Bacterial biota in the human distal esophagus

Zhiheng Pei*†‡§, Edmund J. Bini†‡, Liying Yang†, Meisheng Zhou†, Fritz Francois†‡, and Martin J. Blaser†‡¶

Departments of *Pathology, †Medicine, and ¶Microbiology, New York University School of Medicine, New York, NY 10016; and ‡Department of Veterans Affairs New York Harbor Healthcare System, New York, NY 10010

The esophagus, like other luminal organs of the digestive system, provides a potential environment for bacterial colonization, but little is known about the presence of a bacterial biota or its nature. By using broad-range 16S rDNA PCR, biopsies were examined from the normal esophagus of four human adults. The 900 PCR products cloned represented 833 unique sequences belonging to 41 genera, or 95 species-level operational taxonomic units (SLOTU); 59 SLOTU were homologous with culture-defined bacterial species, 34 with 16S rDNA clones, and two were not homologous with any known bacterial 16S rDNA. Members of six phyla, *Firmicutes*, *Bacteroides*, *Actinobacteria*, *Proteobacteria*, *Fusobacteria*, and *TM7*, were represented. A large majority of clones belong to 13 of the 41 genera (783/900, 87%), or 14 SLOTU (574/900, 64%) that were shared by all four persons. *Streptococcus* (39%), *Prevotella* (17%), and *Veilonella* (14%) were most prevalent. The present study identified ≈56–79% of SLOTU in this bacterial ecosystem. Most SLOTU of esophageal biota are similar or identical to residents of the upstream oral biota, but the major distinction is that a large majority (82%) of the esophageal bacteria are known and cultivable. These findings provide evidence for a complex but conserved bacterial population in the normal distal esophagus.

Colonizing bacteria are integral components in all parts of the human digestive tract, from the oral cavity to the anus, but little is known about the esophagus. A bacterial population is essential for the development of the gastrointestinal mucosal immune system, the maintenance of a normal physiological environment, and for providing essential nutrients (1). Bacteria also play a role in a variety of disease conditions, as exemplified by *Helicobacter pylori* in the stomach (2). Conversely, loss of normal biota is responsible for overgrowth of opportunistic pathogens that normally are inhibited, such as occurs in antibiotic-associated colitis (3). Microenvironment alterations may favor overgrowth of bacteria that produce carcinogenic metabolites (4, 5), promoting tumorigenesis in inflammation-induced cancers, such as adenocarcinoma in mouse experimental colitis models (6).

Quantitative cultivation-based studies indicate that obligate anaerobes constitute an important part of the residential oral biota. In the subgingival crevice of healthy adults, total microscopic bacterial counts averaged 2.7×10^{11} cells per gram of wet weight (7), of which anaerobic bacteria represented approximately two-thirds. However, direct amplification of 16S rDNA revealed much more diverse bacterial populations than did cultivation (8, 9). The majority (52.5%) of directly amplified 16S rDNA sequences were <99% identical to sequences within public databases in contrast to 21.4% of sequences from cultivated bacteria. Of the 16S rDNA sequences identified by direct amplification, 13.5% had <95% identity with known sequences, compared with none from cultivated organisms (8). Such findings suggest that previously unidentified higher-order taxonomic divisions remain within the human ecosystem. The influence of this complex bacterial biota on the human esophagus remains to be determined.

H. pylori appear to be indigenous to the human stomach (10, 11) but have been disappearing in consequence to diminished transmission and to antibiotic usage (12). Because the acidic environment of the stomach is hostile for most bacteria except for *Helicobacter* species, other gastric biota become apparent only in patients with reduced acidity, such as accompanies progressive atrophic gastritis (13). In healthy fasting patients, large numbers of *Enterococcus*, *Pseudomonas*, *Streptococcus*, *Staphylococcus*, and *Stomatococcus* may be isolated in culture when acidity is physiologically reduced, as occurs at night (14). Gastric bacteria are potentially introduced to the distal esophagus by reflux.

The esophagus, as with other luminal organs of the digestive system, represents a potential environment for bacterial colonization, because it has a large mucosal surface downstream of the bacterial species-rich oropharynx. Because of the lack of absolute anatomic or known physiological barriers, bacteria could be introduced into the esophagus by swallowing or by reflux from a colonized stomach. Lack of knowledge about bacterial populations in the esophagus appears related to the belief that bacteria are not responsible for esophageal diseases. Distal esophagitis is generally considered to be a consequence of reflux of gastric acidity and is known as gastroesophageal reflux disease (15). Studies attempting to culture bacteria from luminal washes of the esophagus have reported poor yields; the washes, most likely reflecting transient bacteria of oropharyngeal origin, were either sterile or contained an average of 16 colony-forming units/ml, with no common species found (16, 17). These unsuccessful attempts in detecting bacteria in esophageal washes suggest that the passage of oropharyngeal bacteria is rapid, and/or that the bacteria present in the washes were not cultivatable. Characterization of bacterial populations of the esophagus is not described in textbooks of gastroenterology, microbiology, or infectious diseases. Although the esophagus may become infected with *Candida*, Herpesvirus, *Cryptococcus*, or *Histoplasma*, with the exception of *Mycobacterium* species, bacterial etiologies for the inflammation involving the distal esophagus have not been explored.

