Whole-Genome Relationships among *Francisella* Bacteria of Diverse Origins Define New Species and Provide Specific Regions for Detection

Jean F. Challacombe,[a] Jeannine M. Petersen,[b] La Verne Gallegos-Graves,[a] David Hodge,[c] Segaran Pillai,[d] Cheryl R. Kuske[a]

Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA[a]; Division of Vector-Borne Infectious Diseases, Centers for Disease Control and Prevention, Fort Collins, Colorado, USA[b]; Chemical and Biological Division, Science and Technology Directorate, Department of Homeland Security, Washington, DC, USA[c]; Office of Laboratory Science and Safety, U.S. Food and Drug Administration, Silver Spring, Maryland, USA[d]

ABSTRACT *Francisella tularensis* is a highly virulent zoonotic pathogen that causes tularemia and, because of weaponization efforts in past world wars, is considered a tier 1 biothreat agent. Detection and surveillance of *F. tularensis* may be confounded by the presence of uncharacterized, closely related organisms. Through DNA-based diagnostics and environmental surveys, novel clinical and environmental *Francisella* isolates have been obtained in recent years. Here we present 7 new *Francisella* genomes and a comparison of their characteristics to each other and to 24 publicly available genomes as well as a comparative analysis of 16S rRNA and *sdhA* genes from over 90 *Francisella* strains. Delineation of new species in bacteria is challenging, especially when isolates having very close genomic characteristics exhibit different physiological features—for example, when some are virulent pathogens in humans and animals while others are nonpathogenic or are opportunistic pathogens. Species resolution within *Francisella* varies with analyses of single genes, multiple gene or protein sets, or whole-genome comparisons of nucleic acid and amino acid sequences. Analyses focusing on single genes (16S rRNA, *sdhA*), multiple gene sets (virulence genes, lipopolysaccharide [LPS] biosynthesis genes, pathogenicity island), and whole-genome comparisons (nucleotide and protein) gave congruent results, but with different levels of discrimination confidence. We designate four new species within the genus; *Francisella opportunistica* sp. nov. (MA06-7296), *Francisella salina* sp. nov. (TX07-7308), *Francisella uliginis* sp. nov. (TX07-7310), and *Francisella frigiditurris* sp. nov. (CA97-1460). This study provides a robust comparative framework to discern species and virulence features of newly detected *Francisella* bacteria.

IMPORTANCE DNA-based detection and sequencing methods have identified thousands of new bacteria in the human body and the environment. In most cases, there are no cultured isolates that correspond to these sequences. While DNA-based approaches are highly sensitive, accurately assigning species is difficult without known near relatives for comparison. This ambiguity poses challenges for clinical cases, disease epidemics, and environmental surveillance, for which response times must be short. Many new *Francisella* isolates have been identified globally. However, their species designations and potential for causing human disease remain ambiguous. Through detailed genome comparisons, we identified features that differentiate *F. tularensis* from clinical and environmental *Francisella* isolates and provide a knowledge base for future comparison of *Francisella* organisms identified in clinical samples or environmental surveys.

KEYWORDS *Francisella* species, *Francisella tularensis*, *Francisella novicida*, *Francisella*

Address correspondence to Cheryl R. Kuske, Kuske@lanl.gov.

holarctica, *Francisella philomiragia*, *Francisella opportunistica* (MA06-7296), *Francisella salina* (TX07-7308), *Francisella uliginis* (TX07-7310), *Francisella frigiditurris* (CA97-1460), *Allofrancisella*, whole-genome comparisons, nucleic acid sequence identity, amino acid identity, new *Francisella* species

The species concept for bacteria remains tenuous despite much recent work to delineate consistent and stringent molecular criteria for species identification (1, 2). For decades, a bacterial species was considered to be a collection of isolates having 70% or greater DNA-DNA reassociation values and at least one diagnostic phenotypic trait (3). In recent years, the relative merits of whole-genome- and core genome-based criteria (1, 2, 4–6) have been discussed alongside ecological and adaptive assignments (7–9) and the more traditional criteria. Currently, whole bacterial genomes are readily sequenced, either from isolated cultures or directly from the environment, and the numbers of genome sequences representing historical bacterial taxa are increasing tremendously. The *Francisella* genus exemplifies this situation, since members of this genus are associated with various clinical and environmental sources and include highly virulent human and animal pathogens, opportunistic human pathogens, fish pathogens, tick endosymbionts, and seemingly free-living organisms inhabiting brackish water. Taxonomic relationships in the *Francisella* genus have remained ambiguous, confounding specific detection of pathogenic species in a background of innocuous environmental species (28), thereby making it difficult to discriminate intentional outbreaks from natural incidences (11).

Francisella tularensis is a highly virulent risk 3 (biosafety level 3 [BSL-3]) pathogen that causes the disease tularemia. This pathogen was weaponized during past world wars and is considered a tier 1 priority pathogen. Within *F. tularensis*, three subspecies designations (*tularensis*, *holarctica*, *mediasiatica*) are generally accepted based upon epidemiology, biochemical and genome characteristics, geographic distribution, and virulence (reviewed in references 12, 13, and 14). *F. tularensis* subsp. *tularensis* (also called type A) is found in North America and is the most virulent subspecies. *F. tularensis* subsp. *holarctica* (also called type B) is found in Europe, Japan, and North America and is less virulent. *F. tularensis* subsp. *mediasiatica* is found primarily in Central Asia. While *F. tularensis* subsp. *mediasiatica* causes a type B-like illness in rabbits, documented human cases are lacking from the published literature. In support of these designations are multiple high-quality genomic sequences (15–23) that have enabled detailed comparative studies among the *F. tularensis* subspecies. Indeed, the numerous *F. tularensis* genome sequences have provided a valuable public resource in lieu of sharing cultured isolates, which is often restricted due to the classification of *F. tularensis* as a risk 3, tier 1 pathogen, which requires biosafety level 3 containment.

Over the past 10 years, the known diversity within the *Francisella* genus has expanded significantly, from two major species groups, *F. tularensis* and *F. philomiragia* (24–26), to four species groups, adding *F. novicida* and *F. noatunensis*, and additional potential species and representatives that do not cluster closely with the named species (26–37). The expansion of diversity in the *Francisella* genus is derived from four sources; environmental samples, human clinical samples, fishes, and ticks. First, through environmental sequencing surveys, novel 16S rRNA and succinate dehydrogenase flavoprotein subunit gene (*sdhA*) sequences have been identified and grouped into five clades that potentially represent new species or genera (26, 28). Environmental isolates representative of three of these clades were successfully cultured (38), and high-quality, finished genomes were obtained (32, 33). Additional environmental species have been isolated from cooling system water in China (*Allofrancisella guangzhouensis*, originally described as *Francisella guangzhouensis*) (30, 39, 40), California (*Francisella* sp. strain CA97-1460), and Germany (*Francisella* sp. strain W12-1067) (41). Human clinical isolates have been cultured from skin lesions of otherwise healthy individuals (31, 42) or from compromised patients with variant symptoms, including near-drowning (43), respiratory (44), and cerebrospinal and blood (27) infections. These clinical isolates do not

"type" into the known species groups, although some of them have been designated *F. novicida*-like (42, 44). One isolate (LA11-2445 [31]) appears to be most closely related to *F. halioticida*, a pathogen from abalone (35). Third, multiple *Francisella* isolates, some identified as *F. noatunensis*, have been obtained from farmed and wild marine fish (reviewed in references 34, 45, and 46). Finally, *Wolbachia persica* endosymbionts and other *Francisella*-related organisms have been detected in ticks (47–54). The distribution of these isolates is global, suggesting a worldwide distribution of *Francisella* organisms that inhabit fish, tick, and a variety of other environmental sources.

Designations surrounding the species *F. novicida* have changed since the species was first described in 1955 (29, 55–58). *F. novicida* has been proposed as a fourth subspecies of *F. tularensis*, but genomic, virulence, transmission, and ecological characteristics justify why *F. novicida* should be considered a species separate from *F. tularensis* (13, 57). Many novel isolates residing outside the *F. tularensis*, *F. novicida*, *F. philomiragia*, and *F. noatunensis* species remain uncharacterized or incompletely characterized, with undesignated isolates described as strain code numbers or as *Francisella* cf. species (closest to). Recent attempts to conduct comparative genomic studies among members of the *Francisella* genus either have been conducted on only a few representatives (32, 33, 59), have focused on subspecies within *F. tularensis* (60, 61), or have been conducted on incomplete, lower-quality draft genomes (62, 63), where interpretation of relationships may be confounded by missing genetic features in the incomplete genomes.

DNA- or protein-based detection methods must be able to distinguish pathogenic *F. tularensis* from other *Francisella* species that reside in the environment or in other animal hosts. This is challenging because of the variety of *Francisella* species discussed above and also because their genomes are small (about 2 Mbp) (Table 1) and very highly conserved. To address this gap, genomic comparisons were performed in this study and included clinical and environmental *Francisella* isolates outside the *F. tularensis* clade (Table 1). We conducted whole-genome comparisons on 31 genomes, 7 of which were newly sequenced for this study. We then compared the results of the whole-genome approaches to the ability of the historical single or multigene approaches to achieve species delineation in the genus. Based on these genomic comparisons, we now identify multiple new species and provide a comparative framework for delineation of species in the *Francisella* genus.

RESULTS

Species discrimination using single genes and a set of 36 conserved housekeeping proteins. Table 1 lists the *Francisella* isolates included in this study and their sources, original published descriptions, and genome characteristics. To benchmark our whole-genome comparisons with the prior literature, we constructed 16S rRNA and *sdhA* gene phylogenies from 94 and 92 *Francisella* isolates, respectively (Fig. 1A; see also the full 16S rRNA gene tree in Fig. S1 in the supplemental material; the full *sdhA* tree is shown in Fig. S2 in the supplemental material). The *F. tularensis* isolates generally clustered together in the 16S rRNA gene phylogeny, but sequence conservation was high and bootstrap values were low. Although widely used, the 16S rRNA gene is not a strong species discriminator in this genus; major clades may be defined, but the placement of specific isolates is variable. The *sdhA* gene was a more robust discriminator between the *F. tularensis* clade and the other potential species groups (Fig. 1A; the full tree of 92 representatives is shown in Fig. S2). The *F. tularensis* and *F. novicida* isolates separated from the others (98% bootstrap support), but the support for differentiation of *F. tularensis* isolates from *F. novicida* was only 38% (Fig. S2). The support values for the minor branches in this group varied (Fig. S2). In the *sdhA* phylogeny, four isolates (MA06-7296, TX07-7308, TX07-7310, and CA97-1460) (Table 1) were well separated from both the *F. tularensis* and *F. novicida* isolates (Fig. 1A; Fig. S2).

