Applied and Environmental Microbiology, Sept. 2005, p. 5494–5500 Vol. 71, No. 9
0099-2240/05/$08.00+0 doi:10.1128/AEM.71.9.5494–5500.2005
Copyright © 2005, American Society for Microbiology. All Rights Reserved.

Detection of Diverse New *Francisella*-Like Bacteria in Environmental Samples†

Susan M. Barns,[1] Christy C. Grow,[1]‡ Richard T. Okinaka,[1] Paul Keim,[2] and Cheryl R. Kuske[1]*

Biosciences Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545,[1] and Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011[2]

Received 7 January 2005/Accepted 4 April 2005

Following detection of putative *Francisella* species in aerosol samples from Houston, Texas, we surveyed soil and water samples from the area for the agent of tularemia, *Francisella tularensis*, and related species. The initial survey used 16S rRNA gene primers to detect *Francisella* species and related organisms by PCR amplification of DNA extracts from environmental samples. This analysis indicated that sequences related to *Francisella* were present in one water and seven soil samples. This is the first report of the detection of *Francisella*-related species in soil samples by DNA-based methods. Cloning and sequencing of PCR products indicated the presence of a wide variety of *Francisella*-related species. Sequences from two soil samples were 99.9% similar to previously reported sequences from *F. tularensis* isolates and may represent new subspecies. Additional analyses with primer sets developed for detection and differentiation of *F. tularensis* subspecies support the finding of very close relatives to known *F. tularensis* strains in some samples. While the pathogenicity of these organisms is unknown, they have the potential to be detected in *F. tularensis*-specific assays. Similarly, a potential new subspecies of *Francisella philomiragia* was identified. The majority of sequences obtained, while more similar to those of *Francisella* than to any other genus, were phylogenetically distinct from known species and formed several new clades potentially representing new species or genera. The results of this study revise our understanding of the diversity and distribution of *Francisella* and have implications for tularemia epidemiology and our ability to detect bioterrorist activities.

The genus *Francisella* comprises two species of gram-negative coccobacilli, *F. tularensis* and *F. philomiragia*. *F. tularensis* is the etiologic agent of tularemia in humans and animals and can occur as pneumonic or ulceroglandular disease. *F. tularensis* is highly infectious (exposure to less than 10 organisms can cause disease), and if left untreated, mortality from infection with this pathogen may be as high as 30 to 60% of cases (9). These characteristics made this organism the focus of historical biological warfare research programs in the United States, Japan, and the former Soviet Union (9). The U.S. Centers for Disease Control (CDC) lists *F. tularensis* among the category A potential biological terrorism agents (28), and it is one of the pathogens monitored by the BioWatch aerosol surveillance program (4, 25) for potential bioterrorist attacks.

An average of 124 cases of tularemia were recorded annually in the United States in the last decade (6), and hundreds of cases per year are reported from tularemia-endemic areas of Europe (3, 35). *F. philomiragia* infections affect mainly near-drowning victims and immunocompromised patients but nonetheless can cause severe disease (22). Comprising only these two validly published named species, the currently described diversity of the genus *Francisella* is rather limited (34). Several closely related endosymbiots of ticks, including the named species *Wolbachia persica* (18) as well as species identified only by analysis of DNA sequence data (29), also affiliate with the *Francisellaceae* based on analysis of 16S rRNA sequences. Recent genetic analyses, however, have suggested that considerable diversity within the genus remains to be discovered (23).

Francisella species are known to infect ≈150 species of vertebrate animals and may also be associated with protozoa in the environment (1). This broad host distribution has impeded understanding of *Francisella* ecology and epidemiology (27). *Francisella* strains are extremely difficult to culture from environmental sources (27), and few studies of its natural distribution and diversity have been undertaken (16), usually after human disease outbreaks have been reported (3, 13, 19). The organism can persist in water or mud for at least a year (5), supporting the possibility that environmental matrices may be important reservoirs for this pathogen. Recent and historical outbreaks indicate that environmental exposure to the organism is a significant source of morbidity (3, 14, 21).

PCR-based analyses have been developed and used to detect *Francisella* species in water (16) and tissue samples in a few cases (11, 17, 19, 30, 36), but such analyses have not yet been reported for soil or other environmental samples. To better understand the natural diversity and ecology of this pathogen and its closely related species, it is important to further explore its distribution and diversity in the environment. This information is also critical to understanding the natural background of the pathogen in environmental samples that may be collected for detection and attribution of biological threat agents.

We analyzed soil and water samples for the presence of

* Corresponding author. Mailing address: Los Alamos National Laboratory, Bioscience Division, Mail Stop M888, Los Alamos, NM 87545. Phone: (505) 665-4800. Fax: (505) 665-3024. E-mail: kuske@lanl.gov.

† This is publication number LA-UR-04-8972.

‡ Present address: Section of Microbiology, University of California, Davis, California.

Francisella-like DNA sequences to follow up on the detection of putative *F. tularensis* in Houston, Texas, by BioWatch aerosol monitors in October 2003 (4). We employed a nested approach, using PCR primer sets of increasing specificity to detect and identify *Francisella* sequences present in DNA extracts from the samples. PCR products were then cloned and sequenced to provide additional information on the diversity of species present.

MATERIALS AND METHODS

Environmental sampling and DNA extraction. Three hundred forty-one surface soil samples (30 to 50 ml each) and 23 water samples (approximately 100 ml each) were collected as single-grab samples in small plastic bags or tubes throughout the eastern Houston, Texas, metropolitan area in November 2003. Samples were stored and transported on ice to the Los Alamos National Laboratory and stored at −70°C for approximately 10 days after collection. Total DNA was extracted from approximately 0.25 to 0.35 ml of the soil samples using the MoBio UltraClean soil DNA isolation kit (MoBio Laboratories, Carlsbad, CA), following the manufacturer's protocols. Briefly, soil was suspended in buffer with 0.1-um glass beads (in a 96-well format) and rapidly shaken to disrupt cells. DNA was purified from soil and cell debris by binding to a silica membrane, washing, and elution. For the water samples, cells were pelleted from 50-ml subsamples by centrifugation, followed by extraction of DNA from the pellet using the MoBio kit. Bacterial thermolysates were used as sources of DNA from laboratory *Francisella* strains (Table 1), as previously described (12).