The likelihood of an indigenous esophageal biota is increased by the emerging knowledge of the extent of bacterial colonization of human mucosal surfaces (3, 8, 9, 18), as well as the ability of bacteria to colonize diverse environments including hot springs, volcanoes, and deep-sea vents (19, 20) but has not been sufficiently addressed. We postulated that indigenous bacteria are closely associated with the esophageal mucosa, are not removable by simple washes, and that the majority may not be cultivatable. To examine these hypotheses, we used sequencing-based broad-range16S rDNA PCR to amplify bacteria in biopsies of normal esophagus.

Materials and Methods

Subjects. Patients presenting to the Department of Veterans Affairs Medical Center and Bellevue Hospital, New York, with gastrointestinal symptoms requiring upper gastrointestinal endoscopy were eligible for this study. Those who were willing to

---

participate in studies of upper gastrointestinal microbiology and who signed an informed consent form were recruited for this study. Exclusion criteria included the use of antibiotics in the prior 8 weeks, previous gastric/esophageal surgery, active infection of the oral cavity, and HIV infection. Esophagogastroduodenoscopy was performed, and endoscopic findings recorded for patients who met the above criteria. Esophageal biopsies were obtained 2 cm above the squamocolumnar junction. Each biopsy was examined microscopically. Patients with morphological features of gastroesophageal reflux disease and intestinal/gastric metaplasia (Barrett's esophagus) or with other esophageal pathology were excluded from this aspect of the study. In total, four patients who had normal esophageal histology and met the above criteria were included; all were Hispanic American (three males and one female), ranging in age from 49 to 79 years (see Table 2, which is published as supporting information on the PNAS web site). Tissue sections of esophageal biopsies were examined for the presence of bacteria by microscopy by using the Gram-Twort stain (21). All staining solutions were passed through 0.22-μm filters.

Specimen Processing. Biopsies of ≈1 × 2 × 2 mm were placed in a 1.5-ml screw-top test tube and stored at –70°C. The four specimens were coded as A, B, C, and D, so that the persons who performed subsequent studies were blinded to the pertinent clinical information. DNA was extracted from the biopsy in a PCR-free clean-room by using a tissue DNA extraction kit (Qiagen, Chatsworth, CA). To lyse bacteria present, the biopsy was incubated with lysozyme (20 mg/ml) (Sigma) in 180 μl of buffer containing 20 mM Tris·HCl (pH 8.0), 2 mM EDTA, and 1.2% Triton X-100 for 60 min at 37°C, and then the Qiagen tissue DNA extraction protocol was followed. The DNA-enriched fractions were eluted in 200 μl of H_2O.

16S rDNA Clone Libraries. For each biopsy, 10 libraries were created from independent PCR amplifications of the extracted DNA. For each PCR, 5 μl of the extracted DNA was added to 45 μl of PCR mixture containing 5 μl of 10× PCR buffer (Qiagen), 1.5 mM MgCl$_2$, 200 μM each dNTP, 50 pmol of each primer, and 5 units of *Taq* DNA polymerase. Reactions were run at 94°C for 2 min, followed by 30 cycles of amplification at 94°C for 30 sec, 52°C for 30 sec, and 72°C for 90 sec and a 20-min extension at 72°C. Two primer pairs were used. In primer pair 1 (forward, GGIACTGAGACACIGICCIIACTCCT; and reverse, TTCCIITACIGITACCTTGTTACGACTT), inosine (I) was used at positions of nucleotide ambiguity, because it forms stable base pairs with all four usual bases. Use of inosine-containing primers significantly reduces the complexity accompanying use of conventional degenerate primers (22). As such, both inosine-containing primers perfectly match the consensus sequence-derived 16S rDNA pools composed of 21 well diversified eubacterial groups including *Agrobacterium*, *Aquifex*, *Arthrobacter*, *Bacillus*, *Chlamydia*, *Chlorobium*, *Chloroflexus*, *Chloroplast*, *Clostridium*, *Desulfovibrio*, *Escherichia*, *Flavobacterium*, *Flexibacter*, *Gloeobacter*, *Heliobacterium*, *Leptonema*, *Planctomyces*, *Rhodocyclus*, *Synechococcus*, *Thermotoga*, and *Thermus* (19) but do not have significant 3′ homology with human 18S rDNA and human mitochondrial small subunit rDNA sequences. The expected product is ≈1,200 bp, corresponding to positions 318–1,519 of the *Escherichia coli* 16S rRNA gene. Primer pair 2 (forward, AGAGTTTGATYMTGGCTCAG; and reverse, TACGGYTACCTTGTTACGACTT) produced PCR products of ≈1,500 bp, spanning positions 8–1513 of the *E. coli* 16S rRNA gene. The PCR products were separated from free PCR primers by using a PCR purification kit (Qiagen), ligated with the pGEM T Easy (Promega) cloning vector, and used to transform *E. coli* DH5α competent cells. The cloned inserts underwent sequence

Fig. 1. Microscopic examination of bacterial cells in the esophagus. Esophageal biopsies were fixed in formalin, paraffin-embedded, sectioned, and examined by using Gram-Twort stain. In the biopsy from subject A, monomorphic Gram-negative bacilli were tightly associated with the surface of squamous epithelial cells (×1,000).

analysis using forward PCR primers and using both forward and reverse primers for clones with ambiguous phylogenetic status.