The AMPHORA approach for species comparisons uses 31 protein markers (64). A phylogeny was constructed following this approach; five additional *Francisella*-specific protein markers were included, and all the markers are listed in Table S1 in the

TABLE 1 Origin and features of *Francisella* isolate genomes

Genome	Species	Origin	Size (bp) (chromosome; plasmid[s])	Coding sequences (chromosome; plasmid[s])	NCBI accession number(s)	Reference(s)
Schu S4	*F. tularensis* subsp. *tularensis* type A1a	Human	1,892,819	1,604	NC_006570	21
WY96-3418	*F. tularensis* subsp. *tularensis* type A2a	Human	1,898,476	1,634	NC_009257; NZ_CP010103; NZ_CP010447	17
NE061598	*F. tularensis* subsp. *tularensis* type A1a	Human	1,892,681	1,836	NC_017453	96
DPG 3A-IS[a]	*F. novicida* (*F. tularensis*)	Environmental: warm spring	1,898,140	1,771	NZ_CP012037	76, 97
LVS	*F. tularensis* subsp. *holarctica* type B1	Vaccine strain	1,895,994	1,754	NC_007880	19
OSU18	*F. tularensis* subsp. *holarctica* type B	Beaver	1,895,727	1,555	NC_008369	22, 98
FSC147	*F. tularensis* subsp. *mediasiatica*	Gerbil	1,893,886	1,406	NC_010677	20
Toba04	*F. noatunensis* subsp. *orientalis*	Tilapia	1,847,202	1,595	NC_017909	84
FSC774[b]	*F. noatunensis* subsp. *noatunensis*	Cod	1,704,705	1,875	PRJNA73457	99
U112	*F. novicida*	Water near dead muskrats	1,910,031	1,719	NC_008601	61
Fx1[a]	*F. novicida*	Clinical: blood	1,913,619	1,818	NC_017450	32, 44
PA10-7858[a]	*F. novicida*	Clinical: blood	1,978,958; 41,013	1,921; 57	CP016635; NC_023026	43, 59
AZ06-7470[a]	*F. novicida*	Clinical: lymph node	1,890,780; 34,471	1,777; 51	CP009682; CP009683	100, 101
AL97-2214[a]	*F. novicida*	Human	1,916,455	1,802	CP009653	100
D9876[a]	*F. novicida*	Clinical: lymph node	1,870,206	1,753	NZ_CP009607	25, 86
F6168[a]	*F. novicida*	Clinical: blood	1,923,262; 3,978	1,792; 6	NZ_CP009353; NZ_CP009352	25, 86
TX07-6608[a]	*F. novicida*	Environmental: seawater	1,985,304; 2,621; 3,546; 82,910; 82,739	1,980; 1; 3; 91; 102	JRXS00000000	38
FSC454[b]	*F. hispaniensis*	Clinical: blood	1,885,078	1,823	PRJNA73391	29
3523[a]	*F. novicida*-like (*F. hispaniensis*)	Clinical: foot wound	1,945,310	1,854	NC_017449	32, 42
ATCC 25017[a]	*F. philomiragia*	Environmental: water	2,045,775	1,911; 4	NC_010336; NC_010331	86, 102
ATCC 25015[a]	*F. philomiragia*	Muskrat	2,017,400	1,815	NZ_CP010019	25, 44, 86
GA01-2794[a]	*F. philomiragia*	Human	2,148,038; 4,016	1,994; 5	NZ_CP009440; NZ_CP009441	85, 86
GA01-2801[a]	*F. philomiragia*	Human	2,022,507; 2,402; 8,805	1,944; 2; 8	NZ_CP009444; NZ_CP009445; NZ_CP009446	85, 86
TX07-7308[a]	New species (*F. salina*)	Environmental: seawater	2,035,931	1,976	NC_015696	33, 38
TX07-7310[a]	New species (*F. uliginis*)	Environmental: seawater	2,237,329	2,010	CP016796	38
CA97-1460[a]	New species (*F. frigiditurris*)	Cooling tower: USA	1,855,434; 6,175	1,806; 7	CP009654; CP009655	NA[d]
MA06-7296[a]	New species (*F.opportunistica*)	Clinical: CSF,[c] blood	1,824,527; 3,403	1,747; 5	CP016929; CP016930	27
FSC845[b]	*W. persica* (*F. persica*)	Tick	1,547,984	1,711	PRJNA73171	37, 103, 104
FSC1006	*F. endociliophora*	Environmental: seawater	2,015,987	1,606	NZ_CP009574	105
08HL01032	*A. guangzhouensis*	Cooling tower water: China	1,658,482; 3,045	1,480; 3	NZ_CP010427; NZ_CP010428	30, 39, 40
W12-1067	New, unnamed species	Cooling tower water: Germany	1,702,336	1,476	NZ_AWHF00000000	41

[a]Genomes sequenced at Los Alamos National Laboratory.
[b]Draft genomes: FSC454 (85 contigs), FSC774 (194 contigs), FSC845 (70 contigs).
[c]CSF, cerebrospinal fluid.
[d]NA, not applicable.

supplemental material. The overall relationships among the different *Francisella* isolates were similar to those shown for the *sdhA* tree, but the bootstrap support for each major branch was much higher (100%) (Fig. 1B; see the full tree in Fig. S3 in the supplemental material). *F. tularensis*, *F. philomiragia*, and *F. noatunensis* could readily be discriminated from other potential species by using the *sdhA* gene or 36 marker protein phylogenies. The rest of the clinical and environmental isolates either clustered with named species or represented an individual branch and were not always consistently placed with high confidence by these methods. To further define species groups, to more accurately place the novel isolates, and to identify candidate genome regions for specific detection, we conducted whole-genome comparisons using protein and DNA sequence-based assessments.

Whole-genome comparisons consistently discriminate multiple species and identify new *Francisella* species. Two whole-genome comparison approaches were
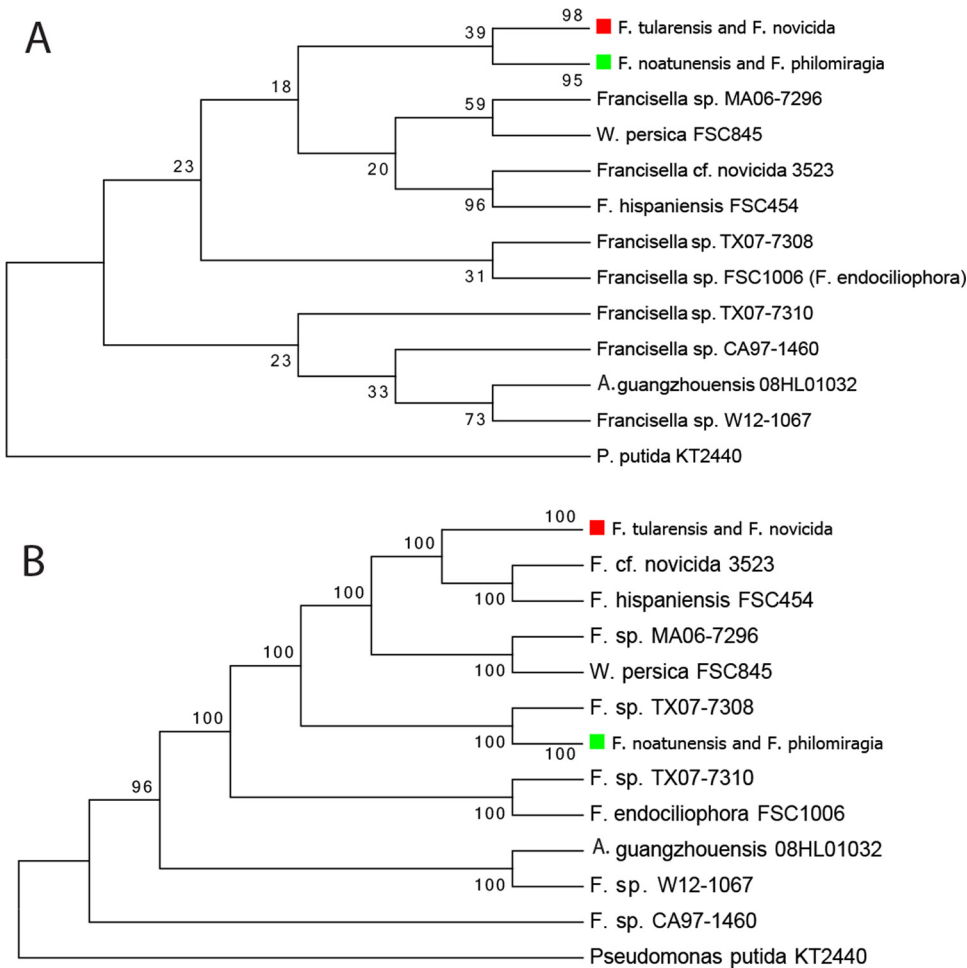
FIG 1 (A) Phylogenetic tree based on the *sdhA* gene from 92 *Francisella* isolates. With the exception of the *F. tularensis* subsp. *mediasiatica* clade, the branches representing the *F. tularensis* clades showed low bootstrap support values (much less than 50%) at the tips, so we collapsed some of the branches with low support (generally less than 10%) for easier viewing of the overall relationships. *P. putida*, *Pseudomonas putida*. (B) Phylogenetic tree based on a set of 36 protein markers (Table S1; Wu and Eisen [64]). Because the focus of this paper is on the *F. novicida*-like genomes and the potentially new species, we did not include all the *F. tularensis* isolates in this tree, just one representative from each subspecies. The red and green squares denote major branches of the trees. The minor branches with low bootstrap support values (generally under 10, indicating identical or nearly identical sequences) were collapsed to make this figure easier to view. The full trees showing all of the branches are presented in Fig. S2 and S3 in the supplemental material.

used to discriminate *F. tularensis* from other *Francisella* species and to distinguish new *Francisella* species. The first was a bidirectional best BLASTP hits analysis to compare the protein translations (coding sequences [CDS]) from each genome to each of the other genomes in a pairwise fashion (see Table S2 in the supplemental material and Fig. 2 for the genomes listed in Table 1). This analysis is based on computing the average amino acid identity (AAI) between pairs of genomes (65). The original presentation of this approach used a more inclusive cutoff for species clustering, 30% amino acid homology across 70% of the translated gene. For the present AAI analysis, we used a different threshold (BLASTP E value = 10^{-5}) to try to better resolve differences among the various *Francisella* members.

