Intrageneric Relationships of Members of the Genus *Fusobacterium* as Determined by Reverse Transcriptase Sequencing of Small-Subunit rRNA

PAUL A. LAWSON, SAHEER E. GHARBIA, HAROUN N. SHAH, DUNCAN R. CLARK, and MATTHEW D. COLLINS

Department of Oral Microbiology, The London Hospital Medical College, London E1 2AD, and Department of Microbiology, AFRC Institute of Food Research, Reading Laboratory, Shinfield, Reading RG2 9AT, United Kingdom

The phylogenetic interrelationships of 14 members of the genus *Fusobacterium* were investigated by performing a comparative analysis of the 16S rRNA sequences of these organisms. The sequence data revealed considerable intrageneric heterogeneity. The four species *Fusobacterium nucleatum* (including *F. nucleatum* subsp. *nucleatum*, *F. nucleatum* subsp. *polymorphum*, "*F. nucleatum* subsp. *fusiformis*," and "*F. nucleatum* subsp. *animalis"), *Fusobacterium alocis*, *Fusobacterium periodonticum*, and *Fusobacterium simiae*, which colonize oral cavities, exhibited high levels of sequence homology with each other and formed a distinct group within the genus. *Fusobacterium mortiferum*, *Fusobacterium varium*, and *Fusobacterium ulcerans* also formed a phylogenetically coherent group, as did the two species *Fusobacterium gonidiaformans* and *Fusobacterium necrophorum*. *Fusobacterium russii* and *Fusobacterium necrogenses* displayed no specific relationship with any of the other fusobacteria. The sequence data are discussed in the context of previous physiological and chemical findings.

The genus *Fusobacterium* consists of obligately anaerobic, gram-negative, nonmotile, rod-shaped organisms that produce acetic and butyric acids as major end products of metabolism (14) and contain straight-chain saturated and monounsaturated long-chain fatty acids (2). Members of this genus have DNA base compositions ranging from 26 to 52 mol% G+C, although the G+C contents of the type strain, *F. nucleatum* ATCC 25557T, *F. nucleatum* ATCC 10953T, *F. nucleatum* subsp. *animalis"), *Fusobacterium alocis*, *Fusobacterium periodonticum*, and *Fusobacterium simiae*, which colonize oral cavities, exhibited high levels of sequence homology with each other and formed a distinct group within the genus. *Fusobacterium mortiferum*, *Fusobacterium varium*, and *Fusobacterium ulcerans* also formed a phylogenetically coherent group, as did the two species *Fusobacterium gonidiaformans* and *Fusobacterium necrophorum*. *Fusobacterium russii* and *Fusobacterium necrogenses* displayed no specific relationship with any of the other fusobacteria. The sequence data are discussed in the context of previous physiological and chemical findings.

The phylogenetic interrelationships of 14 members of the genus *Fusobacterium* were investigated by performing a comparative analysis of the 16S rRNA sequences of these organisms. The sequence data revealed considerable intrageneric heterogeneity. The four species *Fusobacterium nucleatum* (including *F. nucleatum* subsp. *nucleatum*, *F. nucleatum* subsp. *polymorphum*, "*F. nucleatum* subsp. *fusiformis*," and "*F. nucleatum* subsp. *animalis"), *Fusobacterium alocis*, *Fusobacterium periodonticum*, and *Fusobacterium simiae*, which colonize oral cavities, exhibited high levels of sequence homology with each other and formed a distinct group within the genus. *Fusobacterium mortiferum*, *Fusobacterium varium*, and *Fusobacterium ulcerans* also formed a phylogenetically coherent group, as did the two species *Fusobacterium gonidiaformans* and *Fusobacterium necrophorum*. *Fusobacterium russii* and *Fusobacterium necrogenses* displayed no specific relationship with any of the other fusobacteria. The sequence data are discussed in the context of previous physiological and chemical findings.

**MATERIALS AND METHODS**

**Cultures and identification.** The following strains were used in this study: *Fusobacterium nucleatum* subsp. *nucleatum* ATCC 25586 (T = type strain), "*F. nucleatum* subsp. *animalis" NCTC 12276, *F. nucleatum* subsp. *polymorphum* ATCC 10953, 'F. nucleatum subsp. *fusiformis*" NCTC 11326, *Fusobacterium gonidiaformans* ATCC 25563, *Fusobacterium alocis* ATCC 35689, *Fusobacterium mortiferum* ATCC 25557, *Fusobacterium necrophorum* ATCC 25286, *Fusobacterium necrogenses* ATCC 25556, *Fusobacterium ulcerans* NCTC 12111, *Fusobacterium periodonticum* ATCC 33693, *Fusobacterium russii* ATCC 25533, *Fusobacterium simiae* ATCC 33568, and *Fusobacterium varium* NCTC 10560. All strains were cultured and maintained by weekly subculturating on 2.5% (vol/vol) blood agar plates as described previously (14). The identity of each strain was confirmed by using the RapID ANA test system (API System, La Balme-Les-Grottes, France) according to the manufacturer's instructions. Bulk cells were obtained by growing the organisms in liquid medium BM as described previously (7).