Phylogenetic Analysis. The sequences were analyzed by using the preview version of SEQUENCE MATCH at Ribosomal Database Project II (RDP II, http://rdp.cme.msu.edu) (23). All sequences were examined for chimerism by using CHIMERA DETECTION at RDP II and BELLEROPHON (24). For phylogenetic analysis, representative 16S rDNA sequences were aligned by using SEQUENCE ALIGNER at RDP II. Misaligned positions were corrected by using ARB (http://rtfm.arb-home.de). Phylogram of the nucleotide alignment was generated by using the PAUP 4.0B10 [PAUP 4.0B2, Phylogenetic Analysis Using Parsimony (* and Other Methods) Ver. 4] (Sinauer Associates, Sunderland, MA) neighbor-joining method, based on HKY85 distance matrices (25). All sequences that are not classifiable by using the current 16S database at RDP II were deposited in GenBank.

Estimation of SLOTU Richness. The total number of SLOTU that may be present in human distal esophagus and its associated confidence interval were calculated by using a nonparametric richness estimator, Chao 1, as described (26).

Results

Microscopic Examination of Bacterial Biota in the Distal Esophagus. Histological examination of the four biopsies revealed no significant pathological changes. Based on the analogy that chronic gastritis usually is a consequence of the presence of overlaying *H. pylori* in the lumen (27), we examined the distal esophagus to determine whether bacterial cells might be visible. Such a study, if positive, can provide morphological evidence for an indigenous esophageal biota. Bacteria were observed in one of the four biopsies. Bacteria appeared to be closely associated with the epithelial cell surface (Fig. 1). All were Gram-negative bacilli with uniform morphology.

Elimination of Contaminating Sequences. To examine the nature of the bacterial populations present in the distal esophagus and to define their ancestry, we performed universal bacterial 16S PCR on biopsies from four subjects. To avoid operational bias, each biopsy was sampled 10 times by preparation of individual libraries from independent PCR amplifications. On average, each of the 10 libraries yielded ≈700 clones, of which ≈25–30 clones were randomly picked and sequenced; thus, in total, 250–300 clones were analyzed from each biopsy. Because reagents used in DNA extraction and PCRs may contain bacteria or their genomic DNA, and under certain experimental conditions these contaminating DNA may become detectable after PCR amplification, we performed two kinds of controls. The PCR reagent control included all PCR reagents except for the template DNA and was cycled through the identical program as for the esophageal DNA. This control did not generate any visible signals on electrophoresis and ethidium bromide staining. Blind excision of agarose gel at the expected location of the

signal did not yield any 16S rDNA-containing clones on ligation and transformation. The DNA extraction control used water to replace the esophageal biopsy during the extraction. The resultant DNA was amplified, cloned, and sequenced. Visible PCR products, albeit weak, were observed when the DNA from the water controls was amplified by either of the two pairs of primers. Transformation yielded ≈300 clones, of which 72 were picked and sequenced. Eleven SLOTU were found, including *Pseudomonas tolaasii*, *Pseudomonas influorensces*, *Pseudomonas syringae*, *Pseudomonas putida*, uncultured *Duganella* clone CTHB-18 (AF067655), *Stenotrophomonas maltophilia*, *Janthinobacterium lividum*, *Lactobacillus paracasei*, *Propionibacterium acnes*, *Pseudomonas antarctica/meridiana,* and *Brevundimonas bulata* (Table 3, which is published as supporting information on the PNAS web site). These species were excluded when detected in esophageal biopsies.

Yield from 16S PCR Studies. DNA extracted from human case A was amplified by using the inosine-containing primer pair 1. The forward primer was used for sequencing, yielding a readable sequence of >700 bp, ≈650 bp of which was used for phylogenetic analysis. About 250 clones of the ≈7,000 colonies were picked and sequenced, which yielded 186 16S rDNA sequences after eliminating failed sequences and subtraction of the 11 species present in the contamination controls. To evaluate primer-mediated amplification bias and consistency of the bacterial community in the normal distal esophagus, the second pair of primers that do not contain inosine was used to amplify DNA extracted from three additional cases (B, C, and D). The forward primer was used for sequencing, yielding a readable sequence of >900 bp, ≈ 850 bp of which was used for phylogenetic analysis. About 250–300 clones from each specimen were picked and sequenced, of which 205, 264, and 245 16S rDNA sequences for cases B, C, and D, respectively, were obtained after exclusion of failed sequences and subtraction of contamination. Both primer pairs generated similar species distributions from the specimens with which they were used for amplification (Tables 4 and 5, which are published as supporting information on the PNAS web site).

Definition of SLOTU. RDP II SEQUENCE MATCH (Preview Version) compares the inquiry sequence to known 16S rDNA sequences for assignment to the closest taxon, based on similarity scores. In the literature, 97% sequence homology of 16S rDNA has been defined as the species boundary (28); however, its mathematical relation with the RDP II similarity score had not been determined. To define a species boundary based on similarity scores, we analyzed all 900 sequences from the four specimens using linear regression (Fig. 2). Each of the 900 sequences was paired with the best-matched sequences determined by use of SEQUENCE MATCH. The similarity score and percentage of homology for each pair of sequences were calculated. The two measurements were related by the regression equation: $y = 0.03864x - 2.8756$ ($r = 0.93$, $P < 0.0001$), where y indicates the similarity score and x, the percentage of homology. Based on the regression formula, a similarity score of 0.8725 was equivalent to the species boundary, defined by 97% 16S sequence homology. All sequences with similarity scores <0.873 had <97% homology, and those >0.873 (RDP II score is given with three digits after the decimal point) had homology ≥97%, without exception. Based on this validation, we termed sequences defined by this method SLOTU.