PCR survey of environmental DNAs. Initially, all samples were amplified with primers 27F and 787Rb, targeting the small-subunit rRNA genes of all bacteria. Results from this amplification (not shown) indicated that >95% of the extracts were sufficiently pure and contained sufficient DNA to support PCR. Each 25-μl PCR contained 10 mM Tris (pH 8.4), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM deoxynucleoside triphosphates, 0.1 μM each primer, 0.94U of Ampli*Taq* LD polymerase (Perkin-Elmer), 5 μg bovine serum albumin (Boehringer Mannheim), and 1 μl soil or water DNA. Cycling conditions were as follows: 4 min denaturation at 94°C; 40 cycles of 55°C for 45 s, 72°C for 60 s, and 94°C for 30 s; and a final cycle of 55°C for 45 s and 72°C for 5 min (20 min for reactions to be cloned), carried out in a PTC-200 thermal cycler (MJ Research).

To screen for the presence of 16S rRNA gene sequences related to *Francisella*, each soil and water sample was amplified with primers Fr153F0.1 (5′-GCCCA TTTGAGGGGGATACC-3′) and Fr1281R0.1 (5′-GGACTAAGAGTACCTTT TTGAGT-3′), modified from the F11 and F5 primers of Forsman et al. (18) to increase sensitivity and specificity, at an annealing temperature of 60°C. These primers were designed to detect *F. tularensis* and *F. philomiragia* as well as the *Francisella*-like tick endosymbionts related to *Wolbachia persica* (29).

Soil samples that gave positive results with the *Francisella* 16S rRNA gene screen were tested further by amplification of extracted DNAs with a series of primers designed to be more specific for *F. tularensis* (ISFTu2F/R, targeting an insertion element-like sequence; 23kDaF/R, targeting the *23kDa* gene, which is expressed upon macrophage infection; and the Tul4F/R and FopAF/R pairs, targeting genes encoding outer membrane proteins) (36). To differentiate among *F. tularensis* subgroups, primers SdhF (5′-AAGATATATCAACGAGCKTTT-3′) and SdhR (5′-AAAGCAAGACCCATACCATC-3′), targeting a putative succinate dehydrogenase locus (*sdhA*), were designed and used in PCR and sequencing analyses to identify differences among 48 isolates representing *F. tularensis* subspecies and *F. philomiragia*. Primers SdhF/R were used at an annealing temperature of 56°C, and the ISFTu2F/R, 23kDaF/R, Tul4F/R and FopAF/R pairs at used at an annealing temperature of 60°C.

Positive control reactions using DNA from *F. tularensis* LVS (ATCC 29684) (1 and 0.1 pg), as well as negative-control reactions without DNA, were included in each experiment. *F. tularensis* DNA was added to control reactions only after all experimental reaction tubes were sealed to prevent false positives due to contamination. Five microliters of each reaction mixture was analyzed on a 1, 2, or 3% (depending on expected product size) agarose gel. DNA was visualized by ethidium bromide staining and UV transillumination. Putatively positive reactions were repeated to confirm results and obtain products for cloning. PCR amplification of extracts from *Francisella* laboratory isolates was performed as above, using the 27F/1492R primer pair for the 16S rRNA gene (24) and SdhF/R for the *sdhA* gene (Table 1).

Cloning and sequencing of PCR products. PCR products from all of the *Francisella* primer sets were purified by electrophoresis on SeaKem agarose gels, the bands were excised and purified using a Qiaex DNA purification kit (QIAGEN, Inc., Chatsworth, CA), using the manufacturer's protocol. Products were cloned into the pCR4 vector, using the TOPO-TA cloning kit and the manufacturer's protocols (Invitrogen, San Diego, CA). For most soil and water PCR products, 48 to 96 clones were picked and stored in glycerol medium for sequencing. For 16S rRNA and *sdhA* gene sequences from *Francisella* reference isolates, two clones were picked and sequenced on both strands for each isolate and the sequences were compared, in an effort to reduce *Taq* polymerase-induced errors.

Plasmid DNA was isolated from overnight cultures using a solid-phase reverse immobilization procedure (10), and inserts were sequenced using the M13 forward primer or a PCR primer, with the BigDye terminator cycle sequencing reagents (v3.0, Applied Biosystems, Foster City, CA). Sequencing reactions were analyzed on ABI 3700 and 3730 automated sequencers (Applied Biosystems). Preliminary analysis of partial 16S rRNA gene sequences from environmental clone libraries was used to select representatives for full sequencing. Full-length, double-stranded sequence was obtained from these selected environmental and isolate clones using the M13F and M13R primers and additional primers internal to the 16S rRNA gene, 533Fb (5′-GCCAGCAGCNGCGGTAA-3′), 940Fb.Ft (5′-CGGGGACCCGCACAAGC-3′), 910Rb.Ft (5′-GTCCCCGTCAATTCCTT TGAG-3′), and 517Rb0.1 (5′-ATTACCGCIGCTGCTGGC-3′) (modified from reference 24 for use with *Francisella* spp. sequences).

