Francisella tularensis subsp. *tularensis* Group A.I, United States

Dawn N. Birdsell, Anders Johansson, Caroline Öhrman, Emily Kaufman, Claudia Molins, Talima Pearson, Miklós Gyuranecz, Amber Naumann, Amy J. Vogler, Kerstin Myrtennäs, Pär Larsson, Mats Forsman, Andreas Sjödin, John D. Gillece, James Schupp, Jeannine M. Petersen, Paul Keim, and David M. Wagner

We used whole-genome analysis and subsequent characterization of geographically diverse strains using new genetic signatures to identify distinct subgroups within *Francisella tularensis* subsp. *tularensis* group A.I: A.I.3, A.I.8, and A.I.12. These subgroups exhibit complex phylogeographic patterns within North America. The widest distribution was observed for A.I.12, which suggests an adaptive advantage.

Tularemia, caused by the bacterium *Francisella tularensis*, is a potentially severe disease that often causes unspecific symptoms; because of its low infectious dose and ease of dissemination, *F. tularensis* is considered a category A biothreat agent (*1*). Three subspecies of *F. tularensis* have been identified; *F. tularensis* subsp. *tularensis* (type A) has been identified only in North America. Numerous subtyping schemes have subdivided type A into 2 groups, A.I and A.II (*2*–*8*). Group A.II is found primarily in the western United States (*3*,*4*), whereas group A.I is found throughout the central and eastern regions of the country and sporadically in some western states (*3*,*4*,*9*).

Groups A.I and A.II differ in virulence, as do subgroups within A.I, although clinical signs and symptoms can be similar. Human infections involving A.I strains are associated with a higher fatality rate than that for infections involving A.II strains (*4*,*10*); this finding was experimentally confirmed in mice (*11*). Kugeler et al. (*10*) used pulsed-field gel electrophoresis (PFGE) to identify 2 subgroups within A.I, A1a and A1b; this study found A1b strains were associated with higher death rates and were more often isolated from human tissue types that were associated with severe disease. This difference was also experimentally confirmed in mice (*11*,*12*). However, virulence testing is not often used in clinical settings because it is slow, complicated, and expensive. Thus, molecular approaches that can rapidly assign an unknown strain to one of the recognized groups with known differences in virulence may provide valuable information to clinicians.

Because PFGE lacks the phylogenetic resolution of some other testing methods (*6*), we independently identified genetic subgroups within A.I by conducting whole-genome sequencing (WGS) of 13 A.I strains (Figure 1; Table 1, Appendix, wwwnc.cdc.gov/EID/article/20/5/13-1559-T1.htm). The 13 strains were selected on the basis of assignment to PFGE subgroups A1a or A1b (*10*) and to maximize geographic diversity; the previously sequenced A.I strain Schu S4 (*13*) was also included. WGS data were generated, assembled, and analyzed as described in the online Technical Appendix (wwwnc.cdc.gov/EID/article/20/5/13-1559-Techapp1.pdf).

Our whole-genome phylogeny revealed 3 major subgroups within *F. tularensis* subsp. *tularensis* A.I: A.I.3, A.I.8, and A.I.12 (Figure 1). The names we assigned to these subgroups are consistent with previous phylogenetic nomenclature within *F. tularensis* (*14*). With the exception of 1 strain (ND01-1900) that was not assigned to any of the 3 subgroups, all strains previously assigned to PFGE subgroup A1a belonged to the newly designated A.I.12 subgroup (Figure 1; Table 1). In contrast, strains previously assigned to PFGE subgroup A1b were distributed among all 3 of the new subgroups (Figure 1; Table 1). We concluded that results of characterization of subgroups A1a and A1b by PFGE are not in agreement with findings of a robust whole-genome phylogeny and therefore focused the remainder of our analysis on subgroups identified by using WGS.