Classification of the Clones. We first used the program CHIMERA DETECTION at RDP II to detect potential chimeric sequences that may be derivatives of parental sequences from the RDP II 16S rDNA database. None of the 900 sequences exhibited typical features for a chimeric sequence. Because the sensitivity of CHIMERA DETECTION is highly dependent on the completeness of the database that currently contains only a very small pool of all
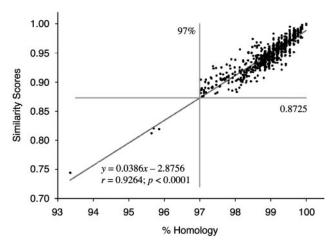

Fig. 2. Correlation of sequence homology with similarity scores. The 900 sequences obtained from this study each were subjected to pairwise comparison with their most closely matched 16S rDNA sequence from the RDP II database. Sequence homology (%) and similarity score using SEQUENCE MATCH at RDP II were calculated for each pair. The relationship between the two measurements was analyzed by linear regression. The 97% line represents the boundary of SLOTU defined by sequence homology (28). The 0.8725 line is the corresponding SLOTU boundary defined by similarity scores, as determined by linear regression in this study.

bacterial 16S sequences, we next examined all sequences for the possibility of chimerism that may be derived from parental sequences in our PCR clonal libraries but not yet available for analysis in the public database, using the program BELLEROPHON (24). BELLEROPHON was specifically developed to detect 16S rDNA chimeras in PCR-clone libraries. Again, none of the 900 sequences were classified as being chimeric. Compared with the database of 16S rDNA derived from cultivated type strains, 738 of the 900 clones had a similarity score >0.873 with type strains. Of the remaining 162 clones, one had a similarity score >0.873 with 16S rDNA derived from a fully defined cultured bacterial nontype strain (Table 5). The 739 clones (82.1%) from these two groups were classified as belonging to culture-defined bacterial species. A total of 59 species were represented in this category (Fig. 3 and Table 5). Of the remaining 161 clones (17.9%), 157 had a similarity score >0.873 only in relation to PCR-generated 16S rDNA clones in the database, which were classified as 16S rDNA clones, representing 34 SLOTU. A binomial name was assigned to each putative species based on the genus name assigned by SEQUENCE MATCH, followed by the GenBank accession number of the best-matched rDNA clone. The remaining four sequences were not significantly homologous (<0.873 similarity score) with culture-defined species nor with existing 16S rDNA clones. These sequences were classified as unknowns, representing two SLOTU (GenBank accession nos: AY394004, AY423746, AY423747, and AY423748). Their designation was based on the putative genus or higher taxon they most closely resembled, followed by the clone number used in this study. The taxonomic assignments were confirmed by phylogenetic analysis (Fig. 3).

Distribution of the Clones at the Phylum and Genus Levels. The 900 clones conformed to six phyla. *Firmicutes* was the most prevalent phylum represented in the distal esophageal biopsies, accounting for 41 SLOTU or 69.6% (626/900) of clones from the four specimens (Fig. 3 and Table 1). *Bacteroides* was the second most prevalent phylum, composed of 23 SLOTU or 20.2% (182) of the clones. *Actinobacteria* (39, 4.3%, eight SLOTU), *Proteobacteria* (20, 2.2%, 14 SLOTU), *Fusobacteria* (20, 2.2%, six SLOTU), and *TM7* (13, 1.4%, three SLOTU), in combination, comprised the remaining 10% of the clones. The six phyla were observed in all four