The genomes from three of the recognized species (*F. tularensis*, *F. novicida* [including *F. novicida*-like isolates], and *F. philomiragia*) generally had the highest AAI values of all the members of the same group (*F. tularensis*, 72 to 96%; *F. novicida*, 70 to 97%; *F. philomiragia*, 80 to 91%; see boxes in Table S2). There was some overlap in coding region conservation among the *F. tularensis* and *F. novicida* groups, with some members having high AAI values with members of both groups. The four genomes labeled
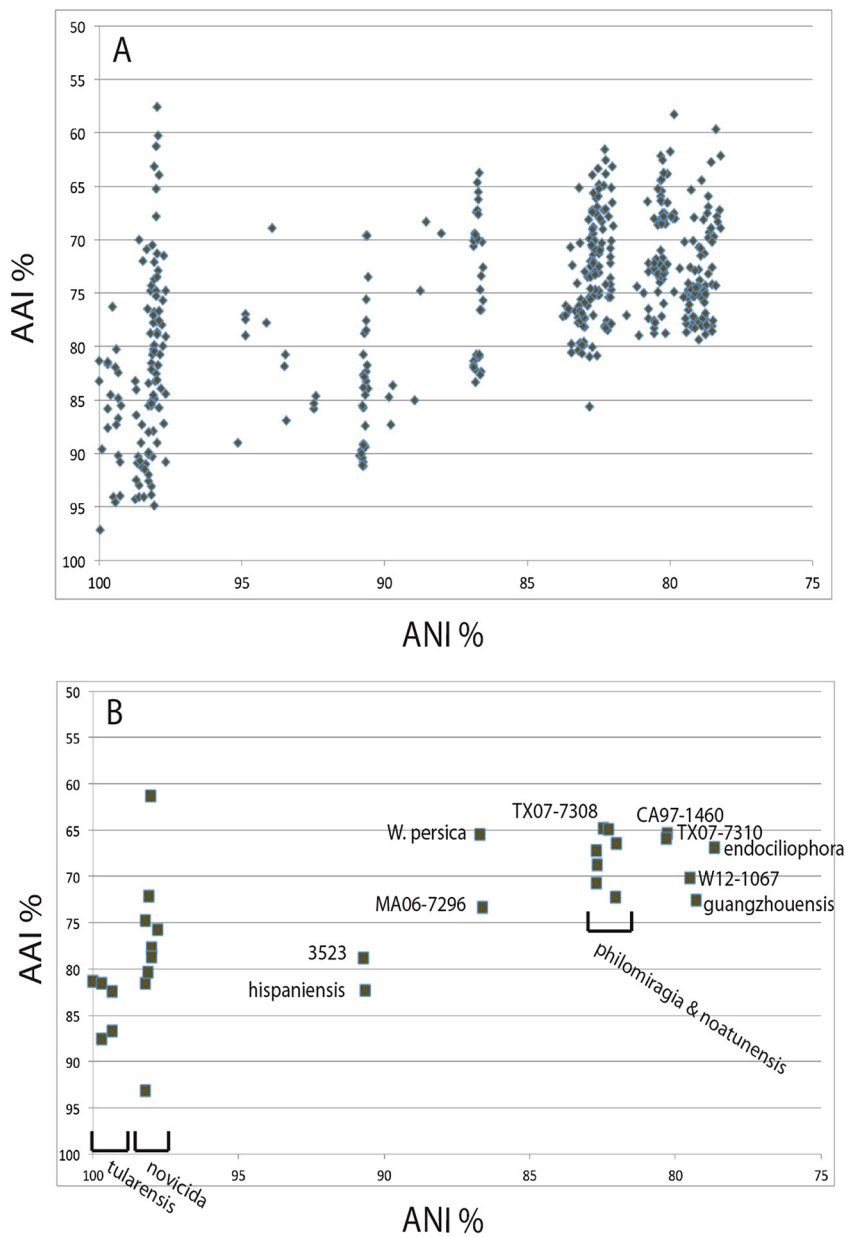
FIG 2 Relationships among calculated ANI and AAI percent between each pair of *Francisella* isolates listed in Table 1 (*n* = 31). (A) Relationships among all of the isolates; (B) calculated values for each isolate compared to the *F. tularensis* subsp. *tularensis* Schu4 genome.

with new species names in Table S2 (CA97-1460, MA06-7296, TX07-7308, and TX07-7310) generally had less than 91% AAI compared to all other genomes and also did not group closely with the *F. tularensis* and *F. novicida* strains in the phylogenetic trees (Fig. 1; see also Fig. S1 to S3 in the supplemental material).

Determining the average nucleotide identity (ANI) is a well-documented method for comparing genomes and assessing species membership through analysis of entire genome nucleotide sequences rather than translated coding sequences (1, 2, 4, 66). In our opinion, because the ANI analysis did not include a functional qualifier (i.e., translated sequence), it was generally less discriminating than the AAI comparison that we performed. All of the *F. tularensis*, *F. novicida*, and *F. novicida*-like genomes had ANI values greater than 97% (see Table S3 in the supplemental material). It was previously reported that an ANI percentage of 95% corresponds to a DNA-DNA hybridization

(DDH) threshold of 69% (66), which would define all these isolates as intraspecific. However, their vastly different phenotypes and differences at the amino acid level warrant keeping them separated (13; this study). The other genomes having ANI values indicating species identity were in the *F. philomiragia* and *F. noatunensis* groups. These genomes had ANI values above 93% compared to each other, but their ANI values were well under 90% compared to the rest of the genomes in Table S3. The *F. novicida*-like 3523 and *F. hispaniensis* FSC454 genomes were only 90 to 91% similar to either of the *F. tularensis* and *F. novicida* groups, but they were 98% similar to each other. The *Francisella* or *Allofrancisella* species lying outside the major *F. tularensis*, *F. novicida*, *F. philomiragia*, and *F. noatunensis* groups (*Allofrancisella guangzhouensis*, *F. endociliophora* FSC1006, *Francisella* sp. W12-1067), *W. persica* FSC845, and four proposed new species (TX07-7308, TX07-7310, CA97-1460, and MA06-7296), all had ANI values below 90% compared to the rest of the genomes, supporting their designation as separate species.

Correspondence between the AAI and ANI values was high when the values were compared (Fig. 2A). However, *Francisella* isolates with a similar percent ANI (*x* axis) exhibited a wide range of percent AAI values (*y* axis) (Fig. 2A). Although the calculated ANI values fell into particular numerical units, the corresponding AAI values for a particular ANI value spanned 20 to 30% variation, suggesting a wide range of amino acid translations that could affect protein composition and function (Fig. 2A).

A comparison of ANI versus AAI among the 31 complete genomes, compared only to the virulent, tularemia-causing strain, *F. tularensis* subsp. *tularensis* Schu4, is shown in Fig. 2B. This result indicates strong support for the new species designations of the *Francisella* spp. MA06-7296, TX07-7308, TX07-7310, and CA97-1460. Other *F. tularensis* isolates formed a cluster most closely related to *F. tularensis* subsp. *tularensis* Schu4; the *F. novicida* isolates formed a separate cluster. The rest of the isolates formed distinct groups with ANI values from 90 to 80% (Fig. 2B).

Francisella pathogenicity and virulence features. The ability of virulent *Francisella* isolates to survive inside host cells and cause disease depends on a set of 16 to 19 genes in the *Francisella* pathogenicity island (FPI) (67, 68). The highly virulent *F. tularensis* strains contain two copies of the FPI, while the less virulent *F. novicida* strains have a single copy (Table 2) (69). All of the genomes, except those of *A. guangzhouensis* 08HL01032 and *Francisella* sp. W12-1067, had at least one copy of the FPI. The W12-1067 genome did not have a complete FPI region, but an FPI-like island was previously identified in the genome (41). As shown in Table 2 and Fig. S4 in the supplemental material, there were four major patterns of FPI gene content and organization in the *Francisella* genomes. Pattern 1 was characteristic of *F. tularensis* subsp. *tularensis*, *F. tularensis* subsp. *mediasiatica*, some of the *F. novicida* and *F. novicida*-like isolates, *F. hispaniensis*, and possibly *W. persica* FSC845. Although the FSC845 genome was lacking a *pdpD* gene, it did have *pmcA*, which makes the FPI most similar to patterns 1 and 2. Since this is a draft genome, it is possible that the *pdpD* gene is in a gap between contigs. Although comparisons of the FPI gene neighborhoods support the pattern 1 designation for the FPIs from *F. tularensis* and *F. novicida*-like genomes, there is a 50-amino-acid insert in the *pdpD* gene in *F. novicida* and *F. tularensis* subsp. *mediasiatica* that is not present in *F. tularensis* subsp. *tularensis* or *F. tularensis* subsp. *holarctica* (68). The *pdpC* and *pdpE* genes in MA06-7296 did not align completely with the Schu S4 query sequences and seemed to be frameshifted (pseudogenes). In keeping with pattern 3, these results suggest that they may not be functional.

Pattern 2 was previously identified in the *F. tularensis* subsp. *holarctica* strains LVS and OSU18 (68, 69). In addition to the LVS and OSU18 genomes, we looked at the FPI genes in the *F. tularensis* subsp. *holarctica* FSC200 and FTFN002-00 genomes (data not shown), which also had FPI pattern 2. FPI patterns 3 and 4 were similar except for a truncated *iglG* gene and a small open reading frame (ORF) on the reverse strand below

TABLE 2 Comparisons of specific gene sets among *Francisella* genomes

Genome	FPI copies:pattern	Functional genes (loci)					*F. novicida* restriction-modification system			
		OppABCDF	SpeADE, AguAB	FTT0794 to FTT0796	FTT1188	FTT1453c, 54c, 58c	Locus 2, type I	Locus 3, type I	Locus 4, type III	Locus 6, type II
SchuS4	2:1	+	+	+	+	+	−	inc	inc	−
DPG 3A-IS	2:1	+	+	+	+	+	−	inc	inc	−
WY96-3418	2:1	+	+	+	+	+	−	inc	inc	−
NEO61598	2:1	+	+	+	+	+	−	inc	inc	−
FSC147	2:1	inc	+	+	+	+	−	inc	inc	−
Fth LVS	2:2	inc	+	+	+	+	−	inc	inc	−
Fth OSU18	2:2	inc	+	+	+	+	−	inc	inc	−
U112	1:1	+	−	−	−	−	+	+	+	+
AZ06-7470	1:1	+	−	−	−	+	−	+	+	−
AL97-2214	1:3	+	+	−	+	−	+	−	−	−
D9876	1:1	+	+	−	−	−	−	−	−	−
F6168	1:3	+	+	−	+	−	+	−	−	−
Fx1	1:1	+	−	−	−	−	−	+	−	−
3523	1:1	+	+	−	−	−	+	+	−	−
FSC454[c]	1:1	+	−	−	−	−	−	inc	−	−
PA10-7858	1:1	+	−	−	−	−	+	inc	+	−
TX07-6608	1:3	−	+	−	+	−	−	+	−	−
ATCC 25017	1:3	+	−	−	−	−	−	inc	inc	−
ATCC 25015	1:3	+	−	−	−	−	−	+	−	−
GA01-2794	1:3	+	−	−	−	−	inc	inc	−	−
GA01-2801	1:3	+	−	−	−	−	−	−	+	inc
TX07-7308	1:4	−	−	−	−	−	−	−	−	−
TX07-7310	1:3	−	−	−	−	−	−	−	−	−
FSC774[c]	1:3	inc	−	−	−	−	−	−	−	−
Toba04	1:4	inc	−	−	−	−	−	−	−	−
FSC845[b]	1:1-2	−	−	−	−	−	−	−	−	−
MA06-7296	1:3	−	−	−	−	−	+	+	−	−
CA97-1460	1:3	−	−	−	−	−	−	−	−	−
08HL01032	None	−	−	−	−	−	−	−	−	−
FSC1006	1:3	−	−	−	−	−	+	+	−	−
W12-1067	inc	−	−	−	−	−	−	−	−	−

[a]Abbreviations and symbols: FPI, *Francisella* pathogenicity island (patterns are illustrated in Fig. S4 in the supplemental material); OppABCDF, oligopeptide ABC transporter locus; SpeADE, spermidine biosynthesis; AguAB, putrescine biosynthesis; +, presence of the locus in each strain (shaded); −, absence of the locus in each strain; inc, missing 1 or more genes but not all (includes genes that are frameshifted, likely pseudogenes).
[b]Draft genome, *pdpD* missing.
[c]Draft genome, missing genes may be in an unsequenced gap.

iglG in pattern 4. Also, the genes in the regions flanking the FPI were different in some genomes.

