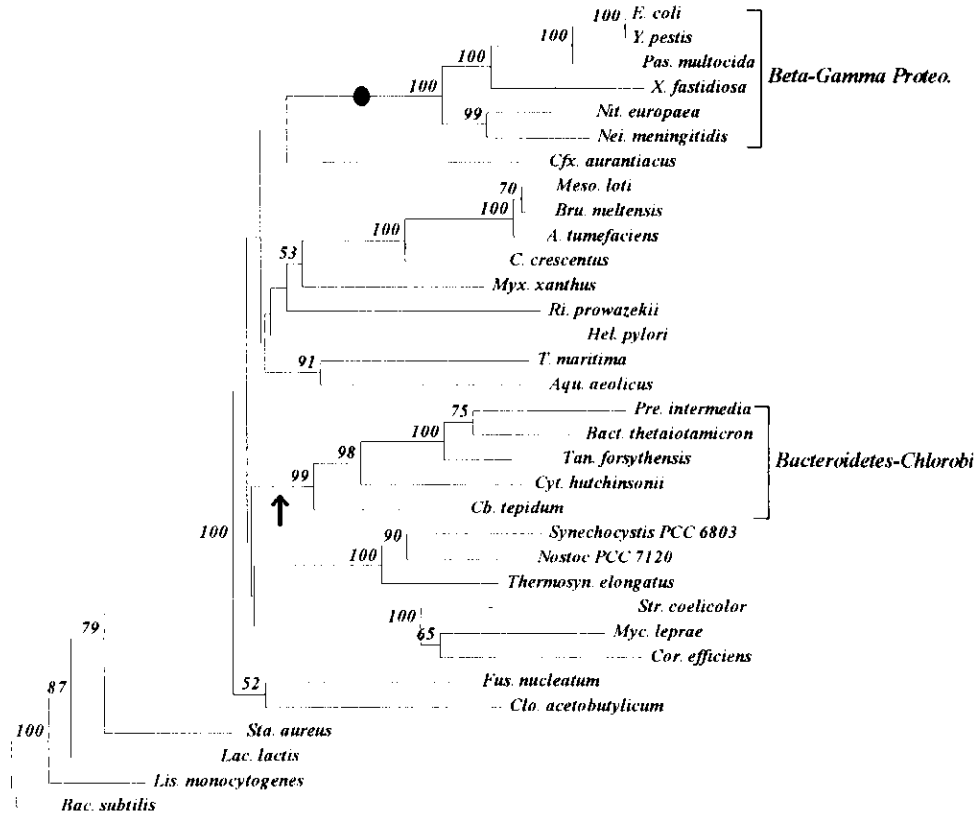

FIG. 5. A neighbor-joining distance tree based on ATPase (α) sequences. The tree is based on 392 aligned positions, which excluded all indels including the signature region in Figure 4. The bootstrap scores >50 are indicated. The arrow and the filled circle mark the evolutionary stages where the inserts in the *Bacteroidetes-Chlorobi* and in the $\beta\gamma$-Proteobacteria were introduced.

sequence of *Cb. tepidum* has been reported (Eisen, Nelson, Paulsen et al. 2001). In phylogenetic trees based on 16S rRNA and different proteins (e.g., RecA, Hsp70, etc.), *Chlorobi* species form a phylogenetically coherent group exhibiting close affinity to the *Bacteroidetes* group (Olsen, Woese, & Overbeek 1994; Olsen & Woese 1993; Gruber, Eisen, Gish et al. 1998; Gupta, Mukhtar, & Singh 1999). However, a specific relationship between these groups based on any molecular or other shared characteristic has not yet been demonstrated. We now describe a number of signature sequences in proteins that are uniquely shared by these bacterial phyla indicating that they are specifically related.

The F_1F_0ATP synthase is a multisubunit complex which synthesizes ATP from ADP and inorganic phosphate (Pi) utilizing the chemiosmotic energy of a transmembrane potential of cellular H^+ or Na^+ gradients (Levya, Bianchet, & Amzel 2003). This complex is located in the cytoplasmic membrane of bacteria, the inner membrane of mitochondria and the thylakoid membranes of chloroplasts. The F_1 portion of this complex is a heteromer made up of five subunits, α, β, γ, δ, and ϵ with the stoichiometry $\alpha_3\beta_3\gamma\delta\epsilon$. The primary structures of most of these subunits are highly conserved across different species. In the ATP synthase (α), an 18 aa insert is present in a highly conserved region that is uniquely shared by all available sequences for *Bacteroidetes* homologs and also by *Cb. tepidum*, but is not present in any other bacteria (Figure 4). A smaller and unrelated insert of 11 aa is also present in this region in various β- and γ-Proteobacteria, which is distinctive of these groups of bacteria (Gupta 2003). The different lengths and sequences of the two signatures indicate that these indels are unrelated and they have been introduced independently. In a phylogenetic tree based on ATP synthase (α) sequences (carried out after excluding all inserts and deletions), the *Bacteroidetes* species formed a well-defined clade which was strongly supported by the bootstrap value (98 out of 100) (Figure 5). Interestingly, the *Cb. tepidum* formed the outgroup of the *Bacteroidetes* group and a clade consisting of *Bacteroidetes* and *Chlorobium* species was again strongly supported by bootstrap analysis (99 of 100). Thus, a specific affinity of the *Bacteroidetes* and *Chlorobium* species as indicated by the shared signature sequence is strongly affirmed by phylogenetic analysis. The identified insert in this protein was thus likely introduced in a common ancestor of these two groups, at the stage marked by the arrow. In the tree shown in Figure 5, the β- and γ-Proteobacteria formed a distinct clade grouping together 100% of the time. The observed insert in these bacteria was thus independently introduced in a common ancestor of these groups (as marked by the filled dot in Figure 5), and it provides a molecular marker for these groups of bacteria (Gupta 2003).

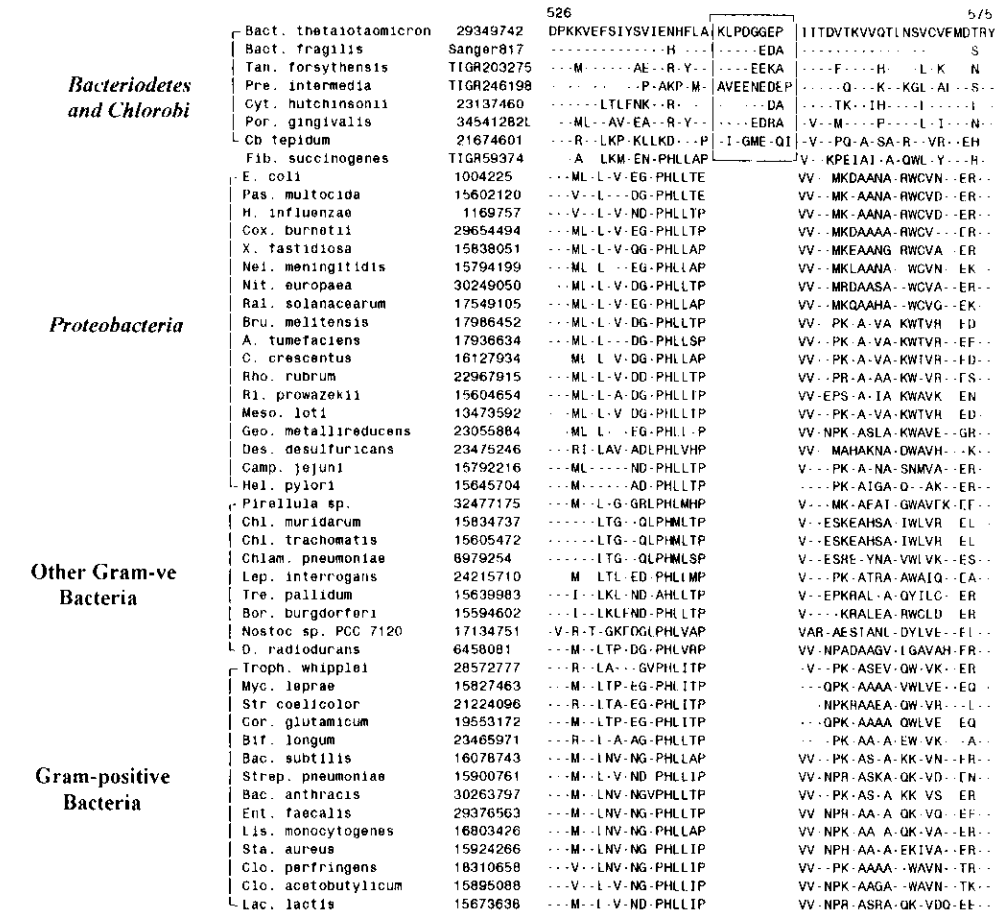

FIG. 6. Partial alignment of FtsK protein sequences showing an 8–9 aa insert that is commonly shared by *Bacteroidetes* and *Cb. tepidum*. Additional abbreviations: *D.*, *Deinococcus*; *H.*, *Haemophilus*.

FtsK is a multifunctional protein that plays a central role in cell division and chromosome segregation in bacteria (Espeli, Lee, & Marians 2003; Capiaux, Lesterlin, Perals et al. 2002). During the cell division process, this protein resides at the septal ring and is required for resolution of chromosome dimers. We have identified an 8–9 aa insert in a conserved region of the FtsK protein that is only found in various *Bacteroidetes* homologs and *Cb. tepidum* but is not present in any other bacteria (Figure 6). The presence of this commonly shared signature thus indicates a specific relationship between these bacteria exclusive of all others. A close and specific affinity of the species belonging to these groups is also clearly evident in a phylogenetic tree based on FtsK protein sequences (Figure 7). In this tree, *Cb. tepidum* formed the outgroup of the *Bacteroidetes* group and this close and specific relationship was strongly supported by the bootstrap score (94 out of 100). The identified insert in the FtsK protein was thus likely introduced in a common ancestor of these bacteria as marked by an arrow in Figure 7.