Analysis of sequence data. Raw data were analyzed using Sequencher (Gene Codes, Inc.) software. The Check_Chimera program (8) was used to screen for chimeric 16S rRNA sequences, which were removed from subsequent analyses. Comparisons were made to database sequences using the RDP Sequence Match program (7) and NCBI BLAST 2.0 (01/05) (2). For phylogenetic analyses of 16S rRNA gene data, sequences were obtained from databases and aligned using Clustal X (33), with final alignment accomplished manually using the GDE multiple sequence editor (8). Sequence analyses were performed on approximately 1,125 aligned nucleotides from each rRNA gene sequence, corresponding to the length of clones obtained from the environmental PCRs.

Phylogenetic trees were inferred using maximum-likelihood analysis (fastDNAml version 1.1, distributed by RDP) (26). Additional distance (minimum evolution) and parsimony analyses (not shown) used to assess support for placement of new sequences were performed using PAUP* (v. 4.0b10 for Macintosh) (31). This program was also used to calculate 16S rRNA gene sequence similarities and infer trees from *sdhA* sequence data. For sequences from amplification reactions using the primers of Versage et al. (36), sequence comparisons and similarity calculations were performed using BLAST analyses (2).

Nucleotide sequence accession numbers. Sequences representative of each new sequence type obtained in this study, as well as from the reference isolates (Table 1), have been deposited in GenBank under accession no. AY968223 to AY968239 (16S rRNA gene sequences from reference isolates); AY968283 to AY968305 (16S rRNA gene sequences from Houston soil clones); AY968240 to AY968282 (*sdhA* sequences from reference isolates); AY968306 to AY968310 (representative *sdhA* sequences from Houston soil clones); and AY973868 to AY973879 (representative sequences from Houston soil samples generated with the primers of Versage et al.).

RESULTS

16S rRNA sequence analysis. To survey broadly for the presence of *Francisella* species and relatives, DNA extracts from 364 Houston soil and water samples were amplified by PCR with primers targeting the small-subunit rRNA genes of *Francisella* spp. and related tick symbionts. This screening indicated that seven soil samples and one water sample contained 16S rRNA gene sequences related to *Francisella*. Five of the seven soil samples and the one positive water sample were obtained from a marshy spoils area, while samples 027 and 045 were obtained from a lawn and the bank of a drainage ditch, respectively.

A clone library was constructed from the PCR products of each sample, and a total of 311 good-quality 16S rRNA gene sequences were obtained and analyzed. All sequences obtained affiliated most closely with sequences from the *Francisella* genus by BLAST analysis, confirming the specificity of the primers. Representative clones for each sequence type obtained

TABLE 1. *Francisella* reference isolates sequenced in this study

Isolate			GenBank accession no.	
Species	Subspecies	Strain(s)	16S rRNA gene	*sdhA* gene
F. tularensis	*tularensis*	FSC 053	AY968223	
F. tularensis	*tularensis*	FSC 054	AY968224	
F. tularensis	*tularensis*	FSC 199	AY968225	
F. tularensis	*tularensis*	FSC 237, Schu S4	AY968226	AY968240
F. tularensis	*tularensis*	MC14		AY968241
F. tularensis	*tularensis*	CO 003111		AY968242
F. tularensis	*tularensis*	KS 00-0948		AY968243
F. tularensis	*tularensis*	MA 00-2970		AY968244
F. tularensis	*tularensis*	MA 00-2972		AY968245
F. tularensis	*tularensis*	MA 00-2973		AY968246
F. tularensis	*tularensis*	MA 00-2987		AY968247
F. tularensis	*tularensis*	OK-CAN		AY968248
F. tularensis	*tularensis*	OK-CHK		AY968249
F. tularensis	*tularensis*	OK-HUG		AY968250
F. tularensis	*tularensis*	OK-OKL-1		AY968251
F. tularensis	*tularensis*	OK-OKL-2		AY968252
F. tularensis	*tularensis*	OK-TUL-1		AY968253
F. tularensis	*tularensis*	SD 00-3146		AY968254
F. tularensis	*tularensis*	SD 00-3147		AY968255
F. tularensis	*holarctica*	FSC 017	AY968227	
F. tularensis	*holarctica*	FSC 022	AY968228	
F. tularensis	*holarctica*	FSC 025	AY968229	
F. tularensis	*holarctica LVS*	FSC 155, ATCC 29684	AY968230	AY968256
F. tularensis	*holarctica*	FSC 257	AY968231	
F. tularensis	*holarctica*	NM 00-2642		AY968257
F. tularensis	*holarctica*	UT 01-1901	AY968232	
F. tularensis	*holarctica*	83A-7152		AY968258
F. tularensis	*holarctica*	84A-3697		AY968259
F. tularensis	*holarctica*	85A-3896		AY968260
F. tularensis	*holarctica*	86A-4765		AY968261
F. tularensis	*holarctica*	89A-2909		AY968262
F. tularensis	*holarctica*	89A-7092		AY968263
F. tularensis	*holarctica*	90A-2057		AY968264
F. tularensis	*holarctica*	91A-3318		AY968265
F. tularensis	*holarctica*	94A-3157		AY968266
F. tularensis	*holarctica*	97A-3245		AY968267
F. tularensis	*holarctica*	99A-2628		AY968268
F. tularensis	*holarctica*	99A-6712		AY968269
F. tularensis	*holarctica*	99A-9419		AY968270
F. tularensis	*holarctica*	Cal26		AY968271
F. tularensis	*holarctica*	Cal28		AY968272
F. tularensis	*mediaasiatica*	FSC 122	AY968233	
F. tularensis	*mediaasiatica*	FSC 147	AY968234	AY968273
F. tularensis	*mediaasiatica*	FSC 148	AY968235	AY968274
F. tularensis	*mediaasiatica*	FSC 149	AY968236	AY968275
F. tularensis	*novicida*	FSC 040, ATCC 15482	AY968237	AY968276
F. tularensis	*novicida*-like	FSC 156 fx1	AY968238	AY968277
F. tularensis	*novicida*-like	FSC 157 fx2		AY968278
F. philomiragia		FSC 037		AY968279
F. philomiragia		FSC 038		AY968280
F. philomiragia		FSC 039		AY968281
F. philomiragia		FSC 144, ATCC 25015	AY968239	AY968282

were sequenced in their entirety (approximately 1,170 bp), and cloned rRNA gene sequences of an additional 17 *Francisella* isolates were determined for reference.