	A	B	C	D	Total
Streptococcus mitis	1	47	6	115	169
Streptococcus oralis	18	0	0	7	25
Streptococcus infantis	0	0	6	2	8
Streptococcus cristatus	0	0	0	1	1
Streptococcus gordonii	0	1	1	0	2
Streptococcus parasanguinis	24	4	29	10	67
Streptococcus thermophilus	34	14	42	10	100
Streptococcus anginosus	0	0	0	1	1
Streptococcus intermedius	0	0	1	0	1
Granulicatella adiacens	7	1	2	5	15
Granulicatella elegans	0	2	0	0	2
Lactobacillus salivarius	0	0	4	0	4
Lactobacillus delbrueckii	0	0	8	0	8
Lactobacillus gasseri	0	1	10	0	11
Lactobacillus fermentum	0	0	2	0	2
Lactobacillus frumenti	0	0	1	0	1
Gemella morbillorum	1	0	0	0	1
Gemella haemolysans	2	1	0	6	9
Gemella sanguinis	3	0	1	0	4
Bulleidia moorei	1	0	2	1	5
Paenibacillus AY167820	0	0	1	0	1
Eubacterium sulsi	0	1	0	1	2
Mogibacterium vescum	0	0	1	0	1
Filifactor alocis	0	0	0	1	1
Peptostreptococcus AF287763	0	0	0	2	2
Veillonella atypica	8	22	21	15	66
Veillonella dispar	26	15	18	3	62
Veillonella 59–8–23(95%)	0	0	3	0	3
Megasphaera micronuciformis	3	8	7	1	19
selenomonas 9–1–7(93%)	0	1	0	0	1
Selenomonas AF287795	0	1	0	0	1
Selenomonas sputigena	0	1	0	0	1
Selenomonas AF385514	0	1	0	1	2
Centipeda periodontii	0	1	0	0	1
Clostridium AF385563	0	1	0	0	1
Clostridiales AF287770	1	3	4	0	8
Clostridiales AF385510	2	0	0	0	2
Clostridium AY278618	1	1	7	1	10
Clostridiales AF385508	0	1	0	0	1
Clostridiales AF201981	0	0	2	2	4
Clostridiales AF287772	0	0	0	1	1
Atopobium parvulum	0	0	3	1	4
Atopobium rimae	0	0	1	0	1
Rothia AF543276	0	0	0	2	2
Rothia mucilaginosus	1	2	14	3	20
Actinomyces graevenitzii	1	0	1	0	2
Actinomyces AF385521	0	0	4	0	4
Actinomyces meyeri	0	0	1	0	1
Actinomyces odontolyticus	1	1	1	2	5
Comamonas testosteroni	0	1	0	0	1
Delftia acidovorans	0	1	0	1	2
Achromobacter piechaudii	0	1	0	0	1
Neisseria AY093465	0	0	0	1	1
Neisseria pharyngis	1	0	0	0	1
Neisseria subflava	1	0	0	0	1
Shewanella baltica	0	0	0	1	1
Yersinia kristensenii	0	0	0	1	1
Haemophilus segnis	0	0	0	1	1
Haemophilus parainfluenzae	1	0	0	0	1
Haemophilus AB098612	0	0	2	0	2
Haemophilus paraphrophilus	1	0	0	2	3
Campylobacter concisus	0	0	3	0	3
Campylobacter showae	1	0	0	0	1
TM7 AF385500	0	1	0	0	1
TM7 AF385520	4	1	5	1	11
TM7 AF385506	0	0	0	1	1
Bacteroidetes AJ289174	0	0	4	0	4
Bacteroides AF385513	1	6	1	3	11
Prevotella AY207061	0	0	0	3	3
Prevotella AY093466	0	0	0	1	1
Prevotella AY005064	0	0	1	0	1
Prevotella AY005065	0	0	1	0	1
Prevotella intermedia	0	1	0	3	4
Prevotella pallens	5	2	2	5	14
Prevotella nigrescens	0	0	0	1	1
Prevotella AJ289182	0	0	1	0	1
Prevotella oris	0	0	2	0	2
Prevotella AF385512	1	2	13	0	16
Prevotella AJ012605	0	0	0	1	1
Prevotella denticola	0	0	2	0	2
Prevotella AF366269	0	11	2	2	15
Prevotella AF385551	0	4	13	0	17
Prevotella melaninogenica	0	5	0	9	14
Prevotella AF432133	1	17	1	0	19
Prevotella veroralis	23	13	9	0	45
Tannerella forsythensis	1	0	0	0	1
Porphyromonas AY008310	0	0	0	1	1
Porphyromonas endodontalis	3	0	0	3	6
Porphyromonas gingivalis	0	0	0	2	2
Leptotrichia AF278621	0	2	0	1	3
Leptotrichia AF385571	0	5	0	2	7
Fusobacterium AF287806	0	0	0	1	1
Fusobacterium simiae	1	0	0	0	1
Fusobacterium naviforme	2	0	0	0	2
Fusobacterium nucleatum	3	0	0	1	6
Total sequences	186	205	264	245	900
Total species	35	40	45	51	95

Phyla groupings (right): Firmicutes, Actinobacteria, Proteobacteria, TM7, Bacteroidetes, Fusobacteria
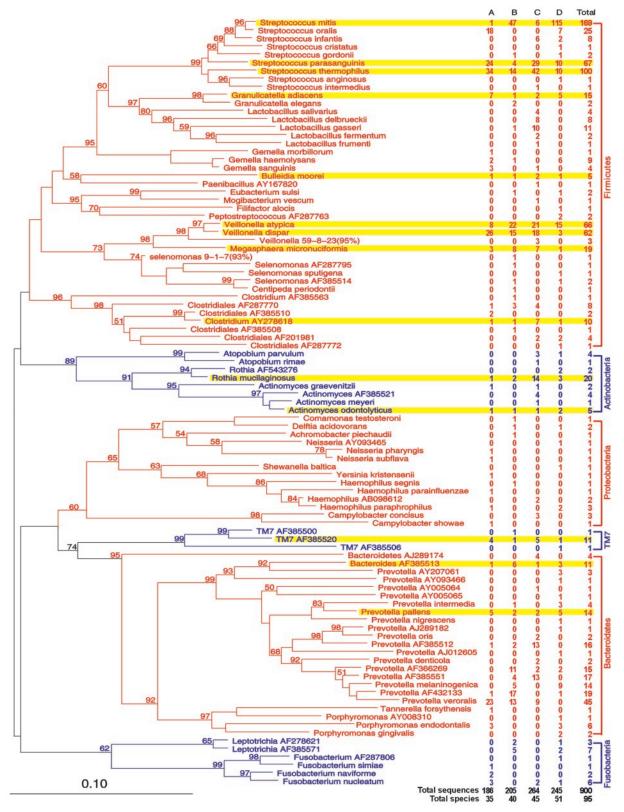
Scale bar: 0.10

Fig. 3. Phylogenetic analysis of bacterial 16S rDNA detected in biopsies of the normal distal esophagus from four persons. Sequences were aligned by using SEQUENCE ALIGNER at RDP II. Misaligned positions were corrected by using ARB. Phylogram was generated by using PAUP 4.0b10 neighbor-joining analysis, based on HKY85 distance matrices. Bootstrap values (based on 500 replicates) are represented at each node when >50%, and the branch length index is represented below the phylogram. Names of SLOTU are located at the termination of each branch. 16S rDNA clones are potential bacterial species whose phylogenetic positions were designated by PCR-amplified 16S sequences only, represented by the closest genus, followed by the GenBank accession number of the best-matched sequence. Unknowns are represented by the closest taxon followed by the serial number of the clone used in this study, as well as the percent sequence identity (in parentheses). The frequency at which a species was detected and its sources are indicated (on the right). The 95 SLOTU belonging to six phyla, contrasted by alternating red and blue print, are shown (on the right). SLOTU shared by all four persons are highlighted in yellow.