Another signature sequence showing a specific relationship between the *Bacteroidetes* and the *Chlorobi* is present in the UvrB protein. The UvrABC nucleotide excision repair complex is responsible for removal of a wide variety of genetic lesions from DNA (Verhoeven, Wyman, Moolenaar et al. 2002). The UvrB protein plays a central role in this process by recognizing DNA damage. We have identified a 1 aa insert in a conserved region of UvrB that is uniquely shared by *Bacteroidetes* species and *Cb. tepidum* but not found in any other bacteria (Figure 8). Similar to the signatures in ATP synthase (α) and FtsK proteins, the insert in UvrB was likely introduced in a common ancestor of *Bacteroidetes* and *Chlorobi* pointing to a specific relationship between these groups.

SIGNATURE SEQUENCES UNIQUELY SHARED BY THE *BACTEROIDETES, CHLOROBI*, AND *FIBROBACTERES*

The phylum *Fibrobacteres* (Class—*Fibrobacters*; Order—*Fibrobacterales*; Family—*Fibrobacteraceae*) consists of a single genus *Fibrobacter* with only two typed species viz. *F. succinogenes* and *F. intestinalis* (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003). Based on its physiological characteristics (e.g., ability to digest complex carbohydrates such as cellulose), *F. succinogenes* was originally known as *Bacteroides*

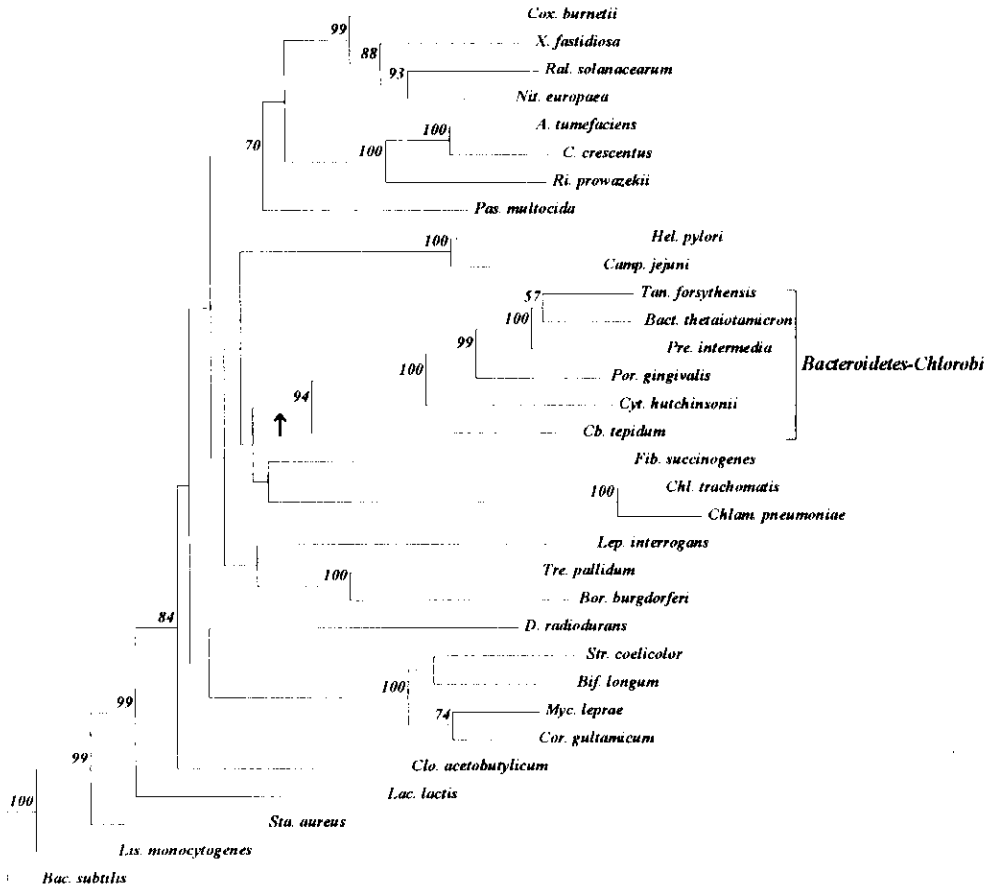

FIG. 7. A neighbor-joining phylogenetic tree based on FtsK protein sequences. The tree is based on 316 aligned positions, which did not include any indels. The bootstrap scores >50 are indicated. The arrow marks the evolutionary stages where the identified signature in this protein was likely introduced.

succinogenes (Hungate 1950). However, subsequent studies based on 16S rRNA has led to its placement in a separate genus and phylum (Garrity & Holt 2001; Amann, Lin, Key et al. 1992; Montgomery, Flesher, & Stahl 1988; Cole, Chai, Marsh et al. 2003). Nevertheless, in phylogenetic trees based on 16S rRNA, the *Chlorobi* and the CFB-division of bacteria are indicated as some of the closest relatives of *Fibrobacter* species (Olsen, Woese, & Overbeek 1994; Corsaro, Valassina, & Venditti 2003). In an earlier study based on signature sequences, *F. succinogenes* was found to branch in a similar position as the Chlamydiae and the CFB-divisions of bacteria, supporting its close relationships to the CFB-group (Griffiths & Gupta 2001). However, a specific relationship of *Fibrobacter* to the *Chlorobi* and the *Bacteroidetes* based on any uniquely shared molecular or other characteristic has not yet been demonstrated. The following sections describe a number of protein signatures that provide evidence to this effect.

The core subunits of the RNA polymerase (RNAP) (i.e., α, β and β') are evolutionarily conserved in sequence, structure, and function in all species ranging from bacteria to humans (Klenk & Zillig 1994; Olsen & Woese 1997). In the β subunit of RNAP (i.e., RpoC), we have identified a 5–7 aa insert in a highly conserved region that is commonly present in various *Bacteroidetes* species, *Cb. tepidum* as well as *F. succinogenes* (Figure 9). This insert is not found in the RpoC homologs from any other bacteria (or from other domains), although the corresponding protein is present in every species examined to date. A phylogenetic tree was also constructed based on RpoC sequences (Figure 10). The alignment used for these studies included only those parts which could be aligned without ambiguity and all sequence gaps, including the signature region, were excluded. In the tree based on RpoC sequences all of the main groups within *Bacteria* viz. Proteobacteria, *Deinococcus-Thermus*, Chlamydiae, Spirochetes, *Actinobacteria*, and *Firmicutes* were resolved with high bootstrap scores. The *Bacteroidetes* species also formed a monophyletic clade grouping together in all 100 bootstraps. However, of particular interest in the present context is the fact that *Cb. tepidum* comprised the outgroup of the *Bacteroidetes* clade and *F. succinogenes* formed the outgroup of the *Bacteroidetes-Chlorobi* clade. Both these relationships were again strongly supported by the bootstrap values (97 of 100 for both). Thus, a specific relationship among *Bacteroidetes*, *Chlorobi*, and *Fibrobacteres* as indicated by the shared presence of the unique insert was strongly affirmed by phylogenetic analysis. The identified insert in the RpoC protein thus was likely introduced in

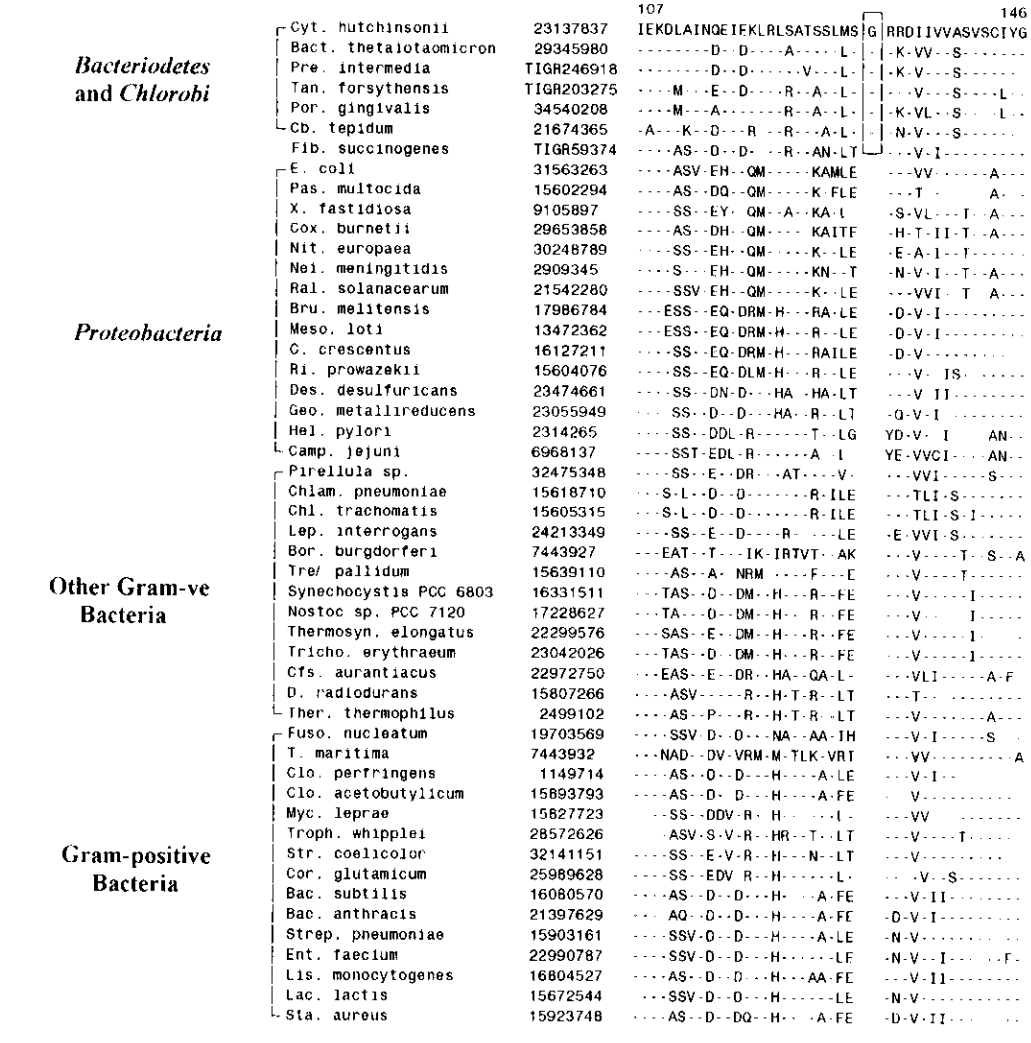

FIG. 8. Partial alignment of UvrB protein sequences showing a 1 aa insert that is commonly shared by *Bacteroidetes* and *Cb. tepidum*, but not found in any other group of bacteria. New abbreviations: *Ther.*, *Thermus*; *Tricho.*, *Trichodesmium*.

a common ancestor of these groups as marked by the arrow (Figure 10).