Phylogenetic analyses revealed a surprising variety of sequence types present in the samples (Fig. 1). None of the sequences obtained from the soil samples were identical to any 16S rRNA sequences available in GenBank or in our reference collection, and the majority of the sequences from the samples did not group closely with previously reported sequences from any *Francisella* isolates. These sequences instead fell into three phylogenetically distinct clusters (Fig. 1; groups II, III, and V) containing only novel sequences from this study. These new clusters showed high sequence similarity (≥99.1%) within each cluster, but they were only more distantly related to sequences from isolates (≤98.5% sequence identity to closest outlying group). Phylogenetic analysis by several methods gave strong support (97 to 100% support in bootstrap analysis) to the coherence of these clusters and their separation from other, previously known sequence types.

Based on levels of sequence identity suggested in previous

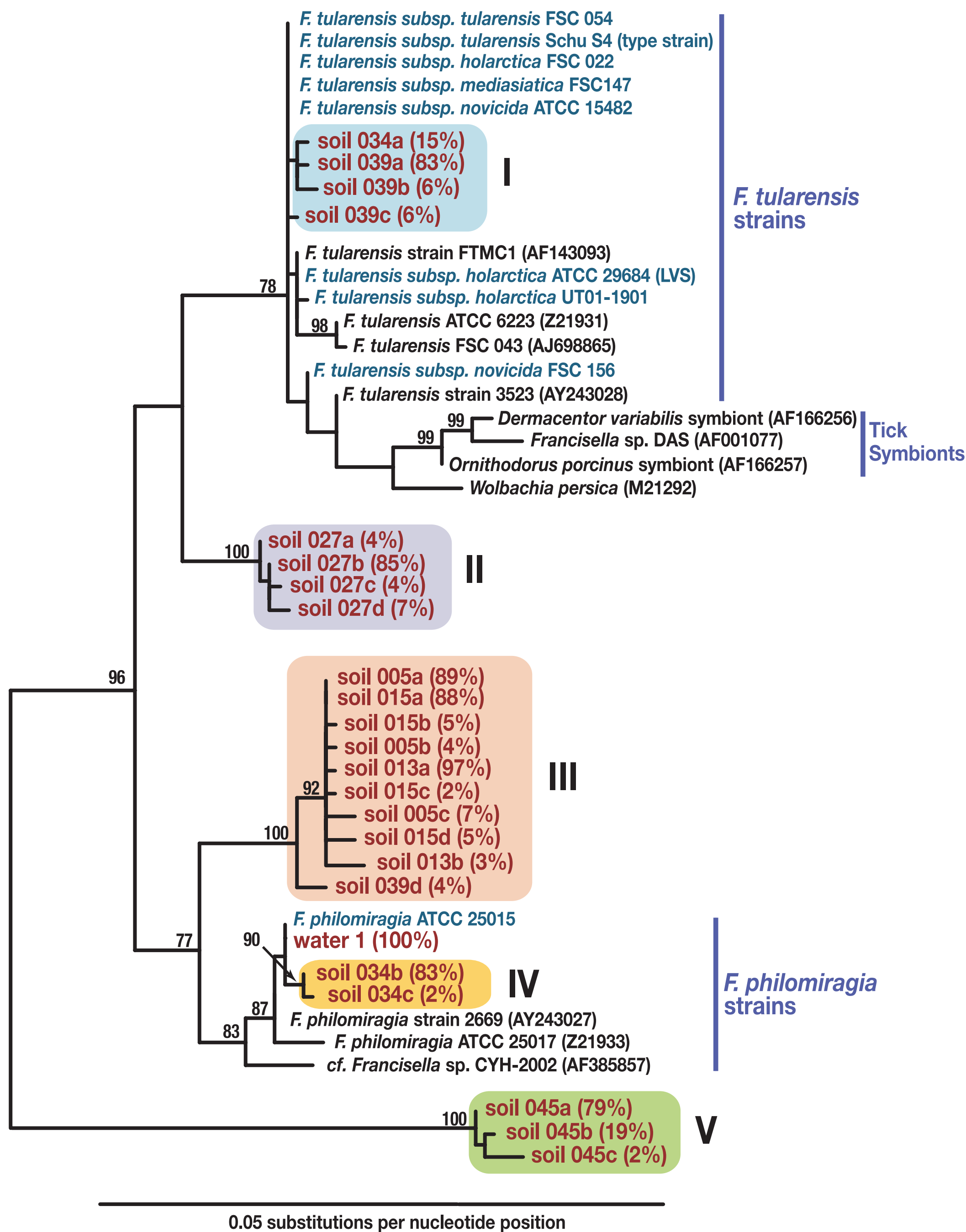

FIG. 1. Phylogenetic tree showing relationships of small-subunit rRNA gene sequences obtained from environmental samples to those of *Francisella* and related species. The tree was rooted using sequences of *Escherichia coli*, *Thiothrix ramosa*, *Caedibacter taenospiralis*, *Piscirikettsia salmonis*, and *Thiomicrospira thyasirae* (not shown). The percentage of 100 bootstrap resamplings that support each topological element in maximum-likelihood analysis is indicated, for values of >70%. Phylogenetic groups of sequences obtained in this study are labeled I to V. Within these groups, sequences representative of the types obtained from Houston soil clone libraries (in red) are labeled with the sample name (soil samples 005 to 045 and one water sample) and by distinct sequence types within each sample (a to d). Numbers in parentheses indicate the percentage of clones with identical sequences obtained from that sample. Sequences in blue text are reference sequences obtained in this study. Sequences in black were obtained from GenBank, with accession numbers shown in parentheses. The scale bar corresponds to 0.05 substitution per nucleotide position.

