Genomic analyses of *Francisella tularensis* strains confirm disease transmission from drinking water sources, Turkey, 2008, 2009 and 2012

A Karadenizli¹, M Forsman (mats.forsman@foi.se)², H Şimşek³, M Taner³, C Öhrman², K Myrtenäs², A Lärkeryd², A Johansson⁴, L Özdemir⁴, A Sjödin²

1. Department of Medical Microbiology, Kocaeli University, Kocaeli, Turkey

2. Division of CBRN Security and Defence, FOI - Swedish Defence Research Agency, Umeå, Sweden

3. Public Health Institution of Turkey, National Tuberculosis Reference Laboratory, Ankara, Turkey

4. The Laboratory for Molecular Infection Medicine Sweden (MIMS), Department of Clinical Microbiology, Umeå University, Umeå, Sweden

Citation style for this article:

Karadenizli A, Forsman M, Şimşek H, Taner M, Öhrman C, Myrtenäs K, Lärkeryd A, Johansson A, Özdemir L, Sjödin A. Genomic analyses of *Francisella tularensis* strains confirm disease transmission from drinking water sources, Turkey, 2008, 2009 and 2012. *Euro Surveill.* 2015;20(21):pii=21136. Available online: <http://www.eurosurveillance.org/ViewArticle.aspx?ArticleId=21136>

Article submitted on 03 September 2014 / published on 28 May 2015

Waterborne epidemics of tularaemia caused by *Francisella tularensis* are increasingly reported in Turkey. We have used whole genome sequencing to investigate if *F. tularensis* isolated from patients could be traced back to drinking water sources. Tonsil swabs from 33 patients diagnosed with oropharyngeal tularaemia in three outbreaks and 140 water specimens were analysed. *F. tularensis* subsp. *holarctica* was confirmed by microagglutination and PCR in 12 patients and five water specimens. Genomic analysis of three pairs of patient and water isolates from outbreaks in Sivas, Çorum, and Kocaeli showed the isolates to belong to two new clusters of the *F. tularensis* B.12 genetic clade. The clusters were defined by 19 and 15 single nucleotide polymorphisms (SNPs) in a multiple alignment based on 507 *F. tularensis* genomes. One synonymous SNP was chosen as a new canonical SNP (canSNP) for each cluster for future use in diagnostic assays. No SNP was identified between the genomes from the patient–water pair of isolates from Kocaeli, one SNP between the pair of isolates from Sivas, whereas the pair from Çorum differed at seven SNPs. These results illustrate the power of whole genome sequencing for tracing *F. tularensis* patient isolates back to their environmental source.

Introduction

Tularaemia is a bacterial zoonosis caused by *Francisella tularensis*, a small Gram-negative coccobacillus. The dose required for human infection depends on the transmission route; less than 10 bacteria may be sufficient for infection via inhalation or the subcutaneous route, while infection via the oral route requires greater than 10⁶ bacteria [1,2]. Several clinical forms of tularaemia (including ulceroglandular/glandular, oculoglandular, oropharyngeal and respiratory) may develop, depending on the infection route. Although ulceroglandular

tularaemia is most frequently reported in the literature, the clinical form most frequently seen in Turkey is oropharyngeal tularaemia [3,4]. Epidemiological surveys have revealed a seasonal pattern of the disease in Turkey, where it is endemic in large areas. Most reported cases occur during the winter (December to March), frequently resulting from ingestion of infected food or contaminated water [3,5].

Two main subspecies of *F. tularensis* are clinically important: the most virulent (subsp. *tularensis*) is found only in North America, while the less virulent (subsp. *holarctica*) occurs in most of Eurasia and North America [6,7]. *F. tularensis* subsp. *holarctica* can be further divided into four major genetic clades (B.4, B.6, B.12 and B.16) based on clade-specific canonical single nucleotide polymorphisms (canSNP) markers [8,9]. The distribution of these genetic clades in the waterborne outbreaks in Turkey seems to be widespread [10].

Four documented outbreaks of tularaemia occurred in Turkey between 1936 and 1953; then, following an epidemiologically silent period, it re-emerged in 1988 and its incidence has strongly increased in the past 20 years [11,12]. From 2005, when tularaemia became a notifiable disease in Turkey, to 2012, several outbreaks and in total 4,773 human infections were reported [13,14]. During that time, the three outbreaks studied here of oropharyngeal tularaemia occurred in the Kocaeli (n=35), Çorum (n=54) and Sivas (n=89) areas. Epidemiological investigations indicated that drinking water was the source of infection, but the genotypes and precise sources remained uncertain. In this study, we investigated possible genetic links between *F. tularensis* subsp. *holarctica* strains isolated from patients and drinking water in the same areas and at the same time.

Methods

Tularaemia diagnostics, isolation of bacteria and DNA preparation from clinical and water samples

The outbreaks in the Çorum, Sivas and Kocaeli provinces (Figure 1) commenced on 21 January 2008 (duration approximately two month), 22 February 2009 (duration approximately three month) and 24 March 2012 (duration approximately 40 days), respectively. All tonsil swab specimens collected from index patients were subjected to culturing, RT-PCR and serology. The samples were cultured for 4 to 10 days on cysteine heart agar supplemented with VCNT inhibitor (6 mg/L vancomycin, 15 mg/L colistin, 10 mg/L trimethoprim and 25