Another signature showing a specific relationship among *Bacteroidetes*, *Chlorobi*, and *Fibrobacteres* has been identified in the protein serine hydroxymethyltransferase (SHMT). SHMT catalyzes the reversible interconversion of serine and glycine by transferring the hydroxymethyl group of L-Ser to 5,6,7,8-tetrahydrofolate to yield Gly and 5,10-methylene-tetrahydorfolate (Fu, Boja, Safo et al. 2003). This reaction is the major source of single-carbon groups required for the synthesis of purine, thymidylate, and methionine. The homologs of SHMT are present in all main groups of bacteria. The alignment of SHMT homologs has led to identification of a 13–16 aa insert in a conserved region that is uniquely shared by the *Bacteroidetes*, *Chlorobi*, and *Fibrobacteres* species, but not found in any other groups of bacteria (Figure 11). Similar to the insert in RpoC, this insert was likely introduced in a common ancestor of these groups, exclusive of all other bacteria. These signatures provide evidence that the *Fibrobacteres*, *Chlorobi*, and *Bacteroidetes* groups (the FCB group) of bacteria are specifically related to each other.

BRANCHING ORDER OF THE FCB GROUP RELATIVE TO OTHER BACTERIAL DIVISIONS

A central aspect of understanding the evolutionary relationships among *Bacteria* is to know how different main divisions or phyla are related to each other and in what order have they branched off from a common ancestor (Gupta & Griffiths 2002; Gupta 2002). In phylogenetic trees based on 16S rRNA or various other molecules, this critical aspect of bacterial phylogeny is not resolved (Ludwig & Klenk 2001; Luding & Schleifer 1999). Consequently, it is presently believed that all main phyla within *Bacteria* have evolved at about the same time from a common ancestor (Ludwig & Klenk 2001; Hugenholtz, Pitulle, Hershberger et al. 1998; Doolittle 1999). However, in our recent work, a new

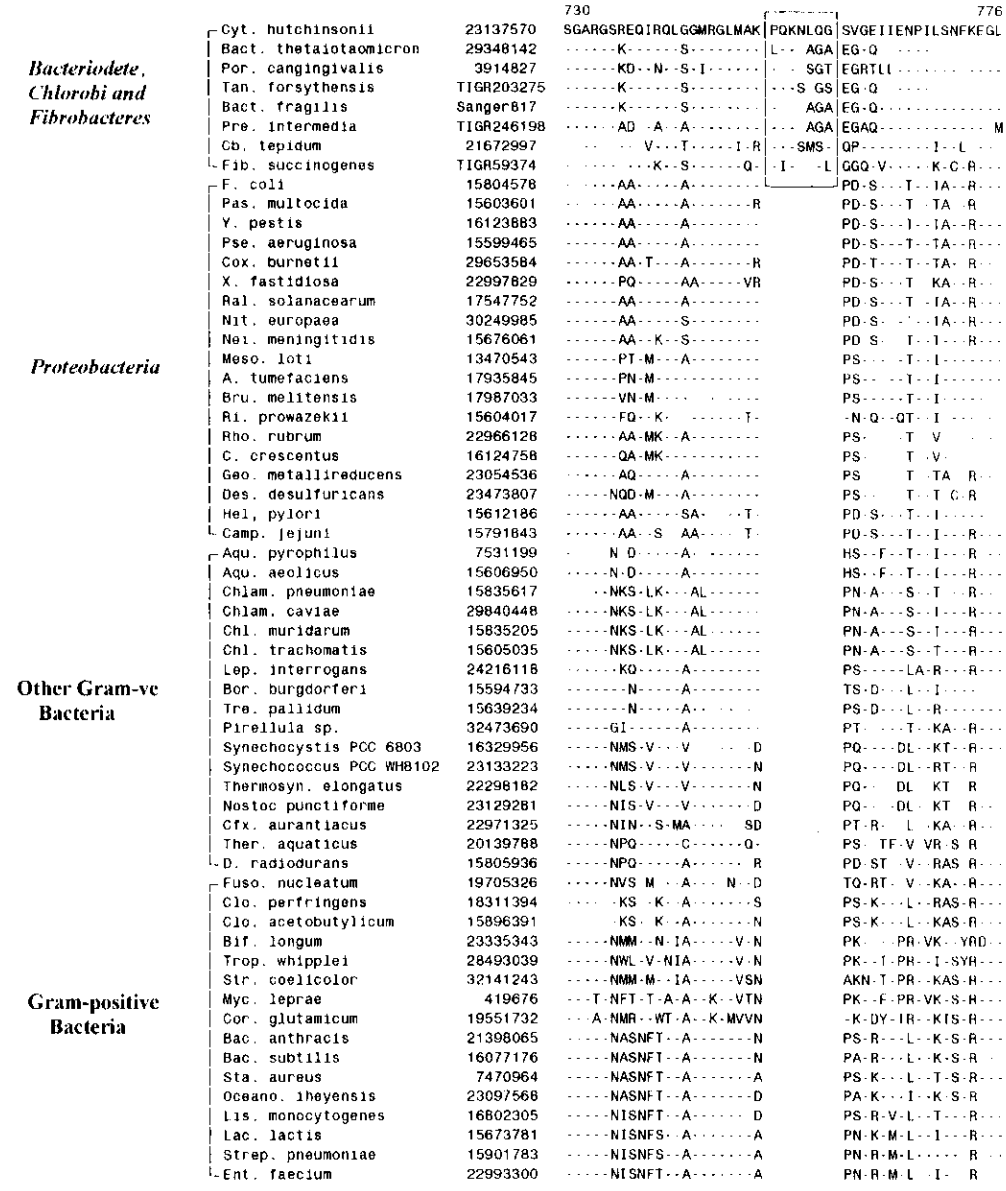

FIG. 9. Partial alignment of RNA polymerase β' (RpoC) subunit sequences showing a 6–7 aa insert (boxed) that is distinctive of the *Bacteroidetes*, *Cb. tepidum*, and *Fibrobacter* species, but not found in any other bacteria.

approach has been described involving the use of main line signatures to deduce the branching order of different groups from a common ancestor (Gupta & Griffiths 2002; Gupta 2001, 2002, 2003). Unlike group-specific signatures, main line signatures are shared by several major groups of bacteria and they are interpreted to have been introduced at major branch-points during the course of evolution (Gupta 1998; Gupta & Griffiths 2002). Once a main line signature has been introduced at a critical branch-point, all groups evolving after that will contain the indel, whereas all groups existing prior to it, or those which did not evolve from this ancestor, will not possess the indel. Thus, based on the presence or absence of various main line signatures in different groups, their relative branching orders from a common ancestor can be logically deduced. In earlier work, a large number of main line signatures have been described that are helpful in determining the branching order of different groups (Gupta 1998, 2000, 2001, 2003; Gupta & Griffiths 2002). Based on the distribution patterns of these signatures, the inferred branching orders of the different main groups is as shown in Figure 12. The signatures on which this model is based, have been described in previous work and the observed pattern is strongly supported by the predicted and observed distribution of various signatures in different genomes (Gupta 1998, 2000, 2001, 2002; Gupta & Griffiths 2002). The results for a few of the main line signatures that are particularly relevant in the present context are discussed in the following sections.

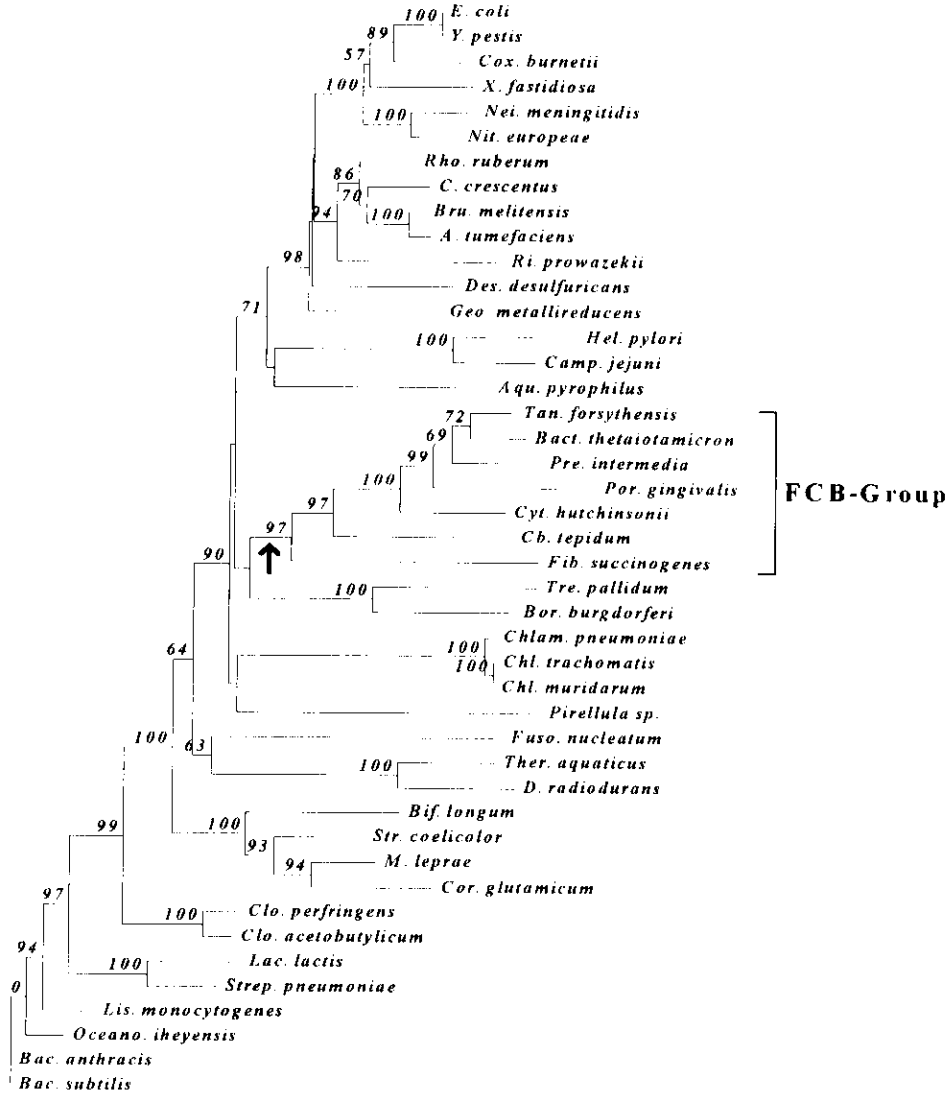

FIG. 10. A neighbor-joining phylogentic tree for RpoC protein sequences. The tree is based on 475 aligned positions, which excluded all indels. The bootstrap scores >50 are indicated. The arrow marks the evolutionary stages where the identified signature was likely introduced.