16S rRNA analysis of the genus (12), the organisms from which these sequences derived probably constitute several new species of *Francisella*. However, there was moderate support for the clustering of group III sequences with those of *F. philomiragia*, suggesting a specific relationship between these groups. Sequences in group V were only rather distantly related (≤93.7% identity) to any other sequences in the database, but were more similar to those of *Francisella* spp. than to

any non-*Francisella* sequences (87.8% identity to the sequence of the closest non-*Francisella* relative, *Caedibacter tenospiralis*). Based on this level of sequence identity, our analysis suggests that the organisms from which sequences in group V derived may constitute a new genus of *Francisella*-related species.

Sequences were also obtained which affiliated closely with those of *F. philomiragia* strains isolated from water. Although the sequences from the one positive water sample from Houston were all identical to that of *F. philomiragia* ATCC 25015, none of the soil-derived sequences were an exact match to any *F. philomiragia* sequence in the database (Fig. 1, group IV). In addition, bootstrap support was high for grouping these sequences to the exclusion of other *F. philomiragia* sequences, and these may represent new strains or subspecies of *F. philomiragia* organisms in these samples.

The majority of sequences from sample 039 and a few from 034 differed by only 1 to 2 nucleotides from sequences of *F. tularensis* isolates (Fig. 1, group I). This level of nucleotide identity is similar to that reported for different subspecies of *F. tularensis* (18), and bootstrap analysis also supported inclusion of the soil sequences within the *F. tularensis* group. These data indicate that the environmental organisms may constitute new subspecies of *F. tularensis*.

***sdhA* sequence analysis.** To further investigate the genetic relationship between the soil organisms and known isolates, we used a PCR primer set capable of discriminating among the subspecies of *F. tularensis*. From available *Francisella* species genome sequence data, the *sdhA* gene, which encodes a putative succinate dehydrogenase, was found to differentiate strains at the subspecies level based on specific single-nucleotide polymorphism (SNP) signatures. The SdhA primer set was tested in PCR against DNAs from a diverse group of 48 *F. tularensis* isolates (including representatives of *F. tularensis* subspp. *tularensis*, *holarctica*, *mediaasiatica*, and *novicida*) and four isolates of *F. philomiragia*. Maximum-parsimony analysis of the *sdhA* sequences from the 48 known isolates revealed 43 SNPs common to all *F. tularensis* isolates that distinguish them from *F. philomiragia* strains (Fig. 2). In contrast, strains of *F. tularensis* clustered much more closely in this analysis and were generally well resolved into subspecies clusters by the sequence of their *sdhA* genes. Three of the subspecies of *F. tularensis* form closely related but distinct clades with single-nucleotide polymorphisms separating *F. tularensis* subsp. *tularensis*, *F. tularensis* subsp. *holarctica*, and *F. tularensis* subsp. *mediaasiatica*. *F. tularensis* subsp. *novicida* isolates form a more distant clade, with two and three SNP resolution between the *F. tularensis* subsp. *novicida* and *F. tularensis* subsp. *novicida*-like isolates and the other biovars. The separation of the individual subspecies of *F. tularensis* by only a limited number of SNP differences is consistent with multilocus variable-number tandem repeat studies indicating the clonal evolution of this species (12, 23).

Of the seven Houston soil samples positive with the 16S rRNA gene primer set, soil samples 015, 027, 034, 039, and 045 gave PCR products and *sdhA*-like sequences with the Sdh primers (Table 2, Fig. 2). None of the sequences fell into *F. tularensis* subsp. *tularensis*, *F. tularensis* subsp. *holarctica*, or *F. tularensis* subsp. *mediaasiatica* clades, but 54 of 56 clones formed a new *F. tularensis* group equally closely related to *F. tularensis* subsp. subsp. *holarctica* and *F. tularensis* subsp. *novicida*, with representatives obtained from each of the five soil samples (Table 2). Within this new subspecies cluster, three sequence types (Fig. 2) were distinguishable by a unique SNP. This result further supports the presence of novel *F. tularensis* subspecies in these samples. In addition, two clones from the 034 soil sample contained sequences that grouped closely with the sequences of the *F. philomiragia* isolates, differing by only two to three SNPs from known isolates (Fig. 2). This result concurs with the recovery of 16S rRNA gene sequences similar to those of *F. philomiragia* from this sample.

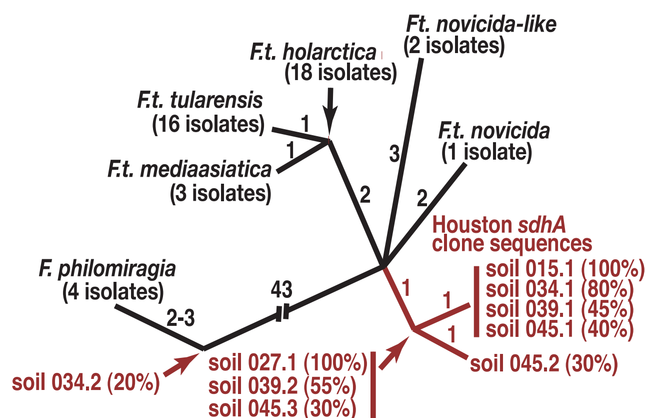

FIG. 2. Maximum-parsimony phylogenetic tree of *sdhA* sequences from Houston soil organisms and *Francisella* reference strains. Sequences representative of the types obtained from Houston soil clone libraries (in red) are labeled with the sample name (soil samples 015 to 045) and sequence type (0.1, 0.2, and 0.3). Numbers in parentheses indicate the percentage of clones with identical sequences obtained from that sample. Sequences in black text are reference sequences obtained in this study. F.t., *F. tularensis*. Numbers above branches indicate the number of single-nucleotide sequence polymorphisms that distinguish each group.

