

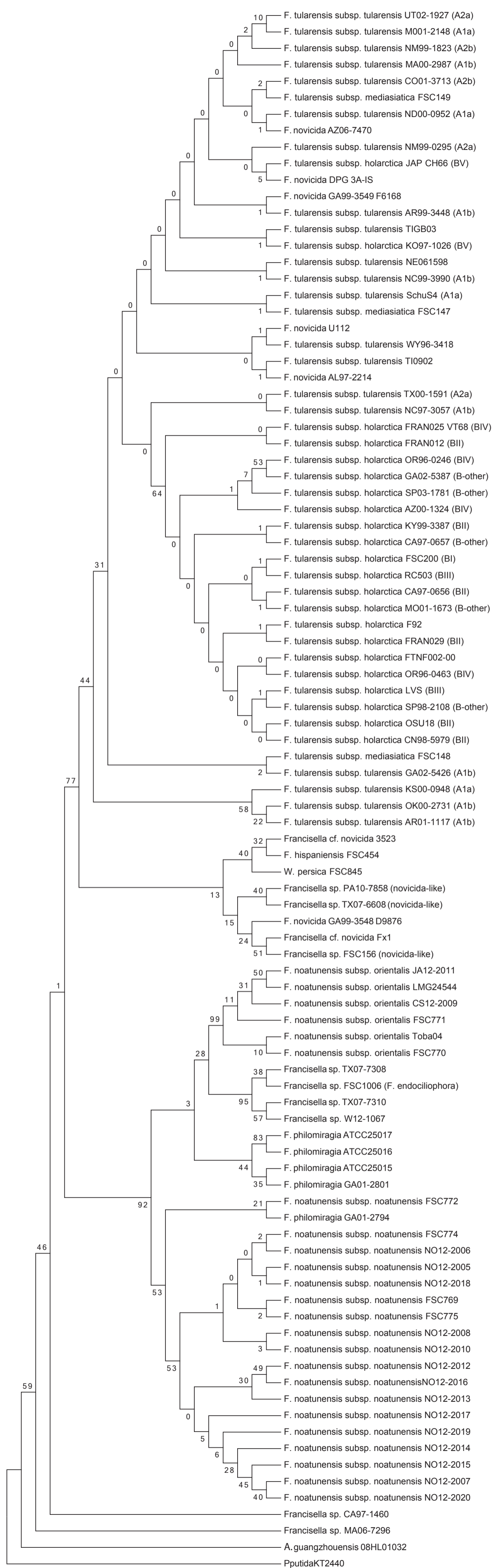
Supplementary Figure Legends

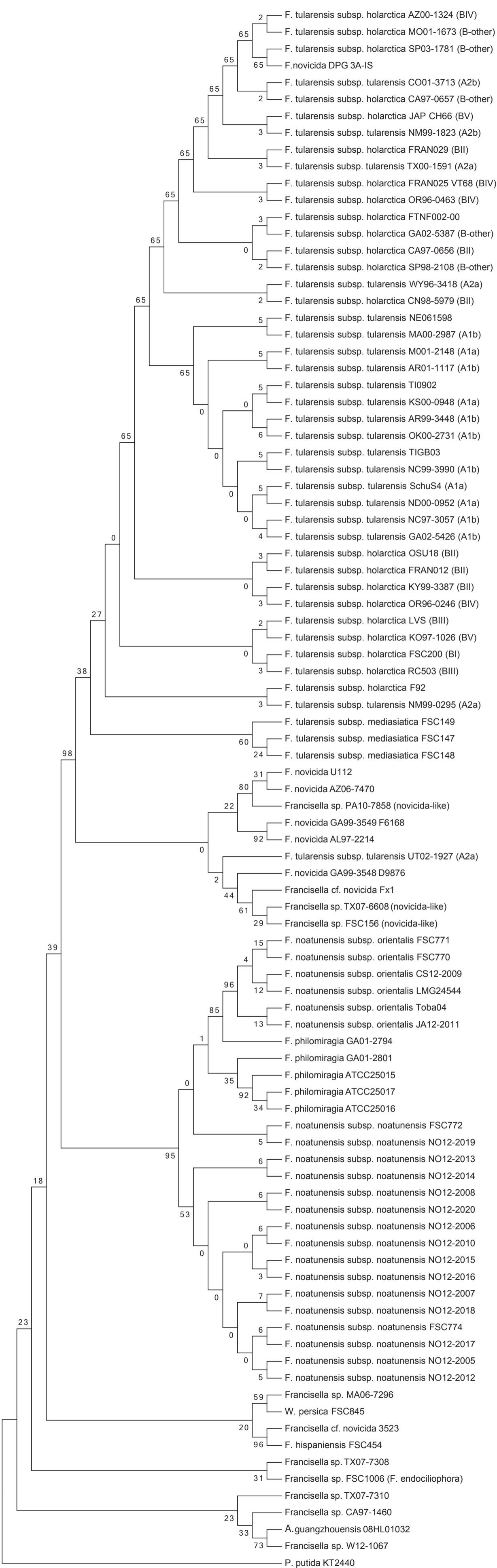
Figure S1. Expanded phylogenetic tree based on 16S ribosomal RNA gene sequences from 94 *Francisella* representatives.