Alanyl-tRNA synthetase (AlaRS) is an essential protein required for protein synthesis, which is present in all bacterial genomes as well as in *Archaea* and eukaryotes. In AlaRS, a 4 aa indel is present in a conserved region which is commonly shared by all homologs from Proteobacteria, Chlamydiae, *Aquificales*, and the FCB group of bacteria, but it is not found in any homolog from other groups of bacteria including *Spirochetes*, *Cyanobacteria*, *Chloroflexi*, *Deinococcus-Thermus*, *Thermotogae*, *Firmicutes*, and *Actinobacteria* (Figure 13) (Gupta & Griffiths 2002; Gupta 2003; Griffiths & Gupta 2004a). Since this indel is also not present in any archael homologs, based on the rooting of the prokaryotic tree between *Archaea* and *Bacteria* (Gupta 1998; Iwabe, Kuma, Hasegawa et al. 1989; Woese, Kandler, & Wheelis 1990; Brown & Doolittle 1995; Klenk, Meier, Durovic et al. 1999) it can be inferred that the bacterial groups lacking this indel are ancestral. Therefore, this insert was likely introduced in a common ancestor of the *Proteobacteria*, *Chlamydiae*, *Aquificales*, and the FCB group of bacteria, after the branching of other groups as shown in Figure 12. Another prominent signature (a large insert of approximately 120 aa) showing exactly the same type of distribution and marking the same branch-point as seen for AlaRS, is present in the RNA polymerase β subunit (RpoB) (Gupta 2003; Klenk, Meier, Durovic et al. 1999). An additional useful signature in the present context is found present in the enzyme inorganic pyrophosphatase (PPiPase). This protein contains a 2 aa insert in a conserved region that is commonly shared by different proteobacteria and the *Aquificales* group of species, but is not found in Chlamydiae,FCB group of bacteria, or any other bacterial phyla (Figure 14) (Griffiths & Gupta 2004a). This signature provides evidence that the Chlamydiae

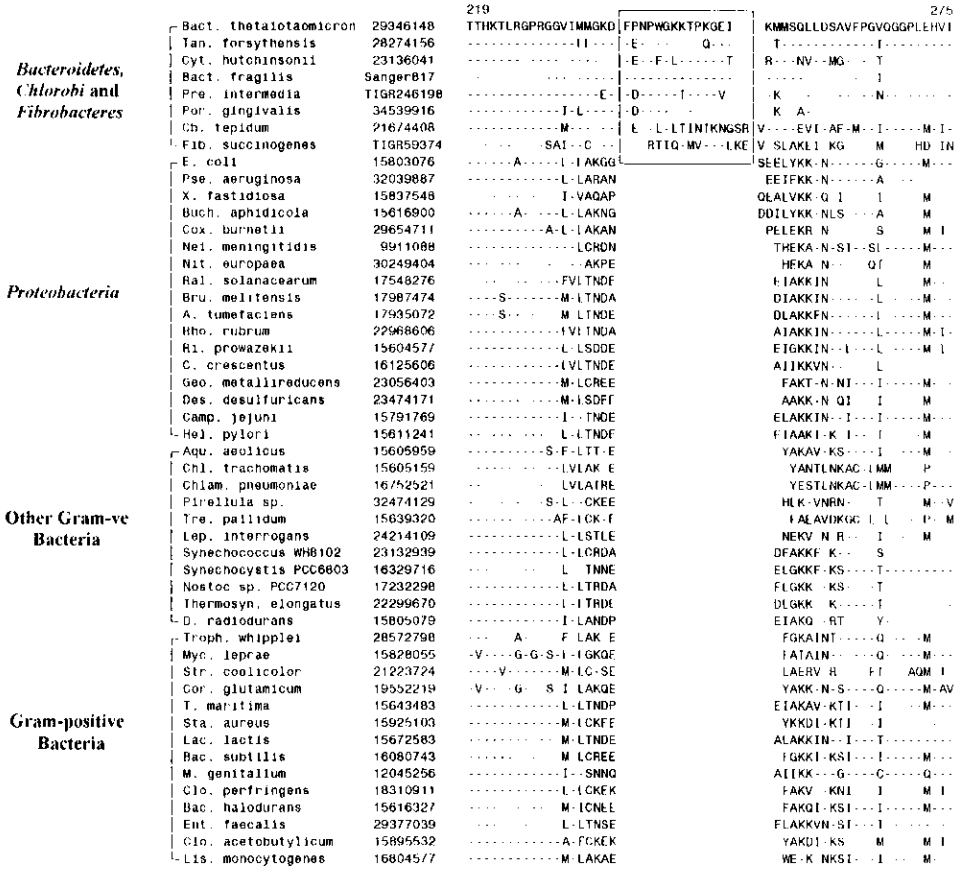

FIG. 11. Partial alignment of Serine hydroxymethyltransferase (SHMT) sequences showing a 14–16 aa insert (boxed) that is uniquely shared by the *Bacteroidetes*, *Cb. tepidum*, and *Fibrobacter* species, but not present in any other bacteria.

and the FCB groups of bacteria have branched off prior to the divergence of *Aquificales* and different groups of Proteobacteria (see Figure 12). Signatures sequences in a number of other proteins (e.g., Hsp70 and CTP synthase), provide evidence that the *Aquificales* have diverged prior to the different divisions of Proteobacteria (Gupta 2000, 2003; Gupta & Griffiths 2002; Griffiths & Gupta 2004a). The results of various signatures strongly indicate that the FCB group of bacteria have diverged at a similar time as the Chlamydiae group and their branching position relative to the other bacterial phyla is as shown in Figure 12 (Griffiths & Gupta 2001; Gupta 2003). One surprising aspect of the branching pattern shown in this figure is the late divergence of the order *Aquificales*. This is unexpected in view of the deep branching of *Aquifex* seen in the rRNA trees (Garrity & Holt 2001; Olsen, Woese, & Overbeek 1994; Olsen & Woese 1993; Cole, Chai, Marsh et al. 2003). However, as Reysenbach has noted in the recent Bergey's manual (Reysenbach 2001), the branching position of *Aquificales* is highly variable in phylogenetic trees and it cannot be reliably resolved by this approach. In contrast, the placement of the *Aquificales* in the position denoted by the signature sequence model is consistently and strongly supported by signature sequences in all of the studied proteins (Griffiths & Gupta 2004a).

CONCLUSIONS

This article describes protein signatures and phylogenetic studies based on many highly conserved proteins which serve to clarify the evolutionary relationships among the *Fibrobacteres*, *Chlorobi*, and *Bacteroidetes* (FCB) groups of bacteria, as well as their relationships to other bacteria.The FCB group of species are presently indicated to form three main phyla within *Bacteria* (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003). However, the criteria for distinguishing the main phyla have not been clearly described (Ludwig & Klenk 2001; Woese, Stackebrandt, Macke et al. 1985; Brenner, Staley, & Krieg 2001). In 16S rRNA trees, upon which bacterial phylogeny is presently based, these groups of species consistently branch within close proximity of one another and they generally exhibit sister-group relationships (Gherna & Woese 1992; Paster, Dewhrist, Olsen et al. 1994; Hugenholtz, Pitulle, Hershberger et al. 1998; Olsen, Woese, & Overbeek 1994; Olsen & Woese 1993). However, based on their branching in phylogenetic trees, which is dependent upon a large number of variables, it has not been possible to discern whether these groups were specifically related to each other or distinct from each other (Ludwig & Klenk 2001; Hugenholtz, Pitulle, Hershberger et al. 1998; Olsen, Woese, & Overbeek 1994; Cole, Chai, Marsh et al. 2003). Of these, the *Bacteroidetes* group is

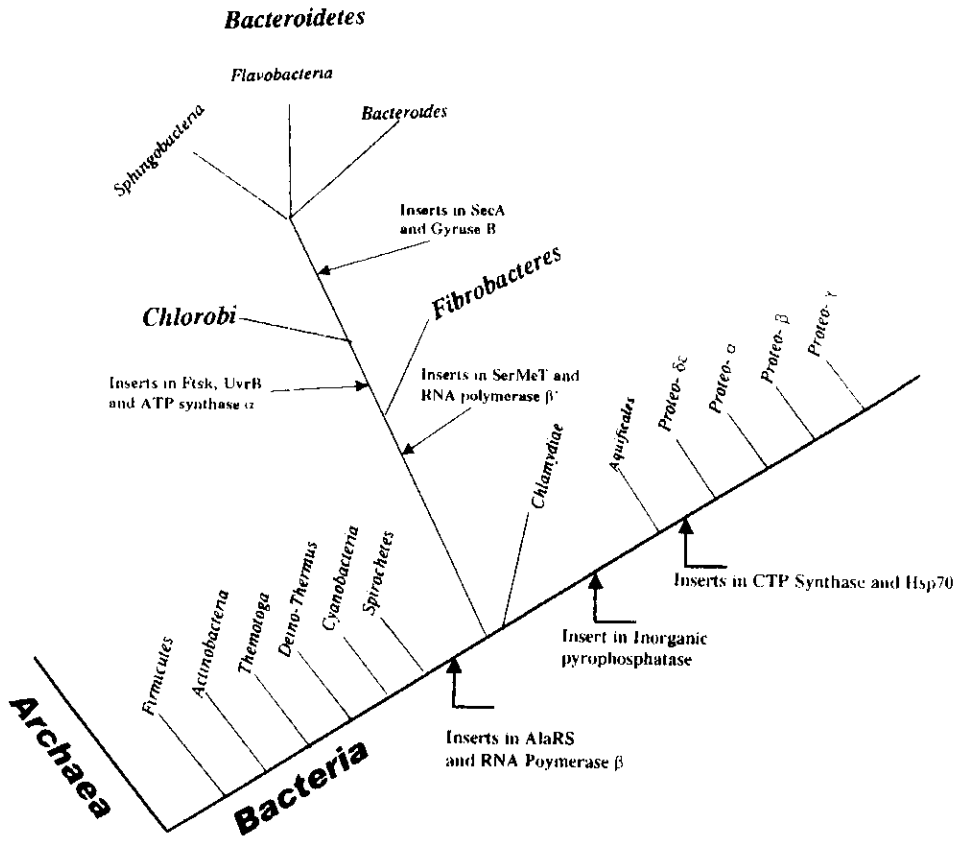

FIG. 12. Branching order of the main bacterial groups based on signature sequences in different proteins. The signature sequences used to deduce the branching order have been described in earlier work (Gupta 1998, 2002, 2003; Gupta & Griffiths 2002) and are not shown here. The thick right angle arrows mark the positions of a few main line signatures discussed in the text (Figures 13 and 14). All of the groups/phyla to the right of the arrows contain these signatures whereas those on the left do not. The evolutionary relationships among the FCB groups of species as deduced through signature sequence and phylogenetic analysis are also shown. The long thin arrows mark the positions where various signatures for these groups of bacteria have been introduced.