We observed several differences among the 3 subgroups in the whole-genome phylogeny (Figure 1). The first split separated the A.I.3 subgroup from the A.I.8 and A.I.12 subgroups; a second split separated the A.I.8 and A.I.12 subgroups. A long branch of 25 single nucleotide polymorphisms (SNPs) led to the A.I.3 subgroup, in which relatedness among the sequenced strains was moderate. A branch of 9 SNPs led to the A.I.8 subgroup, and again, relatedness among the sequenced strains was moderate. The branch leading to subgroup A.I.12 was, by comparison, much longer (37 SNPs), and the sequenced strains were separated only by 3 short branches (1–4 SNPs). This pattern of several short branches without hierarchical structuring is consistent

Author affiliations: Northern Arizona University, Flagstaff, Arizona, USA (D.N. Birdsell, E. Kaufman, T. Pearson, A. Naumann, A.J. Vogler, P. Keim, D.M. Wagner); Umeå University, Umeå, Sweden (A. Johansson); Swedish Defense Research Agency, Umeå (C. Öhrman, K. Myrtennäs, P. Larsson, M. Forsman, A. Sjödin); Centers for Disease Control and Prevention, Fort Collins, Colorado, USA (C. Molins, J.M. Petersen); Center for Agricultural Research, Hungarian Academy of Sciences, Budapest, Hungary (M. Gyuranecz); and Translational Genomics Research Institute, Flagstaff (J. Gillece, J. Schupp, P. Keim)

DOI: http://dx.doi.org/10.3201/eid2005.131559

with a recent radiation, an evolutionary process in response to adaptive change, new ecologic opportunities, or a combination of these factors.

To show more comprehensive phylogenetic patterns, we developed 16 canonical SNP (canSNP) assays as described (online Technical Appendix) and used them to screen 179 *F. tularensis* subsp. *tularensis* A.I strains selected from the collections of the Centers for Disease Control and Prevention (Fort Collins, CO, USA). We selected strains that were representative of all states where A.I infections occur and of all PFGE classification types (Table 1). One limitation of our study is that we did not analyze an equal number of strains from all regions of the country. However, our sample reflects the distribution

Table 2. Melt-MAMA primers targeting canSNPs for new phylogenetic branches in *Francisella tularensis* subsp. *tularensis* A.I in United States*

Subgroup		SchuS4†	SNP state,			Temp,