**rRNA extraction and determination of rRNA sequence.** Cellular RNA was extracted from approximately 2 g (wet weight) of cells as described by Embley et al. (3). 16S rRNA sequences were determined directly from the total RNA by using the Sanger dideoxy termination method (18) and avian myeloblastosis virus reverse transcriptase. The sequences of primers and their target sites were as described by Lane et al. (13) and Embley et al. (3) except that we used an oligomer with the sequence TCTACGCATTTCA at positions 704 to 719 (Escherichia coli numbering system [1]). The products of the sequencing reactions were separated on 55-cm wedge-shaped (0.2- to 0.6-mm) 6% (wt/vol) polyacrylamide denaturing (7 M/L urea) gels at 55°C by using an LKB model Macrophor 2010 sequencing unit operated at 50 W per gel.

**Analysis of sequence data.** The sequences were aligned, and homology values were determined by using the Beckman Microgenie program (17). Nucleotide substitution rates \( K_{ace} \) were calculated (11), and an unrooted phylogenetic tree or network was produced by using the algorithm of Fitch and Margoliash (5) contained in a program written by Felsenstein (4) (PHYLYP version 3.1) for IBM PC computers.

**Nucleotide sequence accession numbers.** The sequences of the microorganisms which were investigated are available for electronic retrieval from GenBank under accession numbers X55401 (for *F. nucleatum* subsp. *nucleatum* ATCC 25586), X55404 (for "*F. nucleatum* subsp. *animalis" NCTC 12276), X55402 (for *F. nucleatum* subsp. *polymorphum* ATCC 10953), X55403 (for "*F. nucleatum* subsp. *fusiformis" NCTC 11326), X55410 (for *F. varium* ATCC 25563), X55406 (for *F. alocis* ATCC 35689), and X55414 (for *F. necrogenses* ATCC 25557). The sequences of all the microorganisms which were investigated are available for electronic retrieval from GenBank under accession numbers X55401 (for *F. nucleatum* subsp. *nucleatum* ATCC 25586), X55404 (for "*F. nucleatum* subsp. *animalis" NCTC 12276), X55402 (for *F. necrophorum* ATCC 10953), X55403 (for "*F. nucleatum* subsp. *fusiformis" NCTC 11326), X55410 (for *F. varium* ATCC 25563), X55406 (for *F. alocis* ATCC 35689), and X55414 (for *F. necrogenses* ATCC 25557).

* Corresponding author.
FIG. 1. Partial 16S rRNA sequences of *F. nucleatum* subsp. *nucleatum* ATCC 25586T (F.n.), *F. nucleatum* subsp. *polymorphum* ATCC 10953T (F.n.p.), “*F. nucleatum* subsp. *fulvisforme*” NCTC 11326T (F.n.f.), “*F. nucleatum* subsp. *animals*” NCTC 12276T (F.n.), *F. periodonticum* ATCC 33963T (F.p.), *F. alocis* ATCC 25286= (F.a.), *F. varium* NCTC 10560* (F.v.), *E. coli* NCTC 1211T (F.e.), and *E. morrisi* ATCC 25557T (F.mo.). The sequences are shown aligned with the sequence of *E. coli* (E.co.) (1). The first and last nucleotides in the sequences correspond to positions 100 and 1,450 of the *E. coli* sequence. N, undetermined nucleotide. Spaces in the sequences correspond to alignment gaps.
FIG. 1—Continued.
FIG. 1—Continued.
1170
E. co. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. np. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. nf. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. na. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. pe. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ri. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ni. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ra. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ru. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ng. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. go. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. nh. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. va. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ul. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. mo. AGCCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC

1320
E. co. GACCGCGGCA CAGCGGGCGG ACCCAUGCU AAAAGCGUCUAA GUCGCCCGA CAGCGGCUA CAGCGGCUA CAGCGGCUA CAGCGGCUA CAGCGGCUA
F. nn. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. nf. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. pe. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ri. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ni. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ra. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ng. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. go. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. nh. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. va. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. ul. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC
F. mo. AAGCGGAGGA AGGGGGGAU GACGCAGAC CUACACCGCC CGUUCGACCA CCAUUCUCA AUCGCGUAC AGAAGAAGC