very large (>5000 species), while the other two in comparison are quite small, and of these the *Fibrobacteres* phylum is made up of only two-typed species (Garrity & Holt 2001; Cole, Chai, Marsh et al. 2003) In discerning the evolutionary relationship among these groups, one of the main problems that had to be faced was that there were no molecular, physiological, or phenotypic criteria available by which species belonging to these groups could be clearly distinguished from all other groups, and which could aid in clarifying their relationships to other groups of bacteria.

The protein signatures described in this review provide new and powerful means both for defining these groups of species in clear molecular terms and for elucidating their evolutionary relationships. Bacterial species belonging to the *Bacteroidetes* phylum are highly divergent in terms of their biochemical, physiological, and phenotypic characteristics, because of which the taxonomic classification or definition of this group of bacteria has proven particularly difficult (Shah 1992a, b; Holmes 1992; Reichenbach 1992). In this context, the conserved inserts in GyrB and SecA proteins that are characteristics of the *Bacteroidetes* provide molecular markers that can be used to define this phylum and for placement of new species into this group. The sequence information for GyrB is available from a large number of *Bacteroidetes* species (>100) covering all main orders and families with this group (Table 1), and all of these are found to contain the signature. Thus, this signature is a highly reliable molecular characteristic of the *Bacteroidetes* phylum. For the SecA signature, only limited sequence information is available at present, but it does include representative species from a number of different orders and families within *Bacteroidetes*. Thus, this signature is also likely to be a distinctive characteristic of the entire group.

Three different protein signatures described here viz. FtsK, UvrB, and ATP synthase (α) are uniquely shared by the *Bacteroidetes* species and *Ch. tepidum*. In phylogenetic trees based on these proteins, *Ch. tepidum* forms the outgroup of the *Bacteroidetes* phylum and a clade comprising of these two groups is strongly supported by bootstrap scores. A close relationship between *Chlorobium* and the CFB group of bacteria is also observed in the rRNA and other protein trees (Hugenholtz, Pitulle, Hershberger et al. 1998, Olsen, Woese, & Overbeek 1994; Olsen & Woese 1993; Gruber, Eisen, Gish et al. 1998; Gupta, Mukhtar,

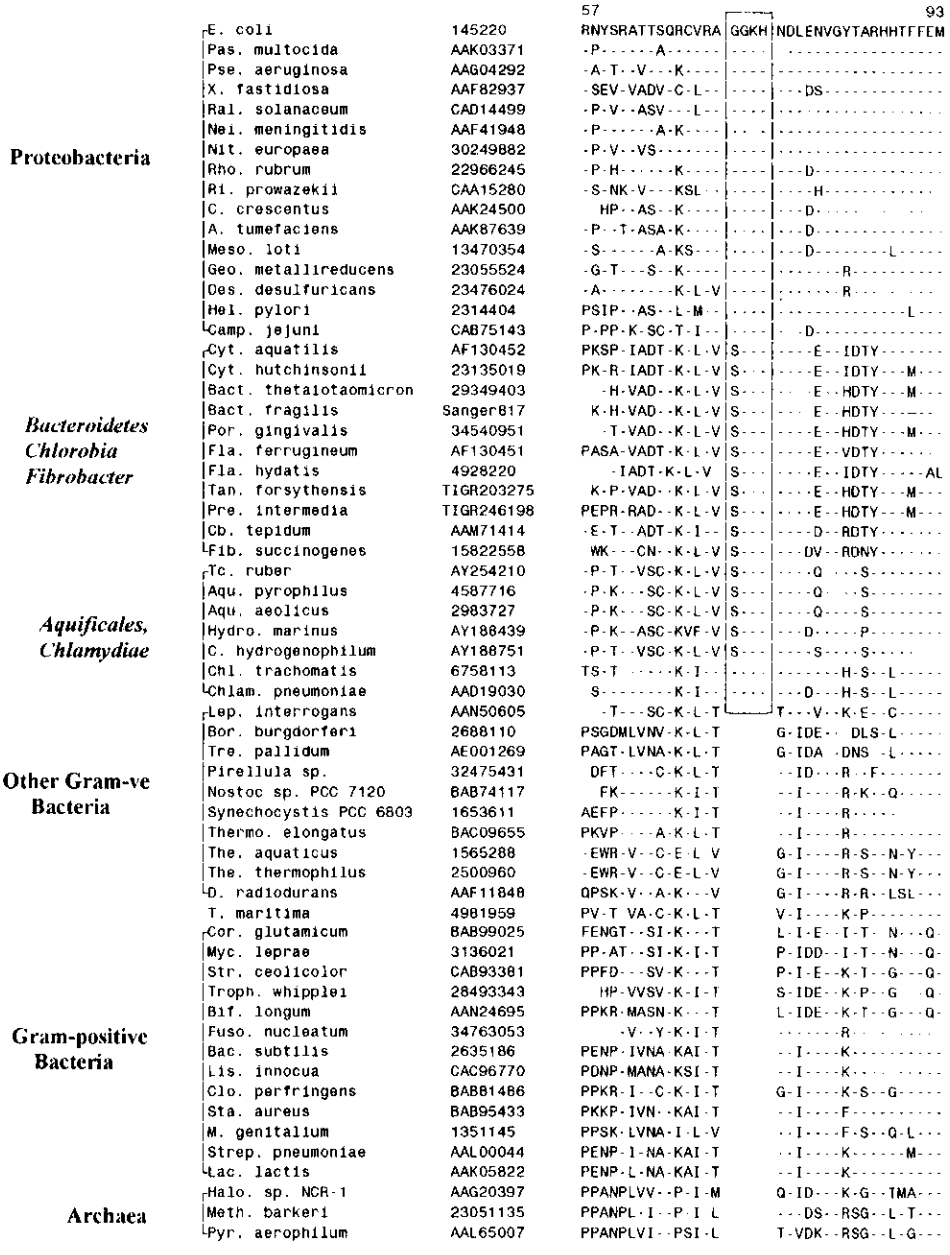

FIG. 13. Partial alignment of Alanyl tRNA synthetase sequences showing a 4 aa insert (boxed) that is commonly shared by various Proteobacteria, Chlamydiae, FCB group of bacteria and *Aquificales* species, but not found in any other bacteria. This insert is also not found in any archaeal homologs and it was likely introduced in a common ancestor of the above bacterial groups after the branching of Spirochetes, Cyanobacteria, *Chloroflexi*, *Deinococcus-Thermus*, *Thermotogae*, *Firmicutes*, and *Actinobacteria*. Another large insert showing similar species distribution is present in the RNA polymerase β subunit (Gupta 2003; Griffiths and Gupta 2004a).

& Singh 1999). These observations provide strong evidence that these two groups are specifically related to each other and the identified signatures were introduced in a common ancestor of these two groups exclusive of other bacteria.

Lastly, the signatures in RpoC and SHMT proteins, which are commonly shared by the *Bacteroidetes*, *Chlorobium*, and *Fibrobacteres* species, provide evidence that these groups shared a common ancestor exclusive of all others. A clade comprising of these groups of species is also strongly supported by phylogenetic analysis based on the above proteins as well as others. A closer affinity of the *Fibrobacteres* to the *Chlorobi* and the CFB-division of bacteria is also seen in the 16S rRNA trees (Olsen, Woese, & Overbeek 1994). Further evidence that these groups of bacteria are specifically related to each other is provided by the main line signatures in different proteins (Griffiths and Gupta 2001). As noted earlier, a large number of main line signatures have been identified and their distribution patterns enable one to deduce the relative branching order of the main