Table 1. Representation of bacterial phyla in the distal esophagus and subgingival crevice

	Number (%) of clones					
	Distal esophagus					Subgingival crevice*
Phylum	A (n = 186)	B (n = 205)	C (n = 264)	D (n = 245)	Total (mean ± SD) (n = 900)	(n = 2252)
Firmicutes	132 (70.9)	128 (62.4)	178 (67.4)	188 (76.7)	626 (69.5 ± 6.0)	659 (26.1)
Bacteroidetes	35 (18.8)	61 (29.8)	49 (18.6)	37 (15.1)	182 (20.2 ± 6.4)	234 (9.3)
Actinobacteria	3 (1.6)	3 (1.5)	25 (9.5)	8 (3.3)	39 (4.3 ± 3.8)	275 (11.0)
Proteobacteria	6 (3.2)	4 (2.0)	5 (1.9)	5 (2.0)	20 (2.2 ± 0.6)	338 (13.4)
Fusobacteria	6 (3.2)	7 (3.4)	2 (0.8)	5 (2.0)	20 (2.2 ± 1.2)	353 (14.0)
TM7	4 (2.1)	2 (1.0)	5 (1.9)	2 (0.8)	13 (1.4 ± 0.6)	34 (1.3)
Spirochaetes	0	0	0	0	0	537 (21.3)
Deferribacteres	0	0	0	0	0	86 (3.4)
Obsidian pool OB11	0	0	0	0	0	6 (0.0)

*From Paster *et al.* (9).

specimens with similar distributions (Table 1). In total, the distal esophageal biopsies from the four individuals yielded sequences representing 41 genera (Table 4). These included one cluster of four clones assigned to unclassified *Bacteroidetes* (order) and five clusters of 16 clones assigned to unclassified *Clostridiales* (order). Phylogenetic status of these clones was analyzed by using near full length (>1,400 bp) sequences. The five clusters of unclassified *Clostridiales* each were assigned a genus status because the clusters differ between each other by >10%.

Members of 13 genera were observed in all four specimens, including *Streptococcus* (38.6%), *Prevotella* (17.4%), *Veillonella* (14.2%), *Rothia* (2.4%), *Megasphaera* (2.1%), *Granulicatella* (1.9%), *Gemella* (1.6%), *TM7* (1.4%), *Actinomyces* (1.3%), *Bacteroides* (1.2%), *Clostridium* (1.2%), *Haemophilus* (0.8%), and *Bulleidia* (0.6%) (Fig. 3 and Table 4). In total, members of the 13 common genera comprised 783 (87.8%) of the 900 esophageal clones analyzed. Members of the other 30 genera were detected in some but not all four specimens, accounting for the remaining 110 (12.2%) of the 900 clones analyzed (Table 4).

Distribution of the Clones at the Species Level. In total, 59 species and 36 putative species, representing 95 SLOTU, were detected in the four specimens. Fourteen SLOTU were found in all four specimens, accounting for the majority (63.8%, 574/900) of the clones analyzed, representing four phyla (Fig. 2). These SLOTU included *Streptococcus mitis* (169), *Streptococcus thermophilus* (100), *Streptococcus parasanguis* (67), *Veillonella atypica* (66), *Veillonella dispar* (62), *Rothia mucilaginosus* (20), *Megasphaera micronuciformis* (19), *Granulicatella adiacens* (15), *Prevotella pallens* (14), *Bacteroides* AF385513 (11), *TM7 AF385520* (11), *Clostridium AY278618* (10), *Bulleidia moorei* (5), and *Actinomyces odontolyticus* (5). Variation also was observed, because, for example, *Prevotella veroralis*, the sixth most abundant SLOTU detected (5.0%), was not detected in case D (Fig. 2 and Table 5).

Estimation of SLOTU Richness. Use of the Chao 1 estimator (26) projects that the normal distal esophageal bacterial biota contains ≈139 SLOTU [95% confidence interval (CI), 121–169]. Based on this prediction, the present study has identified ≈68.3% (95% CI, 56.2–78.5%) of SLOTU in this bacterial ecosystem (Fig. 4, which is published as supporting information on the PNAS web site).

Discussion

Previous culture-based studies suggested that the esophagus is either sterile or contains only few transient bacteria (16, 17). The specimens used in those studies were washes of the luminal esophageal contents, which might explain the failure to detect the presence or richness of the bacterial biota, because washes may not yield bacteria closely associated with the mucosa. Similarly, washing is not a good method to detect transient oral bacteria because the esophagus, unlike the oral cavity, stomach, and colon, does not retain food contents. The cultivation methods used in those studies also may explain the failure to detect bacteria, because many might be fastidious or exist in an uncultivable state. We attempted to overcome these drawbacks by using cultivation-independent PCR to amplify bacteria associated with mucosal surfaces in tissue biopsies.