Figure S2. Expanded phylogenetic tree based on *sdhA* gene sequences from 92 *Francisella* representatives.

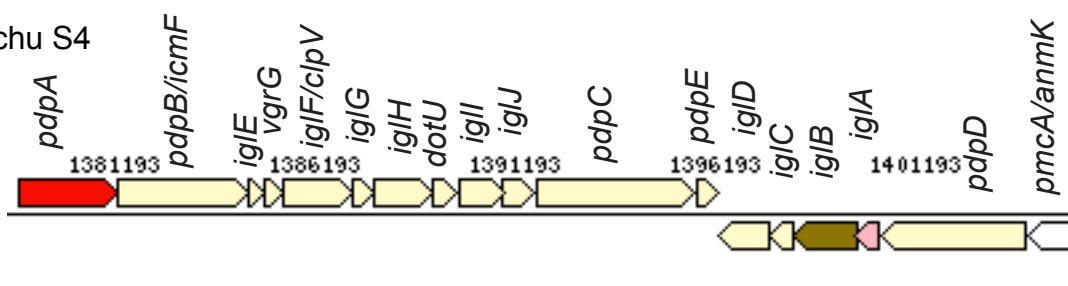
Figure S3. Expanded phylogenetic tree based on 36 protein markers listed in Supplementary Table 1.

Figure S4. Four patterns of gene organization in the *Francisella* Pathogenicity Island (FPI). Pattern 1 is typical of *F. tularensis* subsp. *tularensis* and *F. novicida*. Pattern 2 is found in *F. tularensis* subsp. *holarctica*. Pattern 3 is found in a wide variety of isolates, including *F. philomiragia*, *F. noatunensis*, some *F. novicida*, *F. endociliophora*, and three of the new species (TX07-7310, MA06-7296 and CA97-1460). Pattern 4 is like Pattern 3, but has a truncated *iglG* gene and a small ORF on the complementary strand. Among the genomes that we analyzed, only TX07-7308 and *F. noatunensis* subsp. *orientalis* Toba04 had Pattern 4.



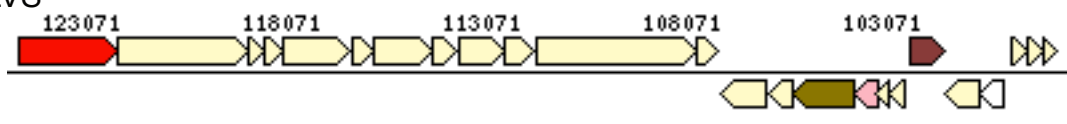


F. tularensis subsp. *tularensis* Schu S4
FTT_1344 - FTT_1361c



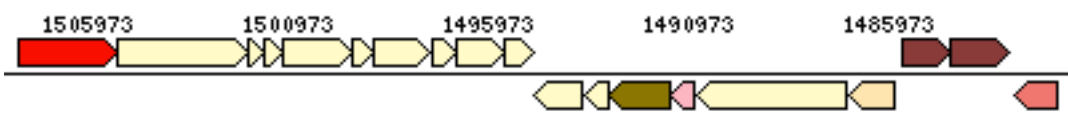
Pattern 1
pmcA is pseudogene
but not in all genomes

F. tularensis subsp. *holarctica* LVS
FTL_0126 - FTL_0107



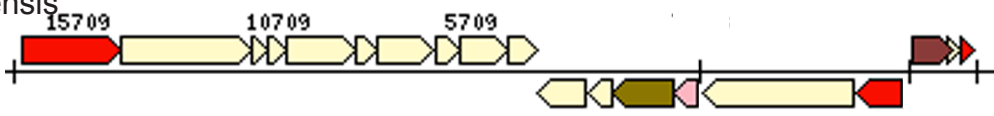
Pattern 2
missing *pmcA*
pdpD is small

F. philomiragia ATCC 25017
Fphi_1377 - Fphi_1359

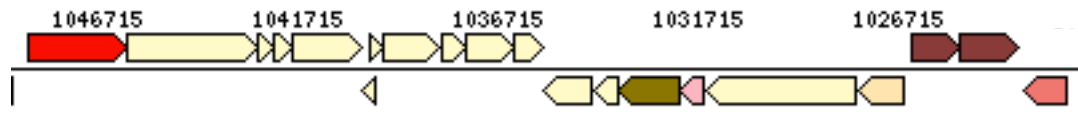


Pattern 3
missing *pdpC*, *pdpE*
flanking regions differ

F. noatunensis subsp. *noatunensis*
representative FPI region

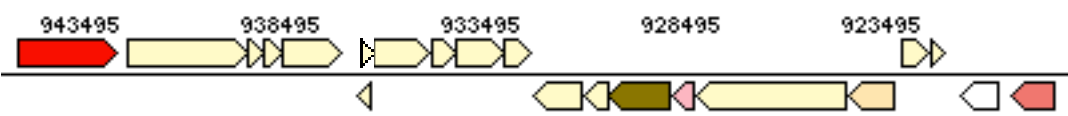


Francisella sp. TX07-7308
F7308_1016 - F7308_0997



Pattern 4
missing *pdpC*, *pdpE*
iglG is truncated
small ORF in reverse
flanking regions differ