We queried representative complete genomes for features that may be associated with tularemia virulence or other defining functions to determine if these might discriminate among *Francisella* species (Table 2). These functions include the oligopeptide transport and spermidine/putrescine biosynthesis (32, 33), seven unique lipopolysaccharides (LPSs) that have been previously identified in *F. tularensis* but not in other *Francisella* species (13, 63), and four restriction-modification (R-M) systems found in *F. novicida* (13, 70).

Several patterns were apparent in this comparison. The previously described seven unique LPS genes were present in all the *F. tularensis* genomes but only sporadically present among the *F. novicida* genomes. Only the *F. tularensis* genomes contained all three genes previously mentioned (13, 71, 72) to be involved in the formation of a cell surface capsule-like structure; these genes were FTT0794 (phosphocholine metabolism), FTT0795 (formylmethyltransferase), and FTT0796 (phosphocholine metabolism) (13, 73). The *F. novicida* U112 genome had only the OppABCDF locus and the four R-M systems identified previously (70). None of the *F. novicida* and *F. novicida*-like genomes had the same pattern as U112, although as a group they generally had similar patterns

of gene content compared to each other. *F. philomiragia* genomes, the other outlying *Francisella* species, and the novel species were notable in their absence of the *F. tularensis*- or *F. novicida*-specific gene sets, although their absence was not complete. Exceptions included the presence of a complete *F. novicida* type I R-M system in *F. philomiragia* 25015, a complete *F. novicida* type III R-M system in *F. philomiragia* GA01-2801, and two complete type I R-M systems in MA06-7296 and *F. endociliophora* FSC1006 (Table 2).

DISCUSSION

The discovery of novel *Francisella* organisms from clinical samples, fish, and natural and man-made environments (e.g., cooling systems) has greatly expanded the documented diversity within the genus. This has resulted in a collection of related organisms for which accurate species designations are lacking and has confounded the ability to specifically detect and identify *F. tularensis* in clinical and environmental samples. Although these isolates have highly conserved genomes, they vary in their pathogenicity and host range. To overcome the safety and security challenges that may preclude sharing of *Francisella* isolates, 7 *Francisella* isolates were sequenced to high quality at Los Alamos National Laboratory and in-depth genomic comparisons were conducted with other publicly available high-quality sequences of genomes in this genus. We found that, consistent with prior studies, the 16S rRNA gene was not an adequate species discriminator in this genus but that other single genes (*sdhA* and the genes listed in Table 2) could discriminate the known *F. tularensis* isolates from *F. novicida* and other *Francisella* species. A multisequence 36-protein-marker approach provided similar species relationships with higher confidence in the assignments. Whole-genome comparisons at the expressed-gene (AAI) and nucleotide (ANI) levels provided robust information to designate at least six, and likely more than six, species groups within this genus. Lastly, we demonstrate the utility of using genome sequences for the designation of new *Francisella* species as well as for providing a resource for comparative analyses in lieu of isolate sharing.

Bacterial species have historically been defined as organisms with at least 70% DNA-DNA hybridization to a "type species" organism, plus a common defining phenotype. In the *Francisella* genus, for which biochemical phenotypes that define genomic clades are difficult to find and organisms having very close genome characteristics may have very different environmental niches and pathogenesis, it is most robust to use genomic characteristics as an initial criterion upon which to define species groups. In agreement with previous reports (74, 75), we found that the highly pathogenic *F. tularensis* isolates grouped together in the single or multigene phylogenetic analyses (Fig. 1) but could not always be discriminated from members of the *F. novicida* group. The pairwise AAI and ANI comparisons (Fig. 2) clearly showed that *F. tularensis* genomes were most closely related to each other and could be separated from the *F. novicida* genomes. The ANI analysis revealed a significant degree of similarity between the *F. tularensis* and *F. novicida* groups (greater than 97%), which satisfies the 95% threshold for species inclusion (4, 66). However, the ANI and AAI correlations illustrated in Fig. 2 show that, at similar ANI values, the members of these species were separated by their protein-coding features, as the corresponding AAI values for a particular ANI value spanned 20 to 30% variation. Based on the whole-genome comparisons, particularly AAI, and the significant differences (e.g., FPI pattern, FTT0794-96 genes) in the gene set comparisons of the *F. tularensis* and *F. novicida* genomes (Table 2), we are in agreement with previous reports (13, 57) that these are distinct species.

An advantage of whole-genome comparisons is the removal of ambiguities from NCBI. One example of this is the *F. novicida* DPG 3A-IS "warm spring" isolate, which was described as an *F. novicida* isolate in the literature (76). All of the comparisons presented here show that, as of November 2016, the DPG 3A-IS genome sequence accessed at NCBI and used for comparison in this study originated from an *F. tularensis* isolate. The original DPG 3A-IS isolate was characterized as *F. novicida* by multiple

methods (97). Independent DNA purification and genome sequencing will be important in order to verify whether the original DPG 3A-IS can be confirmed as *F. novicida*.

In this study, ambiguity was also removed for two genomes from clinical and environmental isolates previously designated "*F. novicida*-like" based on their 16S rRNA gene sequence and various phenotypic methods. Based on single-gene comparisons, as well as the AAI and ANI data sets (Fig. 1; Tables S2 and S3), these genomes (Fx1 and TX07-6608) indeed clustered together with *F. novicida* genomes and generally shared a higher percentage of CDS, and we propose that they should be classified as *F. novicida*. Of the *F. novicida* and *F. novicida*-like isolates included in our comparisons, two were isolated from seawater (PA10-7858 and TX07-6608). Both genomes clustered with the other *F. novicida* genomes in the single-gene and multigene phylogenies and in the AAI comparison. Both isolates have large plasmids. However, unlike the other *F. novicida* genomes, TX07-6608 was missing the OppABCDF operon and had FPI pattern 3 (Table 2). Results from the two protein-based comparison approaches (36 protein markers and AAI whole-genome comparison) showed that one of the isolates originally thought to be *F. novicida*-like (3523) was not closely related to *F. novicida* (Fig. 1B and 2B). In all the genome comparisons presented here, *F. novicida*-like 3523 was found to share most traits with *F. hispaniensis* FSC454. These two isolates shared the highest AAI percentage with each other (89%, 90.5%) and had a bidirectional ANI value of 97.97%, which was the highest that either of these genomes had in any comparison (Fig. 2B; Tables S2 and S3). They also shared the same pattern and gene content in their single FPI. Sjodin et al. (63) suggested that both of these isolates should be designated members of the species *F. hispaniensis* based on whole-genome comparisons. Previous studies reported phenotypic differences between *F. hispaniensis* FSC454 and strain 3523 and concluded that these two strains were different (77). Indeed, these two isolates gave different patterns in the gene set comparisons (Table 2). This situation illustrates a continuing challenge for taxonomic binning of bacterial species in *Francisella*: the lack of a sufficiently large number of isolates and sequenced genomes to make robust designations.

The *F. philomiragia* strains included in our analysis came from a variety of sources. The *F. philomiragia* 25017 isolate was obtained from salt water collected from the Bear River Refuge in Utah, while strain 25015 came from a muskrat at the same location (78); *F. philomiragia* GA01-2794 and GA01-2801 were isolated from humans (79). *F. philomiragia* can cause opportunistic infections in humans with chronic granulomatous disease and in near-drowning patients (25, 80). Despite their various sources or origins, the *F. philomiragia* representatives grouped together in all the phylogenetic trees in the same major branch with the *F. noatunensis* clades. However, the *F. noatunensis* species do have several genetic, phenotypic, and physiological characteristics that discriminate them from *F. philomiragia* (45). The two subspecies of *F. noatunensis*, subsp. *noatunensis* (81) and subsp. *orientalis* (82, 83), were separated in all the trees, reflecting their fish host specificity. The pairwise AAI and ANI values of the *F. philomiragia* and *F. noatunensis* genomes showed the highest similarities to the other members of these groups, and their FPI regions and patterns of other gene sets were very similar. These results support the currently accepted species relationships among the *F. philomiragia* isolates and their similarity to *F. noatunensis* (84).

Our genome comparisons clearly demonstrate that four of the novel clinical and environmental isolates are new species, in addition to the W12-1067 isolate that was previously described as a new species (41). In all the phylogenies, the TX07-7308, TX07-7310, CA97-1460, and MA06-7296 genomes each clustered outside the major *F. tularensis*, *F. novicida*, *F. philomiragia*, and *F. noatunensis* groups. The AAI and ANI (Fig. 2B; Tables S2 and S3) values for these genomes were low compared to those of the other groups. Another similarity among these genomes was the FPI region, which was present in all the genomes as one copy in pattern 3 or 4. The presence/absence patterns of the gene sets in Table 2 was the same in TX07-7308, TX07-7310, *F. noatunensis* FSC774, Toba04, *W. persica* FSC845, the CA97-1460 isolate, and W12-1067. The MA06-7296 and FSC1006 genomes shared the same general features, which set

them apart from *F. tularensis* and the *F. novicida* groups. Collectively, this evidence supports the designation of each of these isolates as a separate species. We formally name them below. Using the genomic framework presented here, the rigor with which these species boundaries are being defined may change as additional isolates are discovered from clinical and environmental sources.

In summary, our results support the following conclusions about *Francisella* species designations. (i) U112, Fx1, PA10-7858, TX07-6608, AZ06-7470, AL97-2214, D9876, and F6168 are members of the *F. novicida* species group. We propose that the *F. novicida*-like isolates in this group be termed *F. novicida* instead of *F. novicida*-like. (ii) The MA06-7296 isolate (and the similar isolate PA05-1188) (27), the two isolates cultured from Galveston, TX, seawater (TX07-7308, TX06-7310), and the water cooling tower isolate, CA97-1460, are distinct enough from all other known *Francisella* species to be designated as separate species. (iii) The 3523 isolate and *F. hispaniensis* FSC454 are clearly members of the same phylogenetic group. Although the two isolates are not identical, the 3523 isolate is more closely related to *F. hispaniensis* FSC454 than it is to *F. novicida*, and we propose that they both be given the same species name.