***F. tularensis*-specific primer sets.** The primer sets described by Versage et al. (36) were originally designed to be used in TaqMan assays. To capture as broad a group of *Francisella* species as possible, the primer pairs were used in standard PCR assays, without use of the corresponding TaqMan probes, and the resulting amplicons were cloned and sequenced. Only the two soil samples, 034 and 039, from which *F. tularensis*-type 16S rRNA gene sequences were obtained were positive with all four primer pairs (Table 2). Sequencing results for products cloned from these reactions showed that both samples contain sequences for these four gene targets that are identical to those from *F. tularensis* strains available in the database. Primer sets 23kDaF/R, FopAF/R, and IsFtu2F/R all produced only sequences matching previously reported *F. tularensis* sequences for these samples. In contrast, cloned sequences from soil sample 034 PCR products were more variable with the Tul4F/R primer pair.

Although a few clones in this library matched the sequence of *F. tularensis tul4* or differed at one position, most clones in this library differed at seven positions from the database sequence. These are predicted to be silent mutations, indicating that these sequences derived from *tul4* homologues in this sample. In addition, the FopAF/R primer set gave positive PCR results with sample 027, suggesting lower specificity, since in this sample only species group II was detected (Fig. 1).

TABLE 2. Clone sequencing results for primer sets used in this study

Sample source/type	Sample no.	16S rRNA primer(s) (no. of groups detected)[a,d]	*sdhA* primers (no. of sequence types)[b,d]	% Identity[c] with *E. tularensis*-specific primer[d]			
				Tul4	FopA	ISFtu	23kDa
Houston/soil	034	I, IV (46)	2 (12)	92–100 (33)	100 (27)	100 (39)	100 (46)
Houston/soil	039	I, III (44)	2 (11)	100 (27)	100 (38)	100 (38)	100 (46)
Houston/soil	005	III (47)	—	—	—	—	—
Houston/soil	013	III (34)	—	—	—	—	—
Houston/soil	015	III (41)	1 (11)	—	—	—	—
Houston/soil	027	II (46)	1 (11)		97 (42)	—	—
Houston/soil	045	V (42)	2 (11)	—	—	—	—
Houston/water	195	F.phil. (11)	n.t.	n.t.	n.t.	n.t.	n.t.

[a] Phylogenetic groups I to V of Fig. 1. F.phil., sequences were identical to that of *F. philomiragia* (ATCC 25015).
[b] —, no detectable product obtained from these samples; n.t., not tested.
[c] Percent identity of sequences obtained from each sample to *F. tularensis* target gene sequences, as determined by BLAST analysis.
[d] Numbers in parentheses are total numbers of sequences analyzed.

Cloned sequences from these products were all identical and differed by two bases from the *F. tularensis* sequence available in the database, both predicted to be silent mutations.

DISCUSSION

This DNA-based survey identified three new bacterial groups related to *Francisella* as well as potential new *F. tularensis* and *F. philomiragia* subspecies in soil samples. It is evident from our analysis that each soil sample contained a mixture of sequence types, potentially obtained from a variety of strains or species in the sample. Unfortunately, since we are analyzing a mixture of DNAs from many organisms rather than from single isolates, it is not possible to correlate sequences from different gene primer sets with individual strains. Overall, however, our results indicate that a wide variety of previously unknown types of *Francisella* are present in these samples, some of which have the potential to be detected by assays designed to be specific for *F. tularensis*.

F. tularensis has been divided into several subspecies based on geographic distribution and disease potential. *F. tularensis* subsp. *tularensis* (also known as biovar A) is considered the most virulent type, and while human cases have been reported only in North America, the pathogen has also been found in Europe (20). *F. tularensis* subsp. *holarctica* (biovar B) causes most cases of tularemia in Europe and also occurs in North America and Japan. Although infection with this subspecies is rarely fatal, it is nonetheless highly infectious and causes significant morbidity in Europe. The other two described subspecies, *F. tularensis* subsp. *mediasiatica* and *F. tularensis* subsp. *novicida*, are most commonly isolated from areas of Central Asia and in North America and Australia, respectively, and are less commonly associated with human disease (36). Our results suggest the existence of additional *F. tularensis* subspecies in soil samples, the pathogenicity of which is unknown. Two clinically significant isolates (Fx1 and Fx2) recovered from immunocompromised patients in Galveston, Texas, and Liberty County, Texas, were characterized as *F. tularensis* subsp. *novicida*-like (Fig. 2) based on growth characteristics, extragenic palindromic sequences, and specific biochemical tests (7). Although these Texas isolates are distinct in sequence from those recovered from the Houston soil samples, they emphasize the existence of other unusual *Francisella* strains that can cause disease in humans.

Francisella species are able to enter a viable but nonculturable state (15) and are particularly refractive to cultivation from environmental samples (27). This has likely contributed to our limited knowledge of their diversity, distribution, and ecology. Using a DNA-based survey in place of cultivation, we have identified genes suggesting the presence of several new *Francisella* species. Our results support the observation made with cultured isolates that *Francisella* species are very diverse (12), but the natural reservoirs for this diverse group of species remain largely unknown (15). This is the first report of the recovery of *Francisella* DNA sequences from soil samples. In a parallel analysis of soil samples from Denver, Colo. (not shown), we have obtained additional sequences that cluster with the novel group II (Fig. 1), suggesting that some of these new groups may indeed be widespread in the environment. Continued DNA-based survey of soil and other environmental samples, coupled with continued culture attempts from these natural sources, will help define the distribution, ecology, transmission, and pathogenicity of the new groups of *Francisella* described here.