In situ staining revealed association of bacteria with the epithelial cell surfaces (Fig. 1), suggesting the presence of residential bacteria in the distal esophagus. For the colon, using fluorescent *in situ* hybridization, Swidsinski *at al*. (29) demonstrated clusters of 10–100 bacteria on the mucosal surfaces in 84% of processed biopsies when bacterial concentrations were 10^6 to 10^7/ml, similar to our observation in the esophagus. However, in the colon, the presence of bacteria on mucosal surfaces could not be proven when concentrations are $<10^6$/ml. Likewise, the failure to visualize bacteria in three of the four cases and the lack of morphologic heterogeneity in the one positive biopsy compared with the findings from 16S PCR amplification of DNA extracted from fresh-frozen tissue could be due to the removal of bacteria from the biopsies by vigorous washes during tissue processing for histological examination.

Analysis of 900 16S rDNA clones revealed a bacteria-rich microbiota in the normal distal esophagus. Six phyla, 41 genus-level taxonomic units, and 95 SLOTU were represented in the microbiota. However, contrary to our hypothesis, the majority (82.1%) of 16S rDNA found in this study represent cultivatable bacterial species. The majority of these species are relatively easy to culture using the methods and conditions used in the previous studies (16, 17). Although the exact reason for the failure to isolate these species is unclear, we speculate that (*i*) our ability to demonstrate a complex bacterial biota from the tissue biopsies suggests that esophageal bacteria are closely related to mucosal surfaces, consistent with, although not conclusive for, the existence of a residential bacterial biota in the normal distal esophagus; or (*ii*) the esophageal bacteria could also exist *in vivo* in a viable but nonculturable state (30), a recently recognized phenomenon in which nonsporulating bacteria persist as dormant vegetative cells with low metabolic activity, especially when they reside in a biofilm (31). As discussed above and in other specimens not part of this study (Z.P., L.Y., R. M. Peek, S. M. Levine, D. T. Pride, and M.J.B., unpublished work), microscopic visualization of bacteria *in situ*, even after tissue preparation, supports the concept of a residential biota (Fig. 5, which is published as supporting information on the PNAS web site). Our preliminary results obtained from direct culture of esophageal biopsies further support the presence of a residential bacterial

biota. A large majority of esophageal bacteria were cultivable, and there were $\approx 10^4$ bacteria per mm^2 mucosal surface of the distal esophagus. They are predominantly α-hemolytic *Streptococcus* species (Fig. 6, which is published as supporting information on the PNAS web site), consistent with the findings from PCR amplification.

The bacterial biota observed in the normal distal esophagus was similar in all four specimens. At the species level, the majority (63.8%) of clones belong to the 14 species present in all four biopsies. Similarly, 87% of the clones belong to 12 common genera. The distributions of clones among the common species and genera also are similar between the biopsies. These data suggest that, although the esophagus is generally viewed as a conduit for food passage, the environment in which the bacteria may reside is relatively stable. Alteration of the microenvironment, such as occurs in reflux esophagitis, Barrett's esophagus, and esophageal carcinoma, could affect the bacterial biota; the *H. pylori* status of the gastric contents also could alter the upstream esophageal biota, due to effects on gastric acidity and its reflux.

The bacterial biota in the normal distal esophagus is similar to that of the oropharynx (8, 9, 32). The majority of clones, especially those that share significant homology only with uncultivated 16S rDNA clones in the RDP II database, are mostly related to oral bacteria, suggesting that transient bacteria predominated in the specimens. However, certain highly abundant oral bacteria, such as members of *Spirochaetes*, were not represented at all in the esophagus. These differences could be due to either selective passage of bacteria from the oropharynx or the selective retention of particular species of oral bacteria by the esophagus. The effects of the condition of the oral cavity on the esophageal bacterial biota remain to be determined. The major distinction from the oral bacteria is that the large majority of the esophageal bacteria belong to known families and are cultivable (8, 9, 32).

Although studies of bacterial biota in other parts of the human microbial ecosystem suggest the presence of yet-to-be-identified higher-order taxonomic divisions, our findings suggest that their relative abundance in the esophagus will be low, even when more samples are analyzed. That the esophageal biota is composed of predominantly known and cultivable species *per se* is unique, compared with other well-studied human biota. This suggests that, in contrast to the stomach, the environment in the normal distal esophagus is temperate and does not require drastic adaptation for bacterial colonization. However, when the environment changes in disease states such as reflux esophagitis, novel species may emerge, similar to the findings in the human mouth (8, 9, 32). Identifying complex bacterial populations in the distal esophagus offers new approaches to understanding bacterial roles as markers of or as pathogenetic factors in esophageal diseases, such as esophagitis of unknown etiology, Barrett's esophagus, and esophageal adenocarcinomas. We currently are examining the change of bacterial biota in diseases of the distal esophagus; the study may reveal alteration of existing normal biota or presence of novel sequences.

Bacterial populations in other parts of the digestive system, including the oral cavity and colon, play important roles in the maintenance of local physiology as well as in disease etiology (1–7, 33). A full assessment of the composition, transience, or stability of this complex bacterial biota in the distal esophagus and associations with disease remains to be determined. Nevertheless, as a result of this study, we now understand the major bacterial populations found in the healthy distal esophagus of adults; extension of those findings and comparison with disease should be fruitful.