F. noatunensis subsp. *orientalis* Toba04
OOM_0944- OOM_0925



Supplementary Table 1. Thirty-six phylogenetic marker genes

Gene	Function
DNA primase (<i>dnaG</i>)	initiation of DNA synthesis
ribosome recycling factor (<i>frr</i>)	recycling ribosomes
translation initiation factor IF-3 (<i>infC</i>)	initiation of protein synthesis
transcription elongation factor (<i>nusA</i>)	formation of stable elongation complexes
phosphoglycerate kinase (<i>pgk</i>)	glycolysis/gluconeogenesis
CTP synthetase (<i>pyrG</i>)	pyrimidine biosynthesis
50S ribosomal protein L1 (<i>rplA</i>)	50S ribosome structure
50S ribosomal protein L2 (<i>rplB</i>)	50S ribosome structure
50S ribosomal protein L3 (<i>rplC</i>)	50S ribosome structure
50S ribosomal protein L4 (<i>rplD</i>)	50S ribosome structure
50S ribosomal protein L5 (<i>rplE</i>)	50S ribosome structure
50S ribosomal protein L6 (<i>rplF</i>)	50S ribosome structure
50S ribosomal protein L11 (<i>rplK</i>)	50S ribosome structure
50S ribosomal protein L7/L12 (<i>rplL</i>)	50S ribosome structure
50S ribosomal protein L14 (<i>rplN</i>)	50S ribosome structure
50S ribosomal protein L16 (<i>rplP</i>)	50S ribosome structure
50S ribosomal protein L19 (<i>rplS</i>)	50S ribosome structure
50S ribosomal protein L20 (<i>rplT</i>)	50S ribosome structure
50S ribosomal protein L27 (<i>rpmA</i>)	50S ribosome structure
DNA-directed RNA polymerase subunit beta (<i>rpoB</i>)	RNA synthesis
30S ribosomal protein S2 (<i>rpsB</i>)	30S ribosome structure
30S ribosomal protein S3, partial (<i>rpsC</i>)	30S ribosome structure
30S ribosomal protein S5 (<i>rpsE</i>)	30S ribosome structure
30S ribosomal protein S9 (<i>rpsI</i>)	30S ribosome structure
30S ribosomal protein S10 (<i>rpsJ</i>)	30S ribosome structure
30S ribosomal protein S11 (<i>rpsK</i>)	30S ribosome structure
30S ribosomal protein S13 (<i>rpsM</i>)	30S ribosome structure
30S ribosomal protein S19 (<i>rpsS</i>)	30S ribosome structure
SsrA-binding protein (<i>smpB</i>)	rescuing stalled ribosomes
elongation factor Ts (<i>tsf</i>)	translation
ATP synthase FOF1 subunit alpha (<i>atpA</i>)	ATP synthesis
DNA topoisomerase IV subunit A (<i>parC</i>)	regulate DNA topology and structure
phosphoglucomutase (<i>pgm</i>)	glycogenesis/glycogenolysis
succinate dehydrogenase, flavoprotein subunit (<i>sdhA</i>)	TCA cycle
triosephosphate isomerase (<i>tpiA</i>)	glycolysis
anthranilate synthase component I (<i>trpE</i>)	Phenylalanine, tyrosine and tryptophan biosynthesis

Thirty-one of these are phylogenetic markers from the AMPHORA (1) database (*dnaG*, *frr*, *infC*, *nusA*, *pgk*, *pyrG*, *rplA*, *rplB*, *rplC*, *rplD*, *rplE*, *rplF*, *rplK*, *rplL*, *rplM*, *rplN*, *rplP*, *rplS*, *rplT*, *rpmA*, *rpoB*, *rpsB*, *rpsC*, *rpsE*, *rpsI*, *rpsJ*, *rpsK*, *rpsM*, *rpsS*, *smpB*, and *tsf*). The additional five markers (*atpA*, *parC*, *pgm*, *sdhA*, and *trpE*) have been used to resolve relationships between *F. novicida* and *F. tularensis* strains (2, 3).

(1) Wu M, Eisen JA. 2008. A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* 9:R151.

(2) 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.

(3) Berrada ZL, Goethert HK, Telford SR. 2008. Isolation of *Francisella tularensis tularensis* subpopulation Ai from Missouri lone star ticks. *Am J Trop Med Hyg* 79:230-230.