As a result of increased concern over terrorist use of agents such as *F. tularensis*, sensitive PCR-based monitoring systems have been developed and deployed to detect the presence of pathogens in environmental samples (25). In addition, high-resolution DNA-based strain typing systems are being developed to provide platforms for epidemiologic and bioforensic analyses (5, 12, 23, 32). The ability of these approaches to differentiate an agent introduced in a bioterrorist attack from naturally occurring strains requires an extensive understanding of the diversity and distribution of the organism and its related species that share genetic traits. This is especially true for the *Francisella* group. *Francisella* species inhabit a wide variety of ecological niches. Species sharing considerable genetic similarity to human and animal pathogens are free-living or are symbionts or pathogens of invertebrates such as insects and amoebae (19, 20) that may not be associated with human disease. Even within the *F. tularensis* group, isolates representing the different subspecies are very similar in genomic characteristics but display a wide range of pathogenicity characteristics.

Currently, little is known about the mechanisms of *Fran-*

cisella virulence (34), but comparison with nonpathogenic environmental isolates may shed light on this area. Further DNA-based studies of environmental samples using the methods described here should enhance our ability to identify and characterize new strains and species of *Francisella*. Isolation and characterization of environmental isolates will contribute significantly to the development of more specific and informative assays for pathogen detection and forensics, as well as for monitoring epidemiology and environmental sources of natural outbreaks of tularemia.

ACKNOWLEDGMENTS

We thank Lori Merrill, Judy Buckingham, Andrew Shearer, Stephanie Redman, Rita Svensson, Jason Farlow, and Cheryl Strout for technical help; Richard Urie, Dustie Rich, Mario Medina, Mark Welch, Patrick Girault, and Carmela Romero for Houston environmental sample collection; and the CDC/DVBID Ft. Collins for providing DNA from their *Francisella* reference collection.

This work was supported by grants to C.R.K., P.K., and R.O. from the DOE Chemical & Biological Nonproliferation program and the DHS.