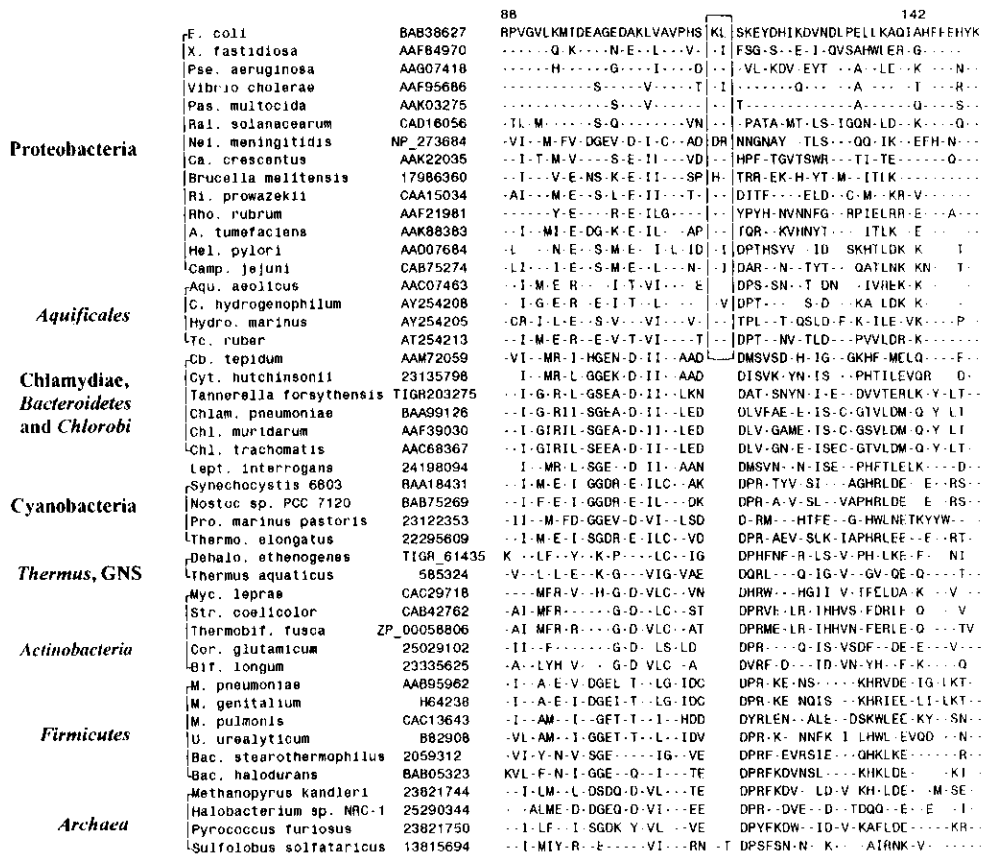

FIG. 14. Partial alignment of Inorganic pyrophosphatase sequences showing a 2 aa insert (boxed) that is commonly shared by various Proteobacteria, and *Aquificales* species, but not found in any other bacteria (Griffiths & Gupta 2004a). This insert was likely introduced in a common ancestor of these bacteria after the branching of *Chlamydiae*, FCB groups of bacteria, *Spirochetes*, *Cyanobacteria*, *Chloroflexi*, *Deinococcus-Thermus*, *Thermotogae*, *Firmicutes*, and *Actinobacteria*.

groups of bacteria. The analyses of these signatures, some examples of which are given here, again provide evidence that the species belonging to the above three groups branch in the same position (Gupta 2001, 2003; Griffiths and Gupta 2001). Thus, a close and specific relationship among these groups is strongly supported by a large number of both group-specific and main line signatures as well as in different phylogenetic trees.

These observations raise the question whether these bacteria should be placed in three independent main groups (viz. *Fibrobacteres*, *Chlorobi*, and *Bacteroidetes*) as is being done currently, or whether these organisms should all be placed in a single phylum (viz. FCB). To my knowledge, no strong arguments or rationales have been provided for the placement of these bacteria into distinct phyla. On the other hand, results presented here strongly indicate that these bacteria are specifically related to each other and they should be placed in a single phylum indicating their relationship. The resolution of this issue requires a clear description of the criteria used for identification the main phyla and applying them uniformly in all cases. As noted above, presently there are no specified criteria for identifying the main bacterial phyla (Ludwig & Klenk 2001; Woese, Stackebrandt, Macke et al. 1985; Brenner, Staley, & Krieg 2001), which has given rise to the problem at hand. To rectify this problem, based on our work with signature sequences, we have suggested a precise definition for the main groups or phyla within *Bacteria*. According to the definition that we have suggested, the main bacterial phyla should correspond to the major branches in the trees, each of which should be clearly distinct from the others in terms of its branching position as well as group-specific signatures or characteristics (Gupta 2000, 2002; Gupta & Griffiths 2002). Based on these criteria, all of the main groups shown in Figure 12, which have diverged from the main bacterial stem at different times, and which can also be distinguished by group-specific signatures, (Gupta 2000; Griffiths & Gupta 2002, 2004b; Gupta, Pereira, Chandrasekera et al. 2003) qualify as the main group or phyla. This definition also suggests that the presently recognized Proteobacterial phylum should be divided into four main phyla, each of which can be clearly distinguished from the other based on their branching order as well as group-specific signatures (Gupta 2000; Gupta & Griffiths 2002). Applying these criteria to the FCB groups of bacteria, all of which branch in the same position and form a monophyletic group based on different signature sequences and phylogenetic trees, indicate that they should all be placed in a single phylum

and not three independent phyla, as is presently done. Within this phylum, the species corresponding to these groups have branched in the following order—*Fibrobacteres* → *Chlorobi* → *Bacteroidetes* (Figure 12)—an inference strongly supported by different signatures and phylogenetic trees. Further, based on the main line signatures that have thus far been identified, the FCB group branch in a similar position as the Chlamydiae. Although Chlamydiae are distinct from the FCB group, based on a large number of signatures (Griffiths & Gupta 2002) and their unique life-cycle (Fields & Barnes 1992), in phylogenetic trees based on 16S rRNA and various proteins, these groups of species generally branch in close proximity to each other (Hugenholtz, Pitulle, Hershberger et al. 1998; Gupta, Mukhtar, & Singh 1999; Olsen, Woese, & Overbeek 1994; Olsen & Woese 1993; Graber, Eisen, Gish et al. 1998; Viale, Arakaki, Soncini et al. 1994). Thus, it is possible that these two groups of bacteria may have diverged from a common ancestor exclusive of the other bacteria. In our proposal, we have suggested that the distinct groups of bacteria branching in the same position (viz., FCB group and Chlamydiae) should be tentatively regarded as subdivisions of the same main branch or phylum (Gupta 2000; Griffith & Gupta 2002). However, if further studies clarify their relative branching orders from the main stem, then these groups can be assigned to distinct main phylum.

ACKNOWLEDGMENTS

Thanks are due to a large number of undergraduate students, particularly Adeel Mahmood, Ghezal Fariq, Chris Hennie, and Vanessa Johari, who have helped me with BLAST searches and generation of sequence alignments. I am also thankful to Julie Paul for excellent secretarial assistance, to Emma Griffiths for helpful discussions, and to Dr. Bhag Singh for assorted computer-related and other assistance. This work was supported by research grant from the Canadian Institute of Health Research and the National Science and Engineereing Research Council of Canada.

REFERENCES

Alexander, B., Andersen, J.H., Cox, R.P., and Imhoff, J.F. 2002. Phylogeny of green sulfur bacteria on the basis of gene sequences of 16S rRNA and of the Fenna—Matthews-Olson protein. *Arch. Microbiol.* 178, 131–140.

Amann, R.I., Lin, C., Key, R., Montgomery, L., and Stahl, D.A. 1992. Diversity among Fibrobacter isolates: Towards a phylogenetic classification. *Syst. Appl. Microbiol.* 15, 23–31.

Anton, J., Rossello-Mora, R., Rodriguez-Valera, F., and Amann, R. 2000. Extremely halophilic bacteria in crystallizer ponds from solar salterns. *Appl. Environ. Microbiol.* 66, 3052–3057.

Baldauf, S.L., and Palmer, J.D. 1993. Animals and fungi are each other's closest relatives: Congruent evidence from multiple proteins. *Proc. Natl. Acad. Sci. USA.* 90, 11558–11562.

Balows, A., Trüper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H. 1992. *The Prokaryotes.* Springer-Verlag, New York.

Barbieri, E., Potenza, L., Rossi, I., Sisti, D., Giomaro, G., Rossetti, S., Beimfohr, C., and Stocchi, V. 2000. Phylogenetic characterization and in situ detection of a Cytophaga-Flexibacter-Bacteroides phylogroup bacterium in Tuber borchii vittad. Ectomycorrhizal mycelium. *Appl. Environ. Microbiol.* 66, 5035–5042.

Blankenship, R.E. 1992. *Origin and early evolution of photosynthesis. Photosynthesis Research.* 33, 91–111.

Blankenship, R.E. 1994. Protein structure, electron transfer and evolution of prokaryotic photosynthetic reaction centers. Antonie van Leeuwenhoek. 65, 311–329.

Boone, D.R., Castenholz, R.W., and Garrity, G.M. 2001. *Bergey's Manual of Systematic Bacteriology.* Springer, New York.

Brenner, D.J., Staley, J.T., and Krieg, N.R. 2001. Classification of Procaryotic Organisms and the concept of Bacterial Speciation. In *Bergey's Manual of Systematic Bacteriology*, Boone, D.R., and Castenholz, R.W., Eds., Springer-Verlag, Berlin, 27–31.

Brown, J.R., and Doolittle, W.F. 1995. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proc. Natl. Acad. Sci. USA.* 92, 2441–2445.

Brown, J.R., and Doolittle, W.F. 1997. Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Rev.* 61, 456–502.

Brown, J.R., Masuchi, Y., Robb, F.T., and Doolittle, W.F. 1994. Evolutionary relationships of bacterial and archaeal glutamine synthetase genes. *J. Mol. Evol.* 38, 566–576.

Capiaux, H., Lesterlin, C., Perals, K., Louarn, J.M., and Cornet, F. 2002. A dual role for the FtsK protein in Escherichia coli chromosome segregation. *EMBO Rep.* 3, 532–536.

Cole, J.R., Chai, B., Marsh, T.L., Farris, R.J., Wang, Q., Kulam, S.A., Chandra, S., McGarrell, D.M., Schmidt, T.M., Garrity, G.M., and Tiedje, J.M. 2003. The Ribosomal Database Project (RDP-II): Previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy. *Nucleic Acids Res.* 31, 442–443.

Corsaro, D., Valassina, M., and Venditti, D. 2003. Increasing diversity within Chlamydiae. *Crit Rev. Microbiol.* 29, 37–78.

Doolittle, W.F. 1999. Phylogenetic classification and the universal tree. *Science* 284, 2124–2128.

Duncan, M.J. 2003. Genomics of oral bacteria. *Crit Rev. Oral Biol. Med.* 14, 175–187.

Eisen, J.A. 1995. The RecA protein as a model molecule for molecular systematic studies of bacteria: Comparison of trees of RecAs and 16S rRNAs from the same species. *J. Mol. Evol.* 41, 1105–1123.