The lack of clarity and consistency in phylogenetic analyses of *Francisella* species is due to the incredible ecological diversity of the *Francisella* genus. In an effort to overcome this significant barrier, we used comparative genomics methods to identify characteristic features in *Francisella* genomes and further distinguish related clades and species groups. Whole-genome comparisons significantly complemented the phylogenetic analyses and enabled us to compile a list of genomic features that discriminate *F. tularensis* from the other species and further refine the relationships among non-*F. tularensis* isolates. These features included seven previously mentioned genes encoding outer surface components of *F. tularensis* (13, 63), genes involved in oligopeptide transport, polyamine biosynthesis, restriction-modification systems (70), and characteristics of the FPI. In the future, the more sequenced representatives that we can obtain from all branches of the phylogeny, the better enabled we will be to accurately classify *Francisella* species using the framework provided here.

Bacteria in the *Francisella* genus are pathogens of animals and fish, and some appear to be free-living in the environment. Sharing *Francisella* isolates can be difficult; some members of the genus require special containment and classification. For example, *F. tularensis* is classified as a risk 3, tier 1 pathogen, requiring biosafety level 3 containment. In lieu of sharing cultured isolates, increased effort on high-throughput sequencing of *F. tularensis* and other *Francisella* bacteria is providing a valuable public resource for comparative genomics and trait identification. The specific genomic features that govern pathogenicity and virulence in different hosts remain uncertain. By sequencing and comparing whole genomes of *Francisella* isolates from different sources, we are beginning to tease apart the relationships among the different species and also to define new species. The four new *Francisella* species described below have not yet been deposited in public culture collections, but we are working toward that end. The annotated whole genomes of these four new species are now publicly available (see Materials and Methods).

Description of *Francisella opportunistica* sp. nov. *F. opportunistica* (op.por.tu.nis'ti.ca N.L. fem. adj. *opportunistica*, opportunistic, referring to the pathogen's ability to opportunistically infect immunocompromised or immune-suppressed humans). The type strain, PA05-1188, was isolated in 2005 from a patient with hemophagocytic syndrome and juvenile rheumatoid arthritis; the MA06-7296 strain was isolated from a patient with end-stage renal disease (27). Sequences of portions of the 16S rRNA and *sdhA* genes are identical for the two isolates (27). Cells are Gram-negative coccobacilli. Phenotypic characterization is as indicated in Table 3. *F. opportunistica* MA06-7296 showed an ANI of <87% compared to whole genomes of *F. tularensis*, *F. novicida*, *F. philomiragia*, and *F. noatunensis* (Fig. 2B; see Tables S2 and S3 in the supplemental material).

TABLE 3 Phenotypic characteristics of new *Francisella* spp.[a]

| New species | Result of biochemical reactivity assay for: | | | | Growth characteristics (48 h) in: | | | | | Agglutination using anti-*F. tularensis* serum[b] | Antimicrobial susceptibility[c] |
	Indole	Urease	Oxidase	Catalase	6.5% NaCl	SBA 25°C	SBA 35°C	CHAB 25°C	CHAB 35°C		
F. opportunistica MA06-7296	−	W/−	+	−	−	NG	NG	+	++	−	Susceptible to all
F. salina TX07-7308	−	−	−	+	+W	+	++	+	++	−	Susceptible to all
F. uliginis TX07-7310	−	−	−	+	+	NG	NG	+	++	−	Susceptible to all
F. frigiditurris CA97-1460[d]	−	−	−	+	+W	NG	NG	+	++	−	ND

[a]Abbreviations and symbols: ND, not determined; W, weak reaction; NG, no growth; −, negative; +, positive; ++, strong positive; SBA, sheep blood agar; CHAB, cysteine heart agar with 9% chocolatized sheep blood.
[b]Polyclonal rabbit serum to whole-killed (formalin-fixed) *F. tularensis*.
[c]Testing performed using Etest strips as described previously (27). Antimicrobials tested were ciprofloxacin, doxycycline, streptomycin, gentamicin, erythromycin, and tetracycline.
[d]Results for MA06-7296 were compiled from reference 27.

Description of *Francisella salina* sp. nov. *F. salina* (sa.li′na, L. fem. adj. *salina*, salty, referring to the original isolation of the type strain, TX07-7308, from brackish seawater and seaweed off the Gulf coast of Galveston, TX, USA, in 2007) (38). Cells are Gram-negative coccobacilli. Phenotypic characterization is as indicated in Table 3. By whole-genome ANI comparisons, *F. salina* TX07-7308 shows an ANI of <90% to *F. tularensis*, *F. novicida*, *F. philomiragia*, and *F. noatunensis* and is most closely related to *F. philomiragia* (Fig. 2B; Tables S2 and S3).

Description of *Francisella uliginis* sp. nov. *F. uliginis* (u.li′gi.nis. L. n. *uligo -inis*, brackish, marshy quality of the Earth; L. gen. n. *uliginis*, of moisture). The type strain, TX07-7310, was isolated from brackish seawater and seaweed off the Gulf coast of Galveston, TX, USA, in 2007 (38). A modified CHAB medium (CHAB-PACCV), containing polymyxin B, amphotericin B, cefepime, cycloheximide, and vancomycin, was used for isolation. Cells are Gram-negative coccobacilli. Phenotypic characterization is as indicated in Table 3. By whole-genome comparisons, *F. uliginis* TX07-7310 shows an ANI of <81% to *F. tularensis*, *F. novicida*, *F. philomiragia*, and *F. noatunensis* (Fig. 2B; Tables S2 and S3).

Description of *Francisella frigiditurris* sp. nov. *F. frigiditurris* (fri.gi.di.tur′ris. L. adj. *frigidus* cold; L. fem. n. *turris* tower: N.L. gen. n. *frigiditurris*, of a cooling tower), referring to the original isolation of this organism from a cooling tower in California, USA. The type strain is *F. frigiditurris* CA97-1460. Cells are Gram-negative coccobacilli. Phenotypic characterization is as indicated in Table 3. By whole-genome ANI comparisons, *F. frigiditurris* CA97-1460 shows an ANI of <80% to *F. tularensis*, *F. novicida*, *F. philomiragia*, and *F. noatunensis* (Fig. 2B; Tables S2 and S3).

MATERIALS AND METHODS

All the *Francisella* isolates were cultured and maintained, and DNA was extracted by the U.S. Centers for Disease Control and Prevention (CDC), in BSL-3 containment (38, 85). Genomic library construction, sequencing, and finishing of *Francisella* genomes were performed at Los Alamos National Laboratory as described previously (32, 76, 86). Genomes sequenced at Los Alamos were of finished quality (less than one error for every megabase of sequence). Genomes were annotated by the Rapid Annotation using Subsystems Technology (RAST) system (87), by NCBI (TX07-7310), or using an Ergatis-based workflow as described previously (86). All other genome sequences were obtained from GenBank.

16S rRNA and *sdhA* nucleotide sequences were obtained from public databases, primarily GenBank, or from our sequenced genomes, and for the genomes not yet sequenced, the 16S rRNA and *sdhA* sequences were obtained as follows. Samples were amplified with the Fr153F0.1 and Fr.1281.R0.1 primers, targeting the 16S rRNA. Each 25-μl reaction mixture consisted of 1× *Taq* LD buffer, 0.2 mM deoxynucleoside triphosphate (dNTP), 0.1 μM each primer, 1.25 U of AmpliTaq LD (Thermo Fisher Scientific), 5 μg bovine serum albumin (Roche), and 1 μl of 1 ng/μl DNA. The same protocol was used for the *sdhA* genes, using the SdhF/SdhR primers. See reference 26 for cycling conditions. Sanger paired-end reads were assembled using CLC Bio Genomics Workbench. All 16S rRNA nucleotide sequences were aligned using ClustalX (88) and trimmed to the same length. The same process was used

for *sdhA* nucleotide sequences. The total number of sequences was 94 for the 16S rRNA gene and 92 for the *sdhA* gene.

Using MEGA 6.06 (89), the *sdhA* nucleotide sequences were aligned by MUSCLE (90) (within MEGA) using default parameters. Maximum likelihood is well suited to the analysis of distantly related sequences, which is the case here. Maximum likelihood trees were constructed in MEGA using 500 bootstrap replicates (91), the Tamura-Nei substitution model with nucleotide substitutions and assuming uniform substitution rates among all sites. The maximum likelihood heuristic method was nearest-neighbor interchange, the initial tree was neighbor joining, and the branch swap filter was set to "very weak" to perform more exhaustive optimization and explore a larger search space (to take into account the larger phylogenetic distances between some of the isolates). All codons were included in the analysis. The same process was followed to construct both the 16S rRNA gene and the *sdhA* trees. Both 16S rRNA gene and *sdhA* trees were annotated in MEGA, collapsing branches with low/zero support.

Phylogenetic analysis of 36 protein markers (obtained from AMPHORA [64], plus additional *Francisella* protein markers), listed in Table S1 in the supplemental material, was performed as follows. For each genome, the protein translations for each gene listed in Table 1 were aligned using ClustalX (88) and trimmed to the same length, and the protein translations of all 36 genes from each genome were concatenated together in the same order. The concatenated sequences were aligned using MUSCLE (90) (within MEGA) using default parameters. Maximum likelihood trees were constructed in MEGA using 500 bootstrap replicates (91) and the Jones-Taylor-Thornton (JTT) substitution model with amino acid substitutions and assuming uniform substitution rates among all sites. The maximum likelihood heuristic method was nearest-neighbor interchange, the initial tree was neighbor-joining, and the branch swap filter was set to "very weak" to perform more exhaustive optimization and explore a larger search space (to take into account the larger phylogenetic distances between some of the isolates).

Protein coding sequences (CDS) from the sequenced genomes in Table 1 were compared using a bidirectional best BLAST hits (Basic Local Alignment Search Tool [92]) analysis as described previously (93). We used an E value cutoff of 10^{-5}. Briefly, the protein translations from each genome were compared to those of each other genome, and only the homologous sequences that were bidirectional best hits in each pairwise comparison were considered to be common to both genomes. While this approach is analogous to the AAI method described in reference 65, the percentages that we obtained were lower than the corresponding AAI values obtained using a threshold of 30% identity and 70% query alignment in the AAI calculator (http://enve-omics.ce.gatech.edu/aai/index). By comparison, the AAI values that we obtained from the AAI calculator were very similar to the ANI values from the ANI calculator (http://enve-omics.ce.gatech.edu/ani/), which uses the method described in reference 66 (data not shown). Because some of the available *Francisella* genomes were of draft quality (63), only representative genomes from this set, containing the fewest contigs and hopefully close to a full set of CDS, were included in our analyses. These genomes included FSC454, which had 85 contigs, FSC774 with 194 contigs, and FSC845 with 70 contigs.

Metabolic pathways were identified and compared among the genomes using BLAST analysis and the Pathway Tools (94). Comparison of *Francisella* pathogenicity islands (FPIs) was performed using the IMG system (95) to compare gene neighborhoods in the publicly available genomes and BLAST (92) analysis to identify and compare FPI genes in newly sequenced genomes. Culture phenotypes of the four new species were conducted using methods described in reference 27.

Accession number(s). The annotated whole genomes of the four new species sequenced in this study are now publicly available through NCBI (GenBank), and their accession numbers are listed in Table 1.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/AEM.02589-16.

TEXT S1, PDF file, 1.0 MB.
DATA SET S1, XLSX file, 0.06 MB.