REFERENCES

1. **Abd, H., T. Johansson, I. Golovliov, G. Sandström, and M. Forsman.** 2003. Survival and growth of *Francisella tularensis* in *Acanthamoeba castellanii*. Appl. Environ. Microbiol. **69:**600–606.
2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389–3402.
3. **Anda, P., J. Segura del Pozo, J. M. Díaz Garcia, R. Escudero, F. J. Garcia Peña, M. C. López Velasco, R. E. Sellek, M. R. Jiménez Chillarón, L. P. Sanchez Serrano, and J. F. Martînez Navarro.** 2001. Waterborne outbreak of tularemia associated with crayfish fishing. Emerg. Infect. Dis. **7:**575–582.
4. **Barton, K., and R. Obey.** 9 October 2003. Officials following up on bacteria detection. http://www.ci.houston.tx.us/departme/health/bacteria%20detection.htm. Houston Department of Health and Human Services and Harris County Public Heath and Environmental Services News Release, Houston, Tex.
5. **Broekhuijsen, M., P. Larsson, A. Johansson, M. Byström, U. Eriksson, E. Larsson, R. G. Prior, A. Sjöstedt, R. W. Titball, and M. Forsman.** 2003. Genome-wide DNA microarray analysis of *Francisella tularensis* strains demonstrates extensive genetic conservation within the species but identifies regions that are unique to the highly virulent *F. tularensis* subsp. tularensis. J. Clin. Microbiol. **41:**2924–2931.
6. **Centers for Disease Control.** 2002. Tularemia—United States, 1990–2000. Morb. Mortal. Wkly. Rep. **51:**181–184.
7. **Clarridge, J. E., 3rd, T. J. Raich, A. Sjöstedt, G. Sandström, R. O. Darouiche, R. M. Shawar, P. R. Georghiou, C. Osting, and L. Vo.** 1996. Characterization of two unusual clinically significant *Francisella* strains. J. Clin. Microbiol. **34:**1995–2000.
8. **Cole, J. R., B. Chai, T. L. Marsh, R. J. Farris, Q. Wang, S. A. Kulam, S. Chandra, D. M. McGarrell, T. M. Schmidt, G. M. Garrity, and J. M. Tiedje.** 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.
9. **Dennis, D. T., T. V. Inglesby, D. A. Henderson, J. G. Bartlett, M. S. Ascher, E. Eitzen, A. D. Fine, A. M. Friedlander, J. Hauer, M. Layton, S. R. Lillibridge, J. E. McDade, M. T. Osterholm, T. O'Toole, G. Parker, T. M. Perl, P. K. Russell, and K. Tonat.** 2001. Tularemia as a biological weapon: medical and public health management. JAMA **285:**2763–2773.
10. **Elkin, C. J., P. M. Richardson, H. M. Fourcade, N. M. Hammon, M. J. Pollard, P. F. Predki, T. Glavina, and T. L. Hawkins.** 2001. High-throughput plasmid purification for capillary sequencing. Genome Res. **11:**1269–1274.
11. **Emanuel, P. A., R. Bell, J. L. Dang, R. McClanahan, J. C. David, R. J. Burgess, J. Thompson, L. Collins, and T. Hadfield.** 2003. Detection of *Francisella tularensis* within infected mouse tissues by using a hand-held PCR thermocycler. J. Clin. Microbiol. **41:**689–693.
12. **Farlow, J., K. L. Smith, J. Wong, M. Abrams, M. Lytle, and P. Keim.** 2001. *Francisella tularensis* strain typing using multiple-locus, variable-number tandem repeat analysis. J. Clin. Microbiol. **39:**3186–3192.
13. **Feldman, K. A., R. E. Enscore, S. L. Lathrop, B. T. Matyas, M. McGuill, M. E. Schriefer, D. Stiles-Enos, D. T. Dennis, L. R. Petersen, and E. B. Hayes.** 2001. An outbreak of primary pneumonic tularemia on Martha's Vineyard. N. Engl. J. Med. **345:**1601–1606.
14. **Feldman, K. A., D. Stiles-Enos, K. Julian, B. T. Matyas, S. R. Telford, 3rd, M. C. Chu, L. R. Petersen, and E. B. Hayes.** 2003. Tularemia on Martha's Vineyard: seroprevalence and occupational risk. Emerg. Infect. Dis. **9:**350–354.
15. **Forsman, M., E. W. Henningson, E. Larsson, T. Johansson, and G. Sandström.** 2000. *Francisella tularensis* does not manifest virulence in viable but non-culturable state. FEMS Microbiol. Ecol. **31:**217–224.
16. **Forsman, M., A. Nyrén, A. Sjöstedt, L. Sjökvist, and G. Sandström.** 1995. Identification of *Francisella tularensis* in natural water samples by PCR. FEMS Microbiol. Ecol. **16:**83–92.
17. **Forsman, M., G. Sandström, and B. Jaurin.** 1990. Identification of *Francisella* species and discrimination of type A and type B strains of *F. tularensis* by 16S rRNA analysis. Appl. Environ. Microbiol. **56:**949–955.
18. **Forsman, M., G. Sandström, and A. Sjöstedt.** 1994. Analysis of 16S ribosomal DNA sequences of *Francisella* strains and utilization for determination of the phylogeny of the genus and for identification of strains by PCR. Int. J. Syst. Bacteriol. **44:**38–46.
19. **Goethert, H. K., I. Shani, and S. R. Telford, 3rd.** 2004. Genotypic diversity of *Francisella tularensis* infecting *Dermacentor variabilis* ticks on Martha's Vineyard, Massachusetts. J. Clin. Microbiol. **42:**4968–4973.
20. **Greub, G., and D. Raoult.** 2004. Microorganisms resistant to free-living amoebae. Clin. Microbiol. Rev. **17:**413–433.
21. **Helvaci, S., S. Gedikoglu, H. Akalin, and H. B. Oral.** 2000. Tularemia in Bursa, Turkey: 205 cases in ten years. Eur. J. Epidemiol. **16:**271–276.
22. **Hollis, D. G., R. E. Weaver, A. G. Steigerwalt, J. D. Wenger, C. W. Moss, and D. J. Brenner.** 1989. *Francisella philomiragia* comb. nov. (formerly *Yersinia philomiragia*) and *Francisella tularensis* biogroup novicida (formerly *Francisella novicida*) associated with human disease. J. Clin. Microbiol. **27:**1601–1608.
23. **Johansson, A., J. Farlow, P. Larsson, M. Dukerich, E. Chambers, M. Byström, J. Fox, M. Chu, M. Forsman, A. Sjöstedt, and P. Keim.** 2004. Worldwide genetic relationships among *Francisella tularensis* isolates determined by multiple-locus variable-number tandem repeat analysis. J. Bacteriol. **186:**5808–5818.
24. **Lane, D. J.** 1991. 16S/23S rRNA sequencing, p. 115–175. *In* E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Wiley and Sons, New York, N.Y.
25. **Marburger, J.** 2003. BioSecurity 2003. Keynote address on national preparedness. [Online.] http://www.ostp.gov/html/10-20-03%20jhm%20BioSecurity%202003.pdf.
26. **Olsen, G. J., H. Matsuda, R. Hagstrom, and R. Overbeek.** 1994. fastDNAmL: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. **10:**41–48.
27. **Petersen, J. M., M. E. Schriefer, K. L. Gage, J. A. Montenieri, L. G. Carter, M. Stanley, and M. C. Chu.** 2004. Methods for enhanced culture recovery of *Francisella tularensis*. Appl. Environ. Microbiol. **70:**3733–3735.
28. **Rotz, L. D., A. S. Khan, S. R. Lillibridge, S. M. Ostroff, and J. M. Hughes.** 2002. Public health assessment of potential biological terrorism agents. Emerg. Infect. Dis. **8:**225–230.
29. **Scoles, G. A.** 2004. Phylogenetic analysis of the *Francisella*-like endosymbionts of Dermacentor ticks. J. Med. Entomol. **41:**277–286.
30. **Sjöstedt, A., U. Eriksson, L. Berglund, and A. Tärnvik.** 1997. Detection of *Francisella tularensis* in ulcers of patients with tularemia by PCR. J. Clin. Microbiol. **35:**1045–1048.
31. **Swofford, D. L.** 2001. PAUP*: Phylogenetic analysis using parsimony (*and other methods), 4th ed. Sinauer Associates, Sunderland, Mass.
32. **Thomas, R., A. Johansson, B. Neeson, K. Isherwood, A. Sjöstedt, J. Ellis, and R. W. Titball.** 2003. Discrimination of human pathogenic subspecies of *Francisella tularensis* by using restriction fragment length polymorphism. J. Clin. Microbiol. **41:**50–57.
33. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The ClustalX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24:**4876–4882.
34. **Titball, R. W., A. Johansson, and M. Forsman.** 2003. Will the enigma of *Francisella tularensis* virulence soon be solved? Trends Microbiol. **11:**118–123.
35. **Tärnvik, A., G. Sandström, and A. Sjöstedt.** 1996. Epidemiological analysis of tularemia in Sweden 1931–1993. FEMS Immunol. Med. Microbiol. **13:**201–204.
36. **Versage, J. L., D. D. Severin, M. C. Chu, and J. M. Petersen.** 2003. Development of a multitarget real-time TaqMan PCR assay for enhanced detection of *Francisella tularensis* in complex specimens. J. Clin. Microbiol. **41:**5492–5499.