1439

FIG. 1—Continued.
RESULTS AND DISCUSSION

The partial 16S rRNA primary sequences of 14 strains of the genus *Fusobacterium* were determined by reverse transcription. The sequencing strategy generated a continuous stretch of approximately 1,330 bases (ranging from position 100 to position 1,450 on the *E. coli* numbering system) for each strain. The sequences of the 14 strains examined are shown aligned with the corresponding *E. coli* sequence (1) in Fig. 1. Figure 1 shows that the sequences of all of the *Fusobacterium* species examined were characterized by a 22-base deletion ranging from position 458 to position 479 of the *E. coli* sequence (1). In order to determine the intragenic relationships of the *Fusobacterium* strains, their sequences were aligned, and similarity values were determined (Table 1). An unrooted phylogenetic tree based on derived $K_{xy}$ (evolutionary distance) values (Table 1) and drawn by using the distance matrix method of Fitch and Margoliash (5) is shown in Fig. 2. All of the *Fusobacterium* strains examined had unique rRNA sequences. Considerable variation in the number of base differences among the different fusobacteria was apparent (Table 1), which was indicative of substantial intragenic heterogeneity. The distance matrix tree procedure confirmed this and revealed the presence of several lines or species groups within the genus (*Fig. 2*). The four recently described subspecies of *F. nucleatum* (viz., *F. nucleatum* subsp. *nucleatum*, "*F. nucleatum* subsp. *anirnalis,*" "*F. nucleatum* subsp. *fusiforme,*" and *F. nucleatum* subsp. *polymorphum*) (8) exhibited relatively high levels of sequence similarity (97.3 to 98.4%). Comparable levels of similarity were found with the species *F. alocis*, *F. periodonticum*, and *F. simiae*. Particularly noteworthy was the exceedingly high level of sequence relatedness (99.5%) between the type strains of *F. nucleatum* subsp. *nucleatum* and *F. periodonticum* (Table 1). The distance matrix analysis confirmed the close affinity among the *F. nucleatum* subspecies, *F. alocis*, *F. periodonticum*, and *F. simiae* and indicated that these taxa represent a distinct group within the genus (*Fig. 2*). The species *F. mortiferum*, *F. ulcers*, and *F. variurn* also formed a separate group on the basis of levels of sequence similarity. These species exhibited ca. 98.5% sequence similarity with each other but significantly lower degrees of similarity with other *Fusobacterium* species (ca. 93 to 94%) (Table 1). The two species *F. gonidiaformans* and *F. necrophorum* also exhibited a specific (albeit loose) relationship on the basis of sequence data (96.5% similarity). By contrast, *F. necrogenes* and *F. russii* formed individual lines of descent and exhibited no specific relationship with any other member of the genus (*Fig. 2*).

Although all of the species examined exhibit relatively high levels of sequence similarity that are consistent with a single genus (ca. 92 to 99%), considerable intragenic heterogeneity is evident (*Fig. 2*). The high levels of sequence relatedness exhibited by the *F. nucleatum* subspecies, *F. alocis*, *F. periodonticum*, and *F. simiae* are in accord with the high levels of phenotypic resemblance of these organisms. All of these species produce indole in peptone-containing medium, ferment glutamate via the 2-oxoglutarate pathway, and contain a peptidoglycan based on meso-
FUSOBACTERIUM INTRAGENERIC RELATEDNESS

FIG. 2. Unrooted phylogenetic tree showing the intrageneric relationships of *Fusobacterium* strains. The tree is based on a comparison of 1,330 bases (Table 1). Bar = 0.005 *K* sub. 

lanthionine, and their growth is inhibited in the presence of bile (6, 9, 16). The high levels of rRNA sequence relatedness exhibited by these species are also in the main consistent with the results of DNA-DNA studies (6), although the low level of genomic similarity (7%) between *F. nucleatum* and *F. periodonticum* reported by Love et al. (15) is questionable; DNA-DNA hybridization studies in our laboratory in which the S1 nuclease procedure was used revealed ca. 63 to 76% DNA homology between these taxa (6). The sequence data also clearly demonstrate that *F. mortiferum*, *F. ulcerans*, and *F. varium* are genealogically closely related and represent a distinct clade within the genus. These species form a chemotaxonomically coherent group. Unlike other fusobacteria, *F. mortiferum*, *F. varium*, and *F. ulcerans* possess an alternative pathway for glutamate catabolism (viz., mesaconate) (6), contain a meso-diaminopimelic acid murine (9), and grow in the presence of 2% oxgall (16). *F. gonidiaformans*, *F. necrogenes*, and *F. necrophorum* differ from *F. mortiferum*, *F. varium*, and *F. ulcerans* in their pathway of glutamate catabolism (10). *F. gonidiaformans* resembles *F. necrogenes* more closely than *F. necrophorum* on the basis of physiological criteria (9, 16). However, our rRNA sequence data reveal results to the contrary. Thus, *F. gonidiaformans* shows a specific association with *F. necrophorum*, whereas *F. necrogenes* forms a separate clade within the genus. The results of our study have done much to clarify the intragenic relationships of members of the genus *Fusobacterium*. Studies on additional members of the genus (viz., *Fusobacterium naviforme*, *Fusobacterium sulci*) and possibly related taxa are currently under way in an attempt to obtain further phylogenetic insight into this complex group of anaerobes.

ACKNOWLEDGMENT

This work was supported by a research grant from the London Hospital Special Trustees (RAC 287).

REFERENCES