1. Cunningham-Rundles, S., Ahrn, S., Abuav-Nussbaum, R. & Dnistrian, A. (2002) *Nutr. Rev.* **60,** S68–S72.
2. Peek, R. M., Jr., & Blaser, M. J. (2002) *Nat. Rev. Cancer* **2,** 28–37.
3. Hopkins, M. J. & Macfarlane, G. T. (2002) *J. Med. Microbiol.* **51,** 448–454.
4. Roediger, W. E., Lawson, M. J. & Radcliffe, B. C. (1990) *Dis. Colon Rectum* **33,** 1034–1036.
5. Mowat, C., Williams, C., Gillen, D., Hossack, M., Gilmour, D., Carswell, A., Wirz, A., Preston, T. & McColl, K. E. (2000) *Gastroenterology* **119,** 339–347.
6. Kado, S., Uchida, K., Funabashi, H., Iwata, S., Nagata, Y., Ando, M., Onoue, M., Matsuoka, Y., Ohwaki, M. & Morotomi, M. (2001) *Cancer Res.* **61,** 2395–2398.
7. Gordon, D. F., Stutman, M. & Loesche, W. J. (1971) *Appl. Microbiol.* **21,** 1046–1050.
8. Kroes, I., Lepp, P. W. & Relman, D. A. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 14547–14552.
9. Paster, B. J., Boches, S. K., Galvin, J. L., Ericson, R. E., Lau, C. N., Levanos, V. A., Sahasrabudhe, A. & Dewhirst, F. E. (2001) *J. Bacteriol.* **183,** 3770–3783.
10. Ghose, C., Perez-Perez, G. I., Dominguez-Bello, M. G., Pride, D. T., Bravi, C. M. & Blaser, M. J. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 15107–15111.
11. Falush, D., Wirth, T., Linz, B., Pritchard, J. K., Stephens, M., Kidd, M., Blaser, M. J., Graham, D. Y., Vacher, S., Perez-Perez, G. I., *et al.* (2003) *Science* **299,** 1582–1585.
12. Tytgat, G. (2000) *J. Gastroenterol. Hepatol.* **15 Suppl,** G30–G33.
13. Guerre, J., Vedel, G., Gaudric, M., Paul, G. & Cornuau, J. (1986) *Pathol. Biol.* **34,** 57–60.
14. Monstein, H. J., Tiveljung, A., Kraft, C. H., Borch, K. & Jonasson, J. (2000) *J. Med. Microbiol.* **49,** 817–822.
15. Freston, J. W., Malagelada, J. R., Petersen, H. & McCloy, R. F. (1995) *Eur. J. Gastroenterol. Hepatol.* **7,** 577–586.
16. Gagliardi, D., Makihara, S., Corsi, P. R., Viana Ade, T., Wiczer, M. V., Nakakubo, S. & Mimica, L. M. (1998) *Dis. Esophagus* **11,** 248–250.
17. Pajecki, D., Zilberstein, B., dos Santos, M. A. A., Ubriaco T. A., Quintanilba, A. G., Cecconello, I. & Gama-Rodrigues, T. (2002) *J. Gastrointestinal Surg.* **6,** 723–729.
18. Suau, A., Bonnet, R., Sutren, M., Godon, J. J., Gibson, G. R., Collins, M. D. & Dore, J. (1999) *Appl. Environ. Microbiol.* **65,** 4799–4807.
19. Pace, N. R. (1997) *Science* **276,** 734–740.
20. Nealson, K. H. & Conrad, P. G. (1999) *Philos. Trans. R. Soc. London B* **354,** 1923–1939.
21. Ollett, W. (1947) *J. Pathol. Bacteriol.* **59,** 357.
22. Knoth, K., Roberds, S., Poteet, C. & Tamkun, M. (1988) *Nucleic Acids Res.* **16,** 10932.
23. 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., *et al.* (2003) *Nucleic Acids Res.* **31,** 442–433.
24. Hugenholtz, T. & Huber, T. (2003) *Int. J. Syst. Evol. Microbiol.* **53,** 289–293.
25. Hasegawa, M., Kishino, J. & Yano, T. (1985) *J. Mol. Evol.* **21,** 160–174.
26. Chao, A. (1987) *Biometrics* **43,** 783–791.
27. Dooley, C. P., Cohen, H., Fitzgibbons, P. L., Bauer, M., Appleman, M. D., Perez-Perez, G. I. & Blaser, M. J. (1989) *N. Engl. J. Med.* **321,** 1562–1566.
28. Stackebrandt, E. G. B. (1994) *Int. J. Syst. Bacteriol.* **44,** 846–849.
29. Swidsinski, A., Ladhoff, A., Pernthaler, A., Swidsinski, S., Loening-Baucke, V., Ortner, M., Weber, J., Hoffmann, U., Schrieber, S., Dietel, M. & Lochs, H. (2002) *Gastroenterology* **122,** 44–54.
30. Colwell, R. R. (2000) in *Nonculturable Microorganisms in the Environment*, eds. Colwell, R. R. & Grimes, D. J. (Am. Soc. Microbiol. Press, Washington, DC), pp. 325–342.
31. Costerton, J. W., Steward, P. S. & Greenberg, E. P. (1999) *Science* **284,** 1318–1322.
32. Kazor, C. E., Mitchell, P. M., Lee, A. M., Stokes, L. N., Loesche, W. J., Dewhirst, F. E. & Paster, B. J. (2003) *J. Clin. Microbiol.* **41,** 558–563.
33. Lederberg, J. (2000) *Science* **288,** 287–293.

MICROBIOLOGY