REFERENCES

1. Konstantinidis KT, Ramette A, Tiedje JM. 2006. The bacterial species definition in the genomic era. Philos Trans R Soc Lond B Biol Sci 361(1474):1929–1940. https://doi.org/10.1098/rstb.2006.1920.

2. Konstantinidis KT, Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci U S A 102:2567–2572. https://doi.org/10.1073/pnas.0409727102.

3. Wayne LG, Brenner DJ, Colwell RR, Grimont PAD, Kandler O, Krichevsky MI, Moore LH, Moore WEC, Murray RGE, Stackebrandt E, Starr MP, Truper HG. 1987. Report of the Ad Hoc Committee on Reconciliation of Approaches to Bacterial Systematics. Int J Syst Bacteriol 37:463–464. https://doi.org/10.1099/00207713-37-4-463.

4. Chan JZ, Halachev MR, Loman NJ, Constantinidou C, Pallen MJ. 2012. Defining bacterial species in the genomic era: insights from the genus Acinetobacter. BMC Microbiol 12:392. https://doi.org/10.1186/1471-2180-12-302.

5. Doolittle WF, Papke RT. 2006. Genomics and the bacterial species problem. Genome Biol 7:116. https://doi.org/10.1186/gb-2006-7-9-116.

6. Riley MA, Lizotte-Waniewski M. 2009. Population genomics and the bacterial species concept. Methods Mol Biol 532:367–377. https://doi.org/10.1007/978-1-60327-853-9_21.

7. Staley JT. 2006. The bacterial species dilemma and the genomic-phylogenetic species concept. Philos Trans R Soc Lond B Biol Sci 361:899–909. https://doi.org/10.1098/rstb.2006.1914.

8. Staley JT. 2009. The phylogenomic species concept for Bacteria and Archaea. Microbe 4:362–365.

9. Vos M. 2011. A species concept for bacteria based on adaptive divergence. Trends Microbiol 19:1–7. https://doi.org/10.1016/j.tim.2010.10.003.

10. Reference deleted.

11. Pedati C, House J, Hancock-Allen J, Colton L, Bryan K, Ortbahn D, Kightlinger L, Kugeler K, Petersen J, Mead P, Safranek T, Buss B. 2015. Notes from the field: increase in human cases of tularemia—Colorado, Nebraska, South Dakota, and Wyoming, January-September 2015. MMWR Morb Mortal Wkly Rep 64:1317–1318. https://doi.org/10.15585/mmwr.mm6447a4.

12. Keim P, Johansson A, Wagner DM. 2007. Molecular epidemiology, evolution, and ecology of Francisella. Ann N Y Acad Sci 1105:30–66. https://doi.org/10.1196/annals.1409.011.

13. Kingry LC, Petersen JM. 2014. Comparative review of Francisella tularensis and Francisella novicida. Front Cell Infect Microbiol 4:35. https://doi.org/10.3389/fcimb.2014.00035.

14. Oyston PC. 2008. Francisella tularensis: unravelling the secrets of an intracellular pathogen. J Med Microbiol 57:921–930. https://doi.org/10.1099/jmm.0.2008/000653-0.

15. Antwerpen MH, Schacht E, Kaysser P, Splettstoesser WD. 2013. Complete genome sequence of a Francisella tularensis subsp. holarctica strain from Germany causing lethal infection in common marmosets. Genome Announc 1(1):e00135-12. https://doi.org/10.1128/genomeA.00135-12.

16. Barabote RD, Xie G, Brettin TS, Hinrichs SH, Fey PD, Jay JJ, Engle JL, Godbole SD, Noronha JM, Scheuermann RH, Zhou LW, Lion C, Dempsey MP. 2009. Complete genome sequence of Francisella tularensis subspecies holarctica FTNF002-00. PLoS One 4(9):e7041. https://doi.org/10.1371/journal.pone.0007041.

17. Beckstrom-Sternberg SM, Auerbach RK, Godbole S, Pearson JV, Beckstrom-Sternberg JS, Deng ZM, Munk C, Kubota K, Zhou Y, Bruce D, Noronha J, Scheuermann RH, Wang AH, Wei XY, Wang JJ, Hao J, Wagner DM, Brettin TS, Brown N, Gilna P, Keim PS. 2007. Complete genomic characterization of a pathogenic A.II strain of Francisella tularensis subspecies tularensis. PLoS One 2(9):e947. https://doi.org/10.1371/journal.pone.0000947.

18. Chaudhuri RR, Ren CP, Desmond L, Vincent GA, Silman NJ, Brehm JK, Elmore MJ, Hudson MJ, Forsman M, Isherwood KE, Gurycová D, Minton NP, Titball RW, Pallen MJ, Vipond R. 2007. Genome sequencing shows that European isolates of Francisella tularensis subspecies tularensis are almost identical to US laboratory strain Schu S4. PLoS One 2:e352. https://doi.org/10.1371/journal.pone.0000352.

19. Green M, Choules G, Rogers D, Titball RW. 2005. Efficacy of the live attenuated Francisella tularensis vaccine (LVS) in a murine model of disease. Vaccine 23:2680–2686. https://doi.org/10.1016/j.vaccine.2004.03.071.

20. Larsson P, Elfsmark D, Svensson K, Wikstrom P, Forsman M, Brettin T, Keim P, Johansson A. 2009. Molecular evolutionary consequences of niche restriction in Francisella tularensis, a facultative intracellular pathogen. PLoS Pathog 5(6):e1000472. https://doi.org/10.1371/journal.ppat.1000472.

21. Larsson P, Oyston PCF, Chain P, Chu MC, Duffield M, Fuxelius HH, Garcia E, Halltorp G, Johansson D, Isherwood KE, Karp PD, Larsson E, Liu Y, Michell S, Prior J, Prior R, Malfatti S, Sjostedt A, Svensson K, Thompson N, Vergez L, Wagg JK, Wren BW, Lindler LE, Andersson SGE, Forsman M, Titball RW. 2005. The complete genome sequence of Francisella tularensis, the causative agent of tularemia. Nat Genet 37:153–159. https://doi.org/10.1038/ng1499.

22. Puiu D, Salzberg SL. 2008. Re-assembly of the genome of Francisella tularensis subsp. holarctica OSU18. PLoS One 3:e3427. https://doi.org/10.1371/journal.pone.0003427.

23. Svensson K, Sjödin A, Byström M, Granberg M, Brittnacher MJ, Rohmer L, Jacobs MA, Sims-Day EH, Levy R, Zhou Y, Hayden HS, Lim R, Chang J, Guenthener D, Kang A, Haugen E, Gillett W, Kaul R, Forsman M, Larsson P, Johansson A. 2012. Genome sequence of Francisella tularensis subspecies holarctica strain FSC200, isolated from a child with tularemia. J Bacteriol 194:6965–6966. https://doi.org/10.1128/JB.01040-12.

24. Johansson A, Forsman M, Sjostedt A. 2004. The development of tools for diagnosis of tularemia and typing of Francisella tularensis. Apmis 112:898–907. https://doi.org/10.1111/j.1600-0463.2004.apm11211-1212.x.

25. Hollis DG, Weaver RE, Steigerwalt AG, Wenger JD, Moss CW, Brenner DJ. 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.

26. Barns SM, Grow CC, Okinaka RT, Keim P, Kuske CR. 2005. Detection of diverse new Francisella-like bacteria in environmental samples. Appl Environ Microbiol 71:5494–5500. https://doi.org/10.1128/AEM.71.9.5494-5500.2005.

27. Kugeler KJ, Mead PS, McGowan KL, Burnham JM, Hogarty MD, Ruchelli E, Pollard K, Husband B, Conley C, Rivera T, Kelesidis T, Lee WM, Mabey W, Winchell JM, Stang HL, Staples JE, Chalcraft LJ, Petersen JM. 2008. Isolation and characterization of a novel Francisella sp. from human cerebrospinal fluid and blood. J Clin Microbiol 46:2428–2431. https://doi.org/10.1128/JCM.00698-08.

28. Kuske CR, Barns SM, Grow CC, Merrill L, Dunbar J. 2006. Environmental survey for four pathogenic bacteria and closely related species using phylogenetic and functional genes. J Forensic Sci 51:548–558. https://doi.org/10.1111/j.1556-4029.2006.00131.x.

29. Huber B, Escudero R, Busse HJ, Seibold E, Scholz HC, Anda P, Kampfer P, Splettstoesser WD. 2010. Description of Francisella hispaniensis sp. nov., isolated from human blood, reclassification of Francisella novicida (Larson et al. 1955) Olsufiev et al. 1959. as Francisella tularensis subsp. novicida comb. nov. and emended description of the genus Francisella. Int J Syst Evol Microbiol 60:1887–1896. https://doi.org/10.1099/ijs.0.015941-0.

30. Qu PH, Chen SY, Scholz HC, Busse HJ, Gu Q, Kämpfer P, Foster JT, Glaeser SP, Chen C, Yang ZC. 2013. Francisella guangzhouensis sp. nov., isolated from air-conditioning systems. Int J Syst Evol Microbiol 63:3628–3635. https://doi.org/10.1099/ijs.0.049916-0.

31. Respicio-Kingry LB, Byrd L, Allison A, Brett M, Scott-Waldron C, Galliher K, Hannah P, Mead P, Petersen JM. 2013. Cutaneous infection caused by a novel Francisella sp. J Clin Microbiol 51:3456–3460. https://doi.org/10.1128/JCM.01105-13.

32. Siddaramappa S, Challacombe JF, Petersen JM, Pillai S, Hogg G, Kuske CR. 2011. Common ancestry and novel genetic traits of Francisella novicida-like isolates from North America and Australia as revealed by comparative genomic analyses. Appl Environ Microbiol 77:5110–5122. https://doi.org/10.1128/AEM.00337-11.

33. Siddaramappa S, Challacombe JF, Petersen JM, Pillai S, Kuske CR. 2012. Genetic diversity within the genus Francisella as revealed by comparative analyses of the genomes of two North American isolates from environmental sources. BMC Genomics 13:422. https://doi.org/10.1186/1471-2164-13-422.

34. Colquhoun DJ, Duodu S. 2011. Francisella infections in farmed and wild aquatic organisms. Vet Res 42:47. https://doi.org/10.1186/1297-9716-42-47.

35. Brevik OJ, Ottem KF, Kamaishi T, Watanabe K, Nylund A. 2011. Franci-

sella halioticida sp. nov., a pathogen of farmed giant abalone (Haliotis gigantea) in Japan. J Appl Microbiol 111:1044–1056. https://doi.org/10.1111/j.1365-2672.2011.05133.x.

36. Ottem KF, Nylund A, Isaksen TE, Karlsbakk E, Bergh Ø. 2008. Occurrence of *Francisella piscicida* in farmed and wild Atlantic cod, Gadus morhua L., in Norway. J Fish Dis 31:525–534. https://doi.org/10.1111/j.1365-2761.2008.00930.x.