Eisen, J.A., Nelson, K.E., Paulsen, I.T., Heidelberg, J.F., Wu, M., Dodson, R.J., DeBoy, R., Gwinn, M.L., Nelson, W.C., Haft, D.H., Hickey, E.K., Peterson, J.D., Durkin, A.S., Kolonay, J.L., Yang, F., Holt, I., Umayam, L.A., Mason, T., Brenner, M., Shea, T.P., Parksey, D., Nierman, W.C., Feldblyum, T.V., Hansen, C.L., Craven, M.B., Radune, D., Vamathevan, J., Khouri, H., White, O., Gruber, T.M., Ketchum, K.A., Venter, J.C., Tettelin, H., Bryant, D.A., and Fraser, C.M. 2002. The complete genome sequence of Chlorobium tepidum TLS, a photosynthetic, anaerobic, green-sulfur bacterium. *Proc. Natl. Acad. Sci. USA* 99, 9509–9514.

Espeli, O., Lee, C., and Marians, K.J. 2003. A physical and functional interaction between Escherichia coli FtsK and topoisomerase IV. *J. Biol. Chem.* 278, 44639–44644.

Felsenstein, J. 1985. Confidence limits in phylogenies: An approach using the bootstap, Evolution. 39, 783–791.

Fields, P.I., And Barnes, R.C. 1992. The Genus Chlamydia, p. 3691–3709. In Balows, A., Trüper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H., Eds., The Prokaryotes. Springer-Verlag, New York.

Frostl, J.M., and Overmann, J. 2000. Phylogenetic affiliation of the bacteria that constitute phototrophic consortia. *Arch. Microbiol.* 174, 50–58.

Fu, T.F., Boja, E.S., Safo, M.K., and Schirch, V. 2003. Role of proline residues in the folding of serine hydroxymethyltransferase. *J. Biol. Chem.* 278, 31088–31094.

Fukushima, M., Kakinuma, K., and Kawaguchi, R. 2002. Phylogenetic analysis of Salmonella, Shigella, and Escherichia coli strains on the basis of the gyrB gene sequence. *J. Clin. Microbiol.* 40, 2779–2785.

Garrity, G.M., and Holt, J.G. 2001. The road map to the manual. In *Bergey's Manual of Systematic Bacteriology*, Boone, D.R., and Castenholz, R.W., Eds., Springer-Verlag, Berlin, 119–166.

Gherna, R., and Woese, C.R. 1992. A Partial phylogenetic analysis of the "flavobacter-bacteroides" phylum: Basis for taxonomic restructuring. *Syst. Appl. Microbiol.* 15, 513–521.

Griffiths, E., and Gupta, R.S. 2001. The use of signature sequences in different proteins to determine the relative branching order of bacterial divisions: Evidence that Fibrobacter diverged at a similar time to Chlamydia and the Cytophaga-Flavobacterium-Bacteroides division. *Microbiology*, 147, 2611–2622.

Griffiths, E., and Gupta, R.S. 2002. Protein signatures distinctive of chlamydia species: Horizontal transfer of cell wall biosynthesis genes glumU from Archaebacteria to Chlamydiae, and murA between Chlamydiae and Streptomyces. *Microbiology*, 148, 2541–2549.

Griffiths, E., and Gupta, R.S. 2004a. Signature sequences in diverse proteins provide evidence for the late divergence of the order Aquificales. *Inter. Microbiol.* 7, 41–52.

Griffiths, E., and Gupta, R.S. 2004b. Distinctive protein signatures provide molecular markers and evidence for the monophyletic nature of the Deinococcus-Thermus phylum. *J. Bacteriol.* 186 (in press).

Gruber, T.M., Eisen, J.A., Gish, K., and Bryant, D.A. 1998. The phylogenetic relationships of Chlorobium tepidum and Chloroflexus auranticus based upon their RecA sequences. *FEMS Microbiol. Lett.* 162, 53–60.

Gupta, R.S. 1995. Evolution of the chaperonin families (Hsp60, Hsp10 and /Tcp-1) of proteins and the origin of eukaryotic cells. *Mol. Microbiol.* 15, 1–11.

Gupta, R.S. 1998. Protein Phylogenies and Signature Sequences: A reappraisal of evolutionary relationships among archaebacteria, eubacteria, and eukaryotes. *Microbiol. Mol. Biol. Rev.* 62, 1435–1491.

Gupta, R.S. 2000. The phylogeny of Proteobacteria: Relationships to other eubacterial phyla and eukaryotes, FEMS *Microbiol. Rev.* 24, 367–402.

Gupta, R.S. 2001. The branching order and phylogenetic placement of species from completed bacterial genomes, based on conserved indels found in various proteins. *Inter. Microbiol.* 4, 187–202.

Gupta, R.S. 2002. Phylogeny of Bacteria: Are we now close to understanding it?. *ASM News* 68, 284–291.

Gupta, R.S. 2003. Evolutionary Relationships Among Photosynthetic Bacteria. *Photosynth. Res.* 76, 173–183.

Gupta, R.S., and Singh, B. 1992. Cloning of the HSP70 gene from Halobacterium marismortui: Relatedness of archaebacterial HSP70 to its eubacterial homologs and model for the evolution of the HSP70 gene. *J. Bacteriol.* 174, 4594–4605.

Gupta, R.S., and Singh, B. 1994. Phylogenetic analysis of 70 kD heat shock protein sequences suggests a chimeric origin for the eukaryotic cell nucleus. *Curr. Biol.* 4, 1104–1114.

Gupta, R.S., and Griffiths, E. 2002. Critical Issues in Bacterial Phylogenies. *Theor. Popul. Biol.* 61, 423–434.

Gupta, R.S., Bustard, K., Falah, M., and Singh, D. 1997. Sequencing of heat shock protein 70 (Dnak) homologs from Deinococcus proteolyticus and Thermomicrobium roseum and their integration in a protein-based phylogeny of prokaryotes. *J. Bacteriol.* 179, 345–357.

Gupta, R.S., Pereira, M., Chandrasekera, C., and Johari, V. 2003. Molecular signatures in protein sequences that are characteristic of Cyanobacteria and plastid homologues. *Int. J. Syst. Evol. Microbial.* 53, 1833–1842.

Gupta, R.S., Mukhtar, T., and Singh, B. 1999. Evolutionary relationships among photosynthetic prokaryotes (Heliobacterium chlorum, Chloroflexus aurantiacus, Cyanobacteria, Chlorobium tepidum and proteobacteria): Implications regarding the origin of photosynthesis. *Mol. Microbiol.* 32, 893–906.

Holdeman, L.V., Kelley, R.W., and Moore, W.E.C. 1984. Family I. Bacteroidaceae Pribam 1933. In *Bergey's Manual of Systematic Bacteriology*, Krieg, N.R., and Holt, J.G., Eds., Williams and Wilkins, Baltimore, 602–662.

Holmes, B. 1992. The Genera Flavobacterium, Sphingobacterium, and Weeksella, p. 3620–3620. In Balows, A., Truper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H., Eds., The Prokaryotes. Springer-Verlag, New York.

Hugenholtz, P., Pitulle, C., Hershberger, K.L., and Pace, N.R. 1998. Novel division level bacterial diversity in a Yellowstone hot spring. *J. Bacteriol.* 180, 366–376.

Humayoun, S.B., Bano, N., and Hollibaugh, J.T. 2003. Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California. *Appl. Environ. Microbiol.* 69, 1030–1042.

Hungate, R.E. 1950. The anaerobid mesophilic celluloytic bacteria, *Bacteriol. Rev.* 14, 1–49.

Igarashi, N., Harada, J., Nagashima, S., Matsuura, K., Shimada, K., and Nagashima, K.V.P. 2001. Horizontal transfer of the photosynthesis gene cluster and operon rearrangement in purple bacteria. *J. Mol. Evol.* 52, 333–341.

Imhoff, J.F. 2003. Phylogenetic taxonomy of the family chlorobiaceae on the basis of 16S rRNA and fmo (Fenna-Matthews-Olson protein) gene sequences. *Int. J. Syst. Evol. Microbiol.* 53, 941–951.

Iwabe, N., Kuma, K., Hasegawa, M., Osawa, S., and Miyata, T. 1989. Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. USA.* 86, 9355–9359.

Karlin, S., and Brocchieri, L. 1998. Heat shock protein 70 family: Multiple sequence comparisons, function, and evolution. *J. Mol. Evol.* 47, 565–577.

Kasai, H., Watanabe, K., Gasteiger, E., Bairoch, A., Isono, K., Yamamoto, S., and Harayama, S. 1998. Construction of the gyrB Database for the Identification and Classification of Bacteria, Genome Inform. Ser. Workshop Genome Inform. 9, 13–21.

Klenk, H.P., and Zillig, W. 1994. DNA-dependent RNA polymerase subunit B as a tool for phylogenetic reconstructions: Branching topology of the archaeal domain. *J. Mol. Evol.* 38, 420–432.

Klenk, H.P., Meier, T.D., Durovic, P., Schwass, V., Lottspeich, F., Dennis, P.P., and Zillig, W. 1999. RNA polymerase of Aquifex pyrophilus: Implications for the evolution of the bacterial rpoBC operon and extremely thermophilic bacteria. *J. Mol. Evol.* 48, 528–541.

Krause, D.O., and Russell, J.B. 1996. How many ruminal bacteria are there? *J. Dairy Sci.* 79, 1467–1475.

Leyva, J.A., Bianchet, M.A., and Amzel, L.M. 2003. Understanding ATP synthesis: Structure and mechanism of he F1-ATPase (Review). *Mol. Membr. Biol.* 20, 27–33.

Lin, C., and Stahl, D.A. 1995. Comparative analyses reveal a highly conserved endoglucanase in the celluloytic genus Fibrobacter. *J. Bacteriol.* 177, 2543–2549.

Ludwig, W., and Klenk, H.-P. 2001. Overview: A phylogenetic backbone and taxonomic framework for prokaryotic systamatics. In *Bergey's Manual of Systematic Bacteriology*, Boone, D.R., and Castenholz, R.W., Eds., Springer-Verlag, Berlin, 49–65.

Ludwig, W., and Schleifer, K.H. 1999. Phylogeny of Bacteria beyond the 16S rRNA Standard. *ASM News* 65, 752–757.

Meyer, T.E., Cusanovich, M.A., and Kamen, M.D. 1986. Evidence against use of bacterial amino acid sequence data for construction of all-inclusive phylogenetic trees. *Proc. Natl. Acad. Sci. USA.* 83, 217–220.