Major	Minor	position	der/anc‡	Primers, 5′ → 3′§	Con¶	°C#
NA	A.I.7	1005448**	C/T	A: TATTTCAATTTTTGCGATGGTAgGT	0.80	55
				D: ggggcggggcggggcTATTTCAATTTTTGCGATGGTAcTC	0.20	
				C: AAGTATGTTGGCAAGTAAAGTGAGAAGA	0.20	
A.I.12	NA	142781††	C/G	A: GCTTATCGCCGACATTCATCtAC	0.20	60
				D: ggggcggggcggggcgggCTTATCGCCGACATTCATCcAG	0.20	
				C: GGTATGGCAAAAAATACTTATGGTACG	0.20	
A.I.12	A.I.13	1833651‡‡	T/C	A: CTTTCAATCATGTAACCATCATTATTTAaGC	0.80	60
				D: cggggcggggcggggcggggCTTTCAATCATGTAACCATCATTATTTAgGT	0.20	
				C: CTTAATGAACTTGGTGTAATGGGTAGATA	0.20	
A.I.12	A.I.16	273622	T/C	A: AAACTTAAAAAAGAGCAAGAACTTAATGATcTC	0.60	60
				D: ggggcggggcggggcgAAACTTAAAAAAGAGCAAGAACTTAATGATaTT	0.15	
				C: CATCTTCATTAAAAGTCTTATTGTTTAAACGC	0.15	
A.I.12	A.I.15	1210286	A/G	A: TCTTAAAACATCGACACTCTCAACcTG	0.80	60
				D: ggggcggggcggggcGATCTTAAAACATCGACACTCTCAACtTA	0.20	
				C: GTATCATTCAGATCATAATGAAGCAACTATC	0.20	
A.I.12	A.I.14	1296147	T/C	A: ATCATACTGGTTATATTGGCGGTcTC	0.80	60
				D: cggggcggggcggggcggggATCATACTGGTTATATTGGCGGTgTT	0.20	
				C: GATGAGTCGCTATTAGCTTCTCGAAAG	0.20	
A.I.8	NA	1150298	G/A	A: TAGTCAATCTTGGAACTCCAGAtAA	0.75	60
				D: ggggcggggcggggcTAGTCAATCTTGGAACTCCAGAaAG	0.15	
				C: TCTATTACTCTAGGGTCAGATAGAAATTC	0.15	
A.I.8	A.I.9	1453599	C/T	A: GCTGCTGCTAGATTAGCTATgCT	0.15	60
				D: ggggcggggcggggcGCTGCTGCTAGATTAGCTATcCC	0.15	
				C: TCAAGCAATCAACAATAATTTTACTAT	0.15	
A.I.8	A.I.10	797599	T/G	A: GATCAATTGGTGGTGTTcCG	0.80	60
				D: ggggcggggcggggcGTGATCAATTGGTGGTGTTtCT	0.20	
				C: AACGTTTTATCCTCTTGAATATCAACTAT	0.20	
A.I.8	A.I.11	1278606	G/A	A: AAGGAACAAAAAACATCATCATTgCT	0.20	60
				D: ggggcggggcggggcAAAAGGAACAAAAAACATCATCATTaCC	0.20	
				C: TCATACTAACAACGGCTATTCAGGGA	0.20	
A.I.3	NA	1233898	T/G	A: GCTTGACAATATTAGCTTATAAAACTATAgTG	0.15	60
				D: ggggcggggcggggcGCTTGACAATATTAGCTTATAAAACTATAaTT	0.15	
				C: TTTTTTCCATATTTCTGTAAAAAATATACTATTATG	0.15	
A.I.3	A.I.4	830715§§	T/C	A: GTTAAGTCGGTAAGTATCGACAAaTC	0.60	60
				D: ggggcggggcggggcGTTAAGTCGGTAAGTATCGACAAgTT	0.20	
				C: CAAATCTTCTAGTATCTCTTTATCTTCAG	0.20	
A.I.3	A.I.5	113671	G/A	A: cgggcgggcgggcgggGCTTGAGTTTATTTTTTGTTTAATGTgTA	0.20	60
				D: GCTTGAGTTTATTTTTTGTTTAATGTaTG	0.20	
				C: GGACAAAACTGTGGACGTTAAGAA	0.20	
A.I.3	A.I.6	580153	G/A	A: cgggcgggcgggcgggTATAATGGTAACTCATGATCAAGAAcAA	0.20	60
				D: TTATAATGGTAACTCATGATCAAGAAaAG	0.20	
				C: ATCTGTCATGATACCAATTCTTGTCG	0.20	