37. Larson MA, Nalbantoglu U, Sayood K, Zentz EB, Cer RZ, Iwen PC, Francesconi SC, Bishop-Lilly KA, Mokashi VP, Sjöstedt A, Hinrichs SH. 2016. Reclassification of *Wolbachia persica* as *Francisella persica* comb. nov. and emended description of the family Francisellaceae. Int J Syst Evol Microbiol 66:1200–1205. https://doi.org/10.1099/ijsem.0.000855.

38. Petersen JM, Carlson J, Yockey B, Pillai S, Kuske C, Garbalena G, Pottumarthy S, Chalcraft L. 2009. Direct isolation of *Francisella* spp. from environmental samples. Lett Appl Microbiol 48:663–667. https://doi.org/10.1111/j.1472-765X.2009.02589.x..

39. Qu P, Deng X, Zhang J, Chen J, Zhang J, Zhang Q, Xiao Y, Chen S. 2009. Identification and characterization of the *Francisella* sp. strain 08HL01032 isolated in air condition systems. Wei Sheng Wu Xue Bao 49:1003–1010. (In Chinese.)

40. Qu PH, Li Y, Salam N, Chen SY, Liu L, Gu Q, Fang BZ, Xiao M, Li M, Chen C, Li WJ. 19 August 2016. Allofrancisella inopinata gen. nov., sp. nov. and Allofrancisella frigidaquae sp. nov., isolated from water-cooling systems and transfer of Francisella guangzhouensis Qu et al. 2013 to the new genus as Allofrancisella guangzhouensis comb. nov. Int J Syst Evol Microbiol https://doi.org/10.1099/ijsem.0.001437.

41. Rydzewski K, Schulz T, Brzuszkiewicz E, Holland G, Lück C, Fleischer J, Grunow R, Heuner K. 2014. Genome sequence and phenotypic analysis of a first German *Francisella* sp. isolate (W12-1067) not belonging to the species Francisella tularensis. BMC Microbiol 14:169. https://doi.org/10.1186/1471-2180-14-169.

42. Whipp MJ, Davis JM, Lum G, de Boer J, Zhou Y, Bearden SW, Petersen JM, Chu MC, Hogg G. 2003. Characterization of a novicida-like subspecies of *Francisella tularensis* isolated in Australia. J Med Microbiol 52:839–842. https://doi.org/10.1099/jmm.0.05245-0.

43. Brett M, Doppalapudi A, Respicio-Kingry LB, Myers D, Husband B, Pollard K, Mead P, Petersen JM, Whitener CJ. 2012. *Francisella novicida* bacteremia after a near-drowning accident. J Clin Microbiol 50:2826–2829. https://doi.org/10.1128/JCM.00995-12.

44. Clarridge JE, Raich TJ, Sjosted A, Sandstrom G, Darouiche RO, Shawar RM, Georghiou PR, Osting C, Vo L. 1996. Characterization of two unusual clinically significant Francisella strains. J Clin Microbiol 34:1995–2000.

45. Birkbeck TH, Feist SW, Verner-Jeffreys DW. 2011. Francisella infections in fish and shellfish. J Fish Dis 34:173–187. https://doi.org/10.1111/j.1365-2761.2010.01226.x.

46. McDermott C, Palmeiro B. 2013. Selected emerging infectious diseases of ornamental fish. Vet Clin North Am Exot Anim Pract 16:261–282. https://doi.org/10.1016/j.cvex.2013.01.006.

47. Petersen JM, Mead PS, Schriefer ME. 2009. *Francisella tularensis*: an arthropod-borne pathogen. Vet Res 40:7. https://doi.org/10.1051/vetres:2008045.

48. Goethert HK, Shani I, Telford SR, III. 2004. Genotypic diversity of *Francisella tularensis* infecting Dermacentor variabilis ticks on Martha's Vineyard, Massachusetts. J Clin Microbiol 42:4968–4973. https://doi.org/10.1128/JCM.42.11.4968-4973.2004.

49. Zhang F, Liu W, Wu XM, Xin ZT, Zhao QM, Yang H, Cao WC. 2008. Detection of *Francisella tularensis* in ticks and identification of their genotypes using multiple-locus variable-number tandem repeat analysis. BMC Microbiol 8:152. https://doi.org/10.1186/1471-2180-8-152.

50. Gehringer H, Schacht E, Maylaender N, Zeman E, Kaysser P, Oehme R, Pluta S, Splettstoesser WD. 2013. Presence of an emerging subclone of *Francisella tularensis* holarctica in Ixodes ricinus ticks from southwestern Germany. Ticks Tick Borne Dis 4:93–100. https://doi.org/10.1016/j.ttbdis.2012.09.001.

51. Kreizinger Z, Hornok S, Dán A, Hresko S, Makrai L, Magyar T, Bhide M, Erdélyi K, Hofmann-Lehmann R, Gyuranecz M. 2013. Prevalence of *Francisella tularensis* and *Francisella*-like endosymbionts in the tick population of Hungary and the genetic variability of *Francisella*-like agents. Vector Borne Zoonotic Dis 13:160–163. https://doi.org/10.1089/vbz.2012.1065.

52. Sréter-Lancz Z, Széll Z, Sréter T, Márialigeti K. 2009. Detection of a novel *Francisella* in Dermacentor reticulatus: a need for careful evaluation of PCR-based identification of Francisella tularensis in Eurasian ticks. Vector Borne Zoonotic Dis 9:123–126. https://doi.org/10.1089/vbz.2008.0010.

53. Noda H, Munderloh UG, Kurtti TJ. 1997. Endosymbionts of ticks and their relationship to *Wolbachia* spp. and tick-borne pathogens of humans and animals. Appl Environ Microbiol 63:3926–3932.

54. Zhang X, Norris DE, Rasgon JL. 2011. Distribution and molecular characterization of *Wolbachia* endosymbionts and filarial nematodes in Maryland populations of the lone star tick (*Amblyomma americanum*). FEMS Microbiol Ecol 77:50–56. https://doi.org/10.1111/j.1574-6941.2011.01089.x.

55. Larson CL, Wicht W, Jellison WL. 1955. A new organism resembling *P. tularensis* isolated from water. Public Health Rep 70:253–258. https://doi.org/10.2307/4589039.

56. Busse HJ, Huber B, Anda P, Escudero R, Scholz HC, Seibold E, Splettstoesser WD, Kampfer P. 2010. Objections to the transfer of Francisella novicida to the subspecies rank of Francisella tularensis—response to Johansson et al. Int J Syst Evol Microbiol 60(Part 8):1718–1720. https://doi.org/10.1099/00207713-60-8-1718.

57. Johansson A, Celli J, Conlan W, Elkins KL, Forsman M, Keim PS, Larsson P, Manoil C, Nano FE, Petersen JM, Sjöstedt A. 2010. Objections to the transfer of *Francisella novicida* to the subspecies rank of *Francisella tularensis*. Int J Syst Evol Microbiol 60:1717–1718. https://doi.org/10.1099/ijs.0.022830-0.

58. Sjostedt AB. 2005. Genus I, Francisella Dorofe'ev 1947, 176AL, p 200–210. *In* Brenner DJ, Krieg NR, Staley JT, Garrity GM (ed), Bergey's manual of systematic bacteriology, 2nd ed, vol 2. Springer, New York, NY.

59. Siddaramappa S, Challacombe JF, Petersen JM, Pillai S, Kuske CR. 2014. Comparative analyses of a putative *Francisella* conjugative element. Genome 57:137–144. https://doi.org/10.1139/gen-2013-0231.

60. Champion MD, Zeng QD, Nix EB, Nano FE, Keim P, Kodira CD, Borowsky M, Young S, Koehrsen M, Engels R, Pearson M, Howarth C, Larson L, White J, Alvarado L, Forsman M, Bearden SW, Sjostedt A, Titball R, Michell SL, Birren B, Galagan J. 2009. Comparative genomic characterization of Francisella tularensis strains belonging to low and high virulence subspecies. Plos Pathog 5(5):e1000459. https://doi.org/10.1371/journal.ppat.1000459.

61. Rohmer L, Fong C, Abmayr S, Wasnick M, Freeman TJL, Radey M, Guina T, Svensson K, Hayden HS, Jacobs M, Gallagher LA, Manoil C, Ernst RK, Drees B, Buckley D, Haugen E, Bovee D, Zhou Y, Chang J, Levy R, Lim R, Gillett W, Guenthener D, Kang A, Shaffer SA, Taylor G, Chen JZ, Gallis B, D'Argenio DA, Forsman M, Olson MV, Goodlett DR, Kaul R, Miller SI, Brittnacher MJ. 2007. Comparison of *Francisella tularensis* genomes reveals evolutionary events associated with the emergence of human pathogenic strains. Genome Biol 8(6):R102. https://doi.org/10.1186/gb-2007-8-6-r102.

62. Ahlinder J, Ohrman C, Svensson K, Lindgren P, Johansson A, Forsman M, Larsson P, Sjodin A. 2012. Increased knowledge of Francisella genus diversity highlights the benefits of optimised DNA-based assays. BMC Microbiol 12:220. https://doi.org/10.1186/1471-2180-12-220.

63. Sjodin A, Svensson K, Ohrman C, Ahlinder J, Lindgren P, Duodu S, Johansson A, Colquhoun DJ, Larsson P, Forsman M. 2012. Genome characterisation of the genus Francisella reveals insight into similar evolutionary paths in pathogens of mammals and fish. BMC Genomics 13:268. https://doi.org/10.1186/1471-2164-13-268.

64. Wu M, Eisen JA. 2008. A simple, fast, and accurate method of phylogenomic inference. Genome Biol 9:R151. https://doi.org/10.1186/gb-2008-9-10-r151.

65. Konstantinidis KT, Tiedje JM. 2005. Towards a genome-based taxonomy for prokaryotes. J Bacteriol 187:6258–6264. https://doi.org/10.1128/JB.187.18.6258-6264.2005.

66. Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. 2007. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. Int J Syst Evol Microbiol 57:81–91. https://doi.org/10.1099/ijs.0.64483-0.

67. Ludu JS, de Bruin OM, Duplantis BN, Schmerk CL, Chou AY, Elkins KL, Nano FE. 2008. The Francisella pathogenicity island protein PdpD is required for full virulence and associates with homologues of the type VI secretion system. J Bacteriol 190:4584–4595. https://doi.org/10.1128/JB.00198-08.

68. Nano FE, Zhang N, Cowley SC, Klose KE, Cheung KK, Roberts MJ, Ludu JS, Letendre GW, Meierovics AI, Stephens G, Elkins KL. 2004. A Franci-

sella tularensis pathogenicity island required for intramacrophage growth. J Bacteriol 186:6430–6436. https://doi.org/10.1128/JB.186.19.6430-6436.2004.

69. Nano FE, Schmerk C. 2007. The Francisella pathogenicity island. Ann N Y Acad Sci 1105:122–137. https://doi.org/10.1196/annals.1409.000.

70. Gallagher LA, McKevitt M, Ramage ER, Manoil C. 2008. Genetic dissection of the Francisella novicida restriction barrier. J Bacteriol 190:7830–7837. https://doi.org/10.1128/JB.01188-08.