Montgomery, L., Flesher, B., and Stahl, D.A. 1988. Transfer of Bacteroides succinogenes (Hungate) to Fibrobacter gen.nov.as Fibrobacter succinogenes comb. Nov. and Description of Fibrobacter intestinalis sp. nov., EPA/600/J-88/357. *Int. J. Syst. Bacteriol.* 38, 430–435.

Morse, R., O'Hanlon, K., and Collins, M.D. 2002. Phylogenetic, amino acid content and indel analyses of the beta subunit of DNA-dependent RNA polymerase of gram-positive and gram-negative bacteria. *Int. J. Syst. Evol. Microbiol.* 52, 1477–1484.

Murray, R.G.E., and Stackebrandt, E. 1995. Taxonomic Note: Implementation of the provisional status canadidatus for incompletely described prokaryotes. *Int. J. Syst. Bacteriol.* 45, 186–187.

NCBI Entrez Taxonomy Homepage. http://www.ncbi.nlm.nih.gov/entrez/query.fegi?db=Taxonomy. 2001.

Nelson, K.E., Fleischmann, R.D., DeBoy, R.T., Paulsen, I.T., Fouts, D.E., Eisen, J.A., Daugherty, S.C., Dodson, R.J., Durkin, A.S., Gwinn, M., Haft, D.H., Kolonay, J.F., Nelson, W.C., Mason, T., Tallon, L., Gray, J., Granger, D.,

Tettelin, H., Dong, H., Galvin, J.L., Duncan, M.J., Dewhirst, F.E., and Fraser, C.M. 2003. Complete genome sequence of the oral pathogenic bacterium Porphyromonas gingivalis strain W83. *J. Bacteriol.* 185, 5591–5601.

Ohkuma, M., Noda, S., Hongoh, Y., and Kudo, T. 2002. Diverse bacteria related to the bacteroides subgroup of the CFB phylum within the gut symbiotic communities of various termites. *Biosci. Biotechnol. Biochem.* 66, 78–84.

Olsen, G.J., and Woese, C.R. 1993. Ribosomal RNA: A key to phylogeny. *FASEB Journal.* 7, 113–123.

Olsen, G.J., and Woese, C.R. 1997. Archaeal genomics: An overview. *Cell.* 89, 991–994.

Olsen, G.J., Woese, C.R., and Overbeek, R. 1994. The winds of (evolutionary) change: Breathing new life into. *Microbiology. Bacteriol.* 176, 1–6.

O'Sullivan, L.A., Weightman, A.J., and Fry, J.C. 2002. New degenerate Cytophaga-Flexibacter-Bacteroides-specific 16S ribosomal DNA-targeted oligonucleotide probes reveal high bacterial diversity in River Taff epilithon. *Appl. Environ. Microbiol.* 68, 201–210.

Overmann, J., and Tuschak, C. 1997. Phylogeny and molecular fingerprinting of green sulfur bacteria. *Arch. Microbiol.* 167, 302–309.

Paster, B.J., Dewhirst, F.E., Olsen, I., and Fraser, G.J. 1994. Phylogeny of bacteroides, prevotella, and prophyromonas spp.and related bacteria. *J. Bacteriol.* 176, 725–732.

Pierson, B.K., and Castenholz, R.W. 1992. The Family Chloroflexaceae. In *The Prokaryotes.* Balows, A., Truper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H., Eds., Springer-Verlag, New York, 3754–3775.

Raymond, J., Zhaxybayeva, O., Gogarten, J.P., Gerdes, S.Y., and Blankenship, R.E. 2002. Whole-genome analysis of photosynthetic prokaryotes. *Science* 298, 1616–1620.

Reichenbach, H. 1992. The Order Cytophagales, p. 3631–3675. In Balows, A., Truper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H., Eds., The Prokaryotes. Springer-Verlag, New York.

Reichenbach, H., and Weeks, O.B. 1981. The Flavobacterium-Cytophaga group. Weinheim, Germany, Verlag Chemie.

Reysenbach, A.-L. 2001. Phylum BI. Aquificae phy. Nov. In *Bergey's Manual of Systematic Bacteriology.* Boone, D.R. and Castenholz, R.W. Eds., Springer-Verlag, Berlin, 359–367.

Rivera, M.C., and Lake, J.A. 1992. Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. *Science* 257, 74–76.

Saitou, N., and Nei, M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Sawada, H., Suzuki, F., Matsuda, I., and Saitou, N. 1999. Phylogenetic analysis of Pseudomonas syringae pathovars suggests the horizontal gene transfer of argK and the evolutionary stability of hrp, gene cluster. *J. Mol. Evol.* 49, 627–644.

Schmidt, M.G., and Kiser, K.B. 1999. SecA: The ubiquitous component of preprotein translocase in prokaryotes, *Microbes. Infect.* 1, 993–1004.

Seo, P.S., and Yokota, A. 2003. The phylogenetic relationships of cyanobacteria inferred from 16S rRNA, gyrB, rpoC1 and rpoD1 gene sequences. *J. Gen. Appl. Microbiol.* 49, 191–203.

Shah, R.N. 1992a. The Genus Bacteroides and Related taxa, p. 3593–3607. In Balows, A., Truper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H., Eds., The Prokaryotes. Springer-Verlag, New York.

Shah, R.N. 1992b. The Genus Prophyromonas, p. 3608–3619. In Balows, A., Truper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H., Eds., The Prokaryotes. Springer-Verlag, New York.

Shah, R.N., and Collins, M.D. 1998. Proposal for reclassification of Bacteriodes asaccharolyticus, Bacteriodes gingivalis, and Bacteroides endodontalis in a new genus, Prophyromonas. *Int. J. Syst. Bacteriol.* 38, 128–131.

Shah, R.N., and Collins, M.D. 1989. Proposal to restrict the genus Bacteroides (Castellani and Chalmers) to Bacteroides fragilis and closely related species. *Int. J. Syst. Bacteriol.* 39, 85–87.

Shah, R.N., and Collins, M.D. 1990. Prevotella, a new genus to include Bacteroides melaninogenicus and related species formerly classified in the genus Bacteroides. *Int. J. Syst. Bacteriol.* 40, 205–208.

Shah, R.N., Gharbia, S.E., and Duerden, B.I. 1998. Bacteroides, Prevotella and Porphyromonas. In *Topley & Wilson's Microbiology and Microbial Infections*, Vol. 2, Balows, A., and Duerden, B.I., Eds., Systematic Bacteriology. Arnold, London, 1305–1330.

Suzuki, M., Nakagawa, Y., Harayama, S., and Yamamoto, S. 1999. Phylogenetic analysis of genus Marinilabilia and related bacteria based on the amino acid sequences of gyrB and emended description of Marinilabilia salmonicolor with Marinilabilia agarovorans as its subjective synonym. *Int. J. Syst. Bacteriol.* 49 Pt 4, 1551–1557.

Suzuki, M., Nakagawa, Y., Harayama, S., and Yamamoto, S. 2001. Phylogenetic analysis and taxonomic study of marine Cytophaga-like bacteria: Proposal for Tenacibaculum gen. nov with Tenacibaculum martitimum comb. nov and Tenacibaculum ovolyticum comb. nov., and description of Tenacibaculum mesophilum sp nov and Tenacibaculum amylolyticum sp nov. *Int. J. Syst. Evol. Microbiol.* 51, 1639–1652.

Truper, H.G., and Pfenning, N. 1992. The Family Chlorobiaceae, p. 3583–3592. In Balows, A., Truper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H., Eds., The Prokaryotes. Springer-Verlag, New York.

Valentin, K. 1997. Phylogeny and expression of he sec A gene from a chromophytic algaimplications for the evolution of plastids and sec—dependent protein translocation. *Curr. Genet.* 32, 300–307.

Van de, P.Y., and De Wachter, R. 1994. TREECON for Windows: A software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput. Appl. Biosci.* 10, 569–570.

Ventosa, A., Nieto, J.J., and Oren, A. 1998. Biology of moderately halophilic aerobic bacteria, *Microbiol. Mol. Biol. Rev.* 62, 504–544.

Verhoeven, E.E., Wyman, C., Moolenaar, G.F., and Goosen, N. 2002. The presence of two UvrB subunits in the UvrAB complex ensures damage detection in both DNA strands. *EMBO J.* 21, 4196–4205.

Viale, A.M., Arakaki, A.K., Soncini, F.C., and Ferreyra, R.G. 1994. Evolutionary relationships among eubacterial groups as inferred from GroEL (chaperonin) sequence comparisons. *Int. J. Syst. Bacteriol.* 44, 527–533.

Wang, J.C. 1996. DNA topoisomerases, *Annu. Rev. Biochem.* 65, 635–692.

Wigley, D.B. 1995. Structure and mechanism of DNA topoisomerases. *Annu. Rev. Biophys. Biomol. Struct.* 24, 185–208.

Woese, C.R. 1987. Bacterial evolution. *Microbiol. Rev.* 51–271.

Woese, C.R., Stackebrandt, E., Macke, R.J., and Fox, G.E. 1985. A phylogenetic definition of he major eubacterial taxa. system. *Appl. Microbiol.* 6, 143–151.

Woese, C.R., Kandler, O., and Wheelis, M.L. 1990. Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. USA.* 87, 4576–4579.

Xu, J., Bjursell, M.K., Himrod, J., Deng, S., Carmichael, L.K., Chiang, H.C. Hooper, L.V., and Gordon, J.I. 2003. A genomic view of the human-Bacteroides thetaiotaomicron symbiosis. *Science* 299, 2074–2076.

Yamamoto, S., Kasai, H., Arnold, D.L., Jackson, R.W., Vivian, A., and Harayama, S. 2000. Phylogeny of the genus Pseudomonas: Intrageneric structure reconstructed from the nucleotide sequences of gyrB and rpoD genes. *Microbiology.* 146 (Pt 10), 2385–2394.

Yáñez, M., Catalán, A., Apráiz, D., Figueras, M.J., and Martínez-Murcia, A.J. 2003. Phylogenetic analysis of members of the genus Aeromonas based on gyrB gene sequences. *Int. J. Syst. Evol. Microbiol.* 53, 875–883.

Zinder, S.H. 1998. Bacterial Diversity, p. 125–147. In Balows, A., and Duerden, B.I., Eds., Arnold, London.