*Melt-MAMA, melt–mismatch amplification mutation assay; SNP, single nucleotide polymporphism; canSNP, canonical SNP; con, concentration, μmol/L; temp, annealing temperature, °C; NA, not applicable; der, derived SNP state; anc, ancestral SNP state; D, derived allele primer; A, ancestral allele primer; C, common primer.
†Genomic position in reference A.I SchuS4 strain (GenBank accession no. NC_006570).
‡SNP states are listed according to their orientation in the SCHU S4 reference genome (GenBank accession no. AJ749949.2).
§Melt-MAMA primer sequences; primer tails and antepenultimate mismatch bases are in lower case.
¶Final concentration of each primer in Melt-MAMA genotyping assays.
#Assay annealing temperature.
**Assay designed on the reverse complement.
††SNP from (*6*).
‡‡Assay supplemented with 0.025 U of Platinum Taq DNA polymerase (Life Technologies, Invitrogen, Frederick, MD, USA).
§§SNP from (*7*).

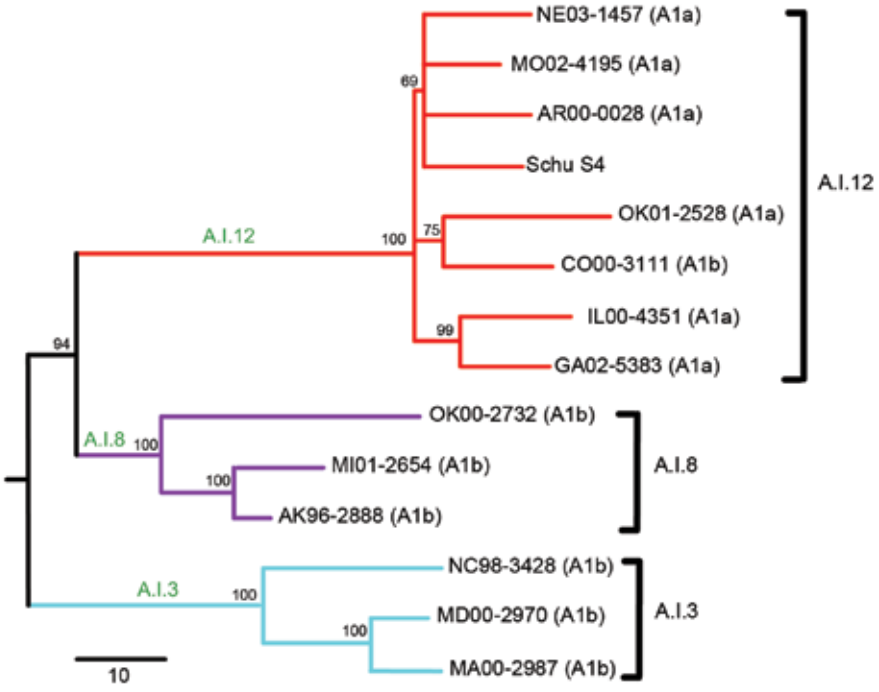

Figure 1. Neighbor-joining tree of 14 *Francisella tularensis* subsp. *tularensis* group A.I strains constructed on the basis of single-nucleotide polymorphisms (SNPs) discovered from whole-genome sequencing. Lines represent major groups within A.I: red, A.I.12; purple, A.I.8; blue, A.I.3. Branch nomenclature for each group is indicated by green text. Bootstrap values for each group and subpopulation are indicated in black font. Pulsed-field gel electrophoresis classifications (A1a and A1b) are indicated for each sequenced strain. A.I strain SchuS4 (GenBank accession no. NC_006570) was included as a reference strain. Scale bar indicates no. SNPs.

of human disease caused by *F. tularensis* subsp. *tularensis* A.I strains: prevalent in the central United States, less common in the eastern United States, and rare in the western United States (*4*). The canSNP assays were based on 12 SNP signatures (Table 2) from the whole-genome phylogeny (Figure 1) and 4 previously described SNP signatures (*6–8*). Using these assays, we assigned the 179 strains to 15 *F. tularensis* subsp. *tularensis* A.I subpopulations, including 8 intervening nodes (Figure 2, panel A). We found 6 subpopulations in the A.I.12 subgroup, 4 in A.I.8, and 4 in A.I.3 (Table 1). To identify broad phylogeographic patterns, we created maps indicating specific states where strains from the 15 subpopulations were isolated (Figure 2, panel B). Within these maps, we created boundaries corresponding to 3 regions within the United States: western, central, and eastern.

Each subgroup exhibited complex yet distinct phylogeographic patterns (Figure 2, panel B). Group A.I.12 strains, assigned to 6 subpopulations (Figure 2, panel A), were isolated throughout the United States: all 6 subpopulations were found in the central region, 3 in the western region, and 5 in the eastern region (Figure 2, panel B, top). Group A.I.8 strains, assigned to 4 subpopulations, were found in the central (3 subpopulations) and western (including Alaska and British Columbia; 3 subpopulations) regions, but only 1 strain was isolated in the eastern region (Figure 2, panel B, middle). For group A.I.3 strains, assigned to 4 subpopulations, distribution differed dramatically from the other subgroups; most strains and all 4 subpopulations occurred in the eastern region and just 1 subpopulation in the central region but none in the western region (Figure 2, panel B, bottom).

Conclusions

The occurrence of the A.I.3 subgroup in the eastern United States could be a recent or ancient event. The subgroup may have been introduced more recently from the central region to a naive niche in the eastern region through importation of rabbits (*Sylvilagus floridanus*) as recently as the 1920s (*3*); before 1937, tularemia was nearly nonexistent in the eastern region (*15*). If the introduction is recent, the current lack of A.I.3 strains in the central United States could be the result of a selective sweep that nearly eliminated this subgroup from its geographic origin. However, most strains and genetic diversity (i.e., subpopulations) within the A.I.3 subgroup are found in the eastern United States, which may reflect a more ancient history in this region involving early introduction and establishment of this subgroup east of the Appalachian Mountains, with only recent spread to the central region.

If we assume that the greatest genetic diversity in a phylogenetic context implies ancient origins, our findings suggest that the central United States is the likely geographic origin of a common ancestor to *F. tularensis* subsp. *tularensis* subgroups A.I.12 and A.I.8 and, perhaps, the A.I group as a whole. The large geographic range of the A.I.12 subgroup and the phylogenetic pattern of a long branch leading to a polytomy with genetic homogeneity point to a possible adaptive advantage for this subgroup.