71. Bandara AB, Champion AE, Wang X, Berg G, Apicella MA, McLendon M, Azadi P, Snyder DS, Inzana TJ. 2011. Isolation and mutagenesis of a capsule-like complex (CLC) from Francisella tularensis, and contribution of the CLC to F. tularensis virulence in mice. PLoS One 6:e19003. https://doi.org/10.1371/journal.pone.0019003.

72. Zarrella TM, Singh A, Bitsaktsis C, Rahman T, Sahay B, Feustel PJ, Gosselin EJ, Sellati TJ, Hazlett KR. 2011. Host-adaptation of Francisella tularensis alters the bacterium's surface-carbohydrates to hinder effectors of innate and adaptive immunity. PLoS One 6:e22335. https://doi.org/10.1371/journal.pone.0022335.

73. Thomas RM, Twine SM, Fulton KM, Tessier L, Kilmury SL, Ding W, Harmer N, Michell SL, Oyston PC, Titball RW, Prior JL. 2011. Glycosylation of DsbA in Francisella tularensis subsp. tularensis. J Bacteriol 193:5498–5509. https://doi.org/10.1128/JB.00438-11.

74. Vogler AJ, Birdsell D, Price LB, Bowers JR, Beckstrom-Sternberg SM, Auerbach RK, Beckstrom-Sternberg JS, Johansson A, Clare A, Buchhagen JL, Petersen JM, Pearson T, Vaissaire J, Dempsey MP, Foxall P, Engelthaler DM, Wagner DM, Keim P. 2009. Phylogeography of Francisella tularensis: global expansion of a highly fit clone. J Bacteriol 191:2474–2484. https://doi.org/10.1128/JB.01786-08.

75. Pandya GA, Holmes MH, Petersen JM, Pradhan S, Karamycheva SA, Wolcott MJ, Molins C, Jones M, Schriefer ME, Fleischmann RD, Peterson SN. 2009. Whole genome single nucleotide polymorphism based phylogeny of Francisella tularensis and its application to the development of a strain typing assay. BMC Microbiol 9:213. https://doi.org/10.1186/1471-2180-9-213.

76. Johnson SL, Minogue TD, Daligault HE, Wolcott MJ, Teshima H, Coyne SR, Davenport KW, Jaissle JG, Chain PS. 2015. Finished genome assembly of warm spring isolate Francisella novicida DPG 3A-IS. Genome Announc 3:e01046-15. https://doi.org/10.1128/genomeA.01046-15.

77. Escudero R, Elía M, Sáez-Nieto JA, Menéndez V, Toledo A, Royo G, Rodríguez-Vargas M, Whipp MJ, Gil H, Jado I, Anda P. 2010. A possible novel Francisella genomic species isolated from blood and urine of a patient with severe illness. Clin Microbiol Infect 16:1026–1030. https://doi.org/10.1111/j.1469-0691.2009.03029.x.

78. Jensen WI, Owen CR, Jellison WL. 1969. Yersinia philomiragia sp. n., a new member of the Pasteurella group of bacteria, naturally pathogenic for the muskrat (Ondatra zibethica). J Bacteriol 100:1237–1241.

79. Versage JL, Severin DD, Chu MC, Petersen JM. 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. https://doi.org/10.1128/JCM.41.12.5492-5499.2003.

80. Wenger JD, Hollis DG, Weaver RE, Baker CN, Brown GR, Brenner DJ, Broome CV. 1989. Infection caused by Francisella philomiragia (formerly Yersinia philomiragia). A newly recognized human pathogen. Ann Intern Med 110:888–892.

81. Ottem KF, Nylund A, Karlsbakk E, Friis-Møller A, Krossøy B. 2007. Characterization of Francisella sp., GM2212, the first Francisella isolate from marine fish, Atlantic cod (Gadus morhua). Arch Microbiol 187:343–350. https://doi.org/10.1007/s00203-006-0198-1.

82. Kamaishi T, Fukuda Y, Nishiyama M, Kawakami H, Matsuyama T, Yoshinaga T, Oseko N. 2005. Identification and pathogenicity of intracellular Francisella bacterium in three-line grunt Parapristipoma trilineatum. Fish Pathol 40:67–72. https://doi.org/10.3147/jsfp.40.67.

83. Birkbeck TH, Bordevik M, Frøystad MK, Baklien A. 2007. Identification of Francisella sp. from Atlantic salmon, Salmo salar L, in Chile. J Fish Dis 30:505–507. https://doi.org/10.1111/j.1365-2761.2007.00837.x.

84. Sridhar S, Sharma A, Kongshaug H, Nilsen F, Jonassen I. 2012. Whole genome sequencing of the fish pathogen Francisella noatunensis subsp. orientalis Toba04 gives novel insights into Francisella evolution and pathogenecity. BMC Genomics 13:598. https://doi.org/10.1186/1471-2164-13-598.

85. Molins CR, Carlson JK, Coombs J, Petersen JM. 2009. Identification of Francisella tularensis subsp. tularensis A1 and A2 infections by real-time

polymerase chain reaction. Diagn Microbiol Infect Dis 64:6–12. https://doi.org/10.1016/j.diagmicrobio.2009.01.006.

86. Johnson SL, Daligault HE, Davenport KW, Coyne SR, Frey KG, Koroleva GI, Broomall SM, Bishop-Lilly KA, Bruce DC, Chertkov O, Freitas T, Jaissle J, Ladner J, Rosenzweig CN, Gibbons HS, Palacios GF, Redden CL, Xu Y, Minogue TD, Chain PS. 2015. Genome sequencing of 18 Francisella strains to aid in assay development and testing. Genome Announc 3:e00147-15. https://doi.org/10.1128/genomeA.00147-15.

87. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75. https://doi.org/10.1186/1471-2164-9-75.

88. Thompson JD, Gibson TJ, Higgins DG. 2002. Multiple sequence alignment using ClustalW and ClustalX. Curr Protoc Bioinformatics Chapter 2:Unit 2.3. https://doi.org/10.1002/0471250953.bi0203s00.

89. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197.

90. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797. https://doi.org/10.1093/nar/gkh340.

91. Pattengale ND, Alipour M, Bininda-Emonds OR, Moret BM, Stamatakis A. 2010. How many bootstrap replicates are necessary? J Comput Biol 17:337–354. https://doi.org/10.1089/cmb.2009.0179.

92. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

93. Challacombe JF, Duncan AJ, Brettin TS, Bruce D, Chertkov O, Detter JC, Han CS, Misra M, Richardson P, Tapia R, Thayer N, Xie G, Inzana TJ. 2007. Complete genome sequence of Haemophilus somnus (Histophilus somni) strain 129Pt and comparison to Haemophilus ducreyi 35000HP and Haemophilus influenzae Rd. J Bacteriol 189:1890–1898. https://doi.org/10.1128/JB.01422-06.

94. Karp PD, Paley S, Romero P. 2002. The Pathway Tools software. Bioinformatics 18(Suppl 1):S225–S232. https://doi.org/10.1093/bioinformatics/18.suppl_1.S225.

95. Markowitz VM, Korzeniewski F, Palaniappan K, Szeto E, Werner G, Padki A, Zhao X, Dubchak I, Hugenholtz P, Anderson I, Lykidis A, Mavromatis K, Ivanova N, Kyrpides NC. 2006. The Integrated Microbial Genomes (IMG) system. Nucleic Acids Res 34:D344–D348. https://doi.org/10.1093/nar/gkj024.

96. Nalbantoglu U, Sayood K, Dempsey MP, Iwen PC, Francesconi SC, Barabote RD, Xie G, Brettin TS, Hinrichs SH, Fey PD. 2010. Large direct repeats flank genomic rearrangements between a new clinical isolate of Francisella tularensis subsp. tularensis A1 and Schu S4. PLoS One 5:e9007. https://doi.org/10.1371/journal.pone.0009007.

97. Whitehouse CA, Kesterson KE, Duncan DD, Eshoo MW, Wolcott M. 2012. Identification and characterization of Francisella species from natural warm springs in Utah, USA. Lett Appl Microbiol 54:313–324. https://doi.org/10.1111/j.1472-765X.2012.03214.x.

98. Petrosino JF, Xiang Q, Karpathy SE, Jiang H, Yerrapragada S, Liu Y, Gioia J, Hemphill L, Gonzalez A, Raghavan TM, Uzman A, Fox GE, Highlander S, Reichard M, Morton RJ, Clinkenbeard KD, Weinstock GM. 2006. Chromosome rearrangement and diversification of Francisella tularensis revealed by the type B (OSU18) genome sequence. J Bacteriol 188:6977–6985. https://doi.org/10.1128/JB.00506-06.

99. Mikalsen J, Olsen AB, Tengs T, Colquhoun DJ. 2007. Francisella philomiragia subsp. noatunensis subsp. nov., isolated from farmed Atlantic cod (Gadus morhua L.). Int J Syst Evol Microbiol 57:1960–1965. https://doi.org/10.1099/ijs.0.64765-0.

100. Birdsell DN, Stewart T, Vogler AJ, Lawaczeck E, Diggs A, Sylvester TL, Buchhagen JL, Auerbach RK, Keim P, Wagner DM. 2009. Francisella tularensis subsp. novicida isolated from a human in Arizona. BMC Res Notes 2:223. https://doi.org/10.1186/1756-0500-2-223.

101. Brett ME, Respicio-Kingry LB, Yendell S, Ratard R, Hand J, Balsamo G, Scott-Waldron C, O'Neal C, Kidwell D, Yockey B, Singh P, Carpenter J, Hill V, Petersen JM, Mead P. 2014. Outbreak of Francisella novicida bacteremia among inmates at a Louisiana correctional facility. Clin Infect Dis 59:826–833. https://doi.org/10.1093/cid/ciu430.

102. Zeytun A, Malfatti SA, Vergez LM, Shin M, Garcia E, Chain PS. 2012. Complete genome sequence of Francisella philomiragia ATCC 25017. J Bacteriol 194:3266. https://doi.org/10.1128/JB.00413-12.

103. Weisburg WG, Dobson ME, Samuel JE, Dasch GA, Mallavia LP, Baca O,

Mandelco L, Sechrest JE, Weiss E, Woese CR. 1989. Phylogenetic diversity of the Rickettsiae. J Bacteriol 171:4202–4206. https://doi.org/10.1128/jb.171.8.4202-4206.1989.

104. Suitor ECJ, Weiss E. 1961. Isolation af a Rickettsialike microorganism (*Wolbachia Persica*, N. SP.) from Argas Persicus (Oken). J Infect Dis 108:95–106.

105. Sjödin A, Ohrman C, Bäckman S, Lärkeryd A, Granberg M, Lundmark E, Karlsson E, Nilsson E, Vallesi A, Tellgren-Roth C, Stenberg P, Thelaus J. 2014. Complete genome sequence of *Francisella endociliophora* strain FSC1006, isolated from a laboratory culture of the marine ciliate *Euplotes raikovi*. Genome Announc 2(6):e01227-14. https://doi.org/10.1128/genomeA.01227-14.