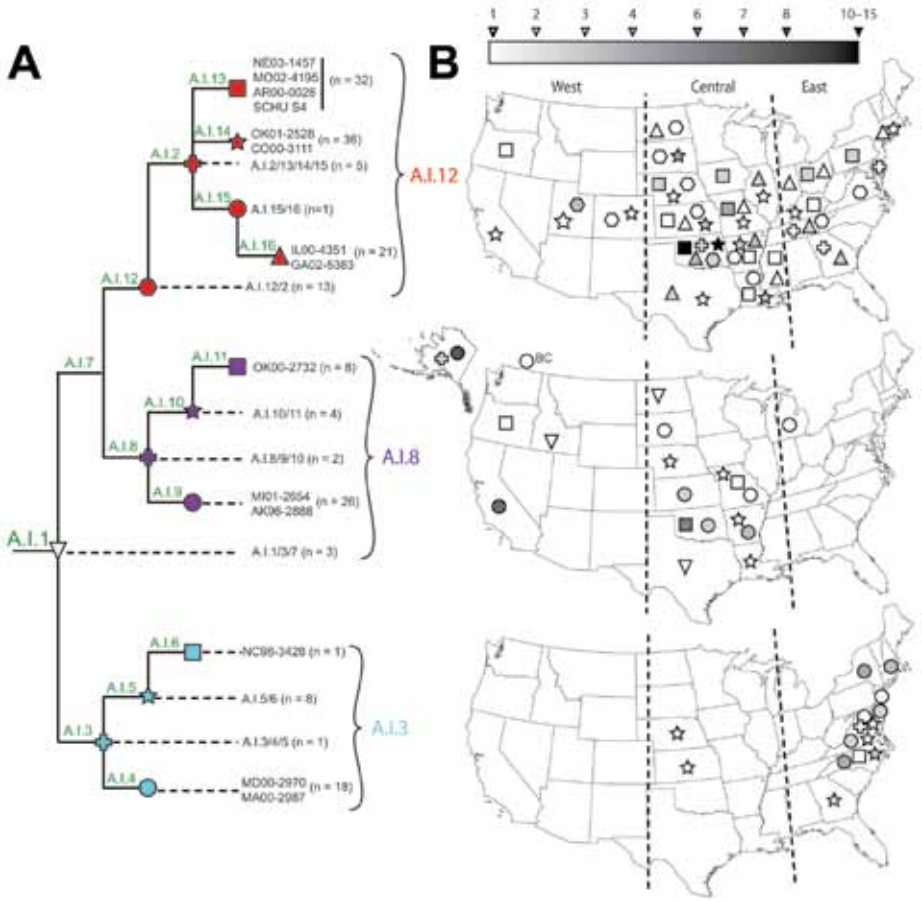

Figure 2. Geographic alignment of 179 geographically diverse *Francisella tularensis* subsp. *tularensis* A.I strains, by subgroup, United States. A) Canonical single-nucleotide polymorphism (canSNP) topology of 15 intervening and terminal subpopulations defined by screening of 16 canSNPs. Colors indicate major subgroups within A.I: red, A.I.12; purple, A.I.8; blue, A.I.3. Subpopulations are indicated by symbols; n values indicate number of strains assigned to each subpopulation. B) Geographic distribution of strains from the 15 subpopulations, shown by corresponding symbols as in panel A and aligned by subgroup (top, A.I.12; middle, A.I.8; bottom, A.I.3). Vertical lines indicate boundaries of the 3 regions: western, central, and eastern. Subgroups are mapped on the basis of geographic origin at the state level. Gradients correspond to number of strains associated with each symbol (i.e., darker symbols indicate a higher number of strains). The basal A.I.1/3/7 subgroup (inverted triangle) cannot be meaningfully assigned to 1 of the 3 main subgroups; thus, this subgroup is arbitrarily represented on the A.I.8 map. BC, British Columbia, Canada.

This advantage may be related to difference in virulence among A.I strains, as suggested by previous testing in mice of 2 A.I.12 strains that exhibited lower virulence than that of 2 A.I.3 strains (*11*). Further research is needed to determine whether the genomic differences that define this subgroup are associated with known *F. tularensis* virulence determinants.

Acknowledgments

We thank Mia Champion for her assistance with initial analyses and Laurel Respicio-Kingry for her assistance with SNP genotyping.

This work was supported in part by the US Department of Homeland Security Science and Technology Directorate through award HSHQDC-10-C-00139, the Swedish Civil Contingencies Agency through TA#014-2010-01, the Laboratory for Molecular Infection Medicine Sweden, and Västerbotten County Council. M.G. was supported by the Lendület program of the Hungarian Academy of Sciences.

Dr Birdsell is a postdoctoral fellow at the Center for Microbial Genetics and Genomics, Northern Arizona University. Her primary research interest is the evolution of *F. tularensis*.

References

1. Rotz LD, Khan AS, Lillibridge SR, Ostroff SM, Hughes JM. Public health assessment of potential biological terrorism agents. Emerg Infect Dis. 2002;8:225–30. http://dx.doi.org/10.3201/eid0802.010164
2. Johansson A, Farlow J, Larsson P, Dukerich M, Chambers E, Byström M, et al. Worldwide genetic relationships among *Francisella tularensis* isolates determined by multiple-locus variable-number tandem repeat analysis. J Bacteriol. 2004;186:5808–18. http://dx.doi.org/10.1128/JB.186.17.5808-5818.2004
3. Farlow J, Wagner DM, Dukerich M, Stanley M, Chu M, Kubota K, et al. *Francisella tularensis* in the United States. Emerg Infect Dis. 2005;11:1835–41. http://dx.doi.org/10.3201/eid1112.050728
4. Staples JE, Kubota KA, Chalcraft LG, Mead PS, Petersen JM. Epidemiologic and molecular analysis of human tularemia, United States, 1964–2004. Emerg Infect Dis. 2006;12:1113–8. http://dx.doi.org/10.3201/eid1207.051504
5. Svensson K, Larsson P, Johansson D, Byström M, Forsman M, Johansson A. Evolution of subspecies of *Francisella tularensis.* J Bacteriol. 2005;187:3903–8. http://dx.doi.org/10.1128/JB.187.11. 3903-3908.2005
6. Pandya GA, Holmes MH, Petersen JM, Pradhan S, Karamycheva SA, Wolcott MJ, et al. Whole genome single nucleotide polymorphism based phylogeny of *Francisella tularensis* and its application to the development of a strain typing assay. BMC Microbiol. 2009;9:213. http://dx.doi.org/10.1186/1471-2180-9-213
7. Svensson K, Granberg M, Karlsson L, Neubauerova V, Forsman M, Johansson A. A real-time PCR array for hierarchical identification of *Francisella* isolates. PLoS ONE. 2009;4:e8360. http://dx.doi.org/10.1371/journal.pone.0008360

8. Vogler AJ, Birdsell D, Price LB, Bowers JR, Beckstrom-Sternberg SM, Auerbach RK, et al. Phylogeography of *Francisella tularensis*: global expansion of a highly fit clone. J Bacteriol. 2009;191:2474–84. http://dx.doi.org/10.1128/JB.01786-08
9. Keim P, Johansson A, Wagner DM. Molecular epidemiology, evolution, and ecology of *Francisella.* Ann N Y Acad Sci. 2007;1105:30–66. http://dx.doi.org/10.1196/annals.1409.011
10. Kugeler KJ, Mead PS, Janusz AM, Staples JE, Kubota KA, Chalcraft LG, et al. Molecular epidemiology of *Francisella tularensis* in the United States. Clin Infect Dis. 2009;48:863–70. http://dx.doi.org/10.1086/597261
11. Molins CR, Delorey MJ, Yockey BM, Young JW, Sheldon SW, Reese SM, et al. Virulence differences among *Francisella tularensis* subsp. *tularensis* clades in mice. PLoS ONE. 2010;5:e10205. http://dx.doi.org/10.1371/journal.pone.0010205
12. Twine SM, Shen H, Kelly JF, Chen W, Sjostedt A, Conlan JW. Virulence comparison in mice of distinct isolates of type A *Francisella tularensis*. Microb Pathog. 2006;40:133–8. http://dx.doi.org/10.1016/j.micpath.2005.12.004
13. Larsson P, Oyston PC, Chain P, Chu MC, Duffield M, Fuxelius HH, et al. The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. Nat Genet. 2005;37:153–9. http://dx.doi.org/10.1038/ng1499
14. Gyuranecz M, Birdsell DN, Splettstoesser W, Seibold E, Beckstrom-Sternberg SM, Makrai L, et al. Phylogeography of *Francisella tularensis* subsp. *holarctica*, Europe. Emerg Infect Dis. 2012;18:290–3. http://dx.doi.org/10.3201/eid1802.111305
15. Ayres JC, Feemster R. Epidemiology of tularemia in Massachusetts with a review of the literature. N Engl J Med. 1948;238:187–94. http://dx.doi.org/10.1056/NEJM194802052380606

Address for correspondence: David M. Wagner, Center for Microbial Genetics and Genomics, Northern Arizona University, Flagstaff, AZ 86011-4073, USA; email: dave.wagner@nau.edu

Article DOI: http://dx.doi.org/10.3201/eid2005.131559

Francisella tularensis subsp. *tularensis* Group A.I, United States

Technical Appendix

Genome sequencing and assembly

Sequencing of the 13 *F. tularensis* genomes was performed using an Illumina GA IIx instrument (Illumina Inc., San Diego, CA, USA) (100-bp pair-end reads) at the Translational Genomics Research Institute (TGen; Flagstaff, Arizona, USA) and the sequences were assembled using Abyss v1.3.3 (*1*). To ensure high-quality data, the genome sequences were filtered to minimize uncertain sequence positions in the phylogenetic analysis. After assembly, sequence reads were re-mapped to their corresponding genome sequence using bowtie2 v2.0.0 (*2*) and subsequent SNP-calling by samtool mpileup (*3*) and VarScan v2.3.2 (*4*) using default parameters except p-value=0.9. Using the SNP information, positions for nucleotides that were supported by <90% of the aligned reads were replaced by the nucleotide symbol "N." After filtering, a multiple genome alignment was calculated by the progressive Mauve algorithm (*5*) using the 13 filtered genome sequences and two public *F. tularensis* genome sequences. These were strain SCHU S4 (acc. AJ749949.2), representing the A.I clade, and strain WY96-3418 (acc. CP000608.1), representing clade A.II and also serving as an outgroup for the phylogenetic analyses. A second filter was then applied to remove all positions within 30-bp of gaps ("-") or uncertain positions ("n") to minimize potential misalignment errors. One SNP found to be incorrectly called due to inaccurate mapping in a repetitive region was manually excluded from the alignment.

Whole genome phylogeny

Based on the filtered and aligned genome data, a Neighbor-Joining tree was inferred using MEGA5 software 2 (Figure 1 in article main text) with gaps/missing data treated as complete deletions.

The naming of the branch leading up to major groups separated by deeply rooted splits was based on SNP nomenclature: A.I.12, A.I.8, and A.I.3 (Figure 1 in article main text). All short read archives were submitted to SRA (NCBI BioProject Accessions: PRJNA187553, PRJNA187555, PRJNA187556, PRJNA187557, PRJNA187558, PRJNA187559, PRJNA187562, PRJNA187563, PRJNA187564, PRJNA187565, PRJNA187567, PRJNA187568, PRJNA187569).

Single Nucleotide Polymorphism (SNP) identification for the development of new canonical SNP assays

SNPs were identified by mapping paired-end reads to a high quality reference genome (*F. tularensis* SCHU S4, acc. AJ749949.2) (*6*) using BWA short read alignment software (*7*) followed by SNP-calling using samtool pileup (*3*) and VarScan v2.2 (min-var-freq 0.9, min-reads 5 and min-coverage 20) (*4*). Finally, SNPs defining the three clades A.I.12, A.I.8, and A.I.3 were confirmed using an in-house Perl script based on their presence in a multiple alignment of de novo assembled genome sequences. From this information 16 canonical SNP (canSNP) assays were created as previously described (*8*).

References

1. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJ, Birol I. ABySS: a parallel assembler for short read sequence data. Genome Res. 2009;19:1117–23. PubMed http://dx.doi.org/10.1101/gr.089532.108

2. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9. PubMed http://dx.doi.org/10.1038/nmeth.1923

3. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25:2078–9. PubMed http://dx.doi.org/10.1093/bioinformatics/btp352

4. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics. 2009;25:2283–5. PubMed http://dx.doi.org/10.1093/bioinformatics/btp373

5. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004;14:1394–403. PubMed http://dx.doi.org/10.1101/gr.2289704

6. Larsson P, Oyston PC, Chain P, Chu MC, Duffield M, Fuxelius HH, et al. The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. Nat Genet. 2005;37:153–9. PubMed http://dx.doi.org/10.1038/ng1499

7. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60. PubMed http://dx.doi.org/10.1093/bioinformatics/btp324

8. Birdsell DN, Pearson T, Price EP, Hornstra HM, Nera RD, Stone N, et al. Melt analysis of mismatch amplification mutation assays (Melt-MAMA): a functional study of a cost-effective SNP genotyping assay in bacterial models. PLoS ONE. 2012;7:e32866. PubMed http://dx.doi.org/10.1371/journal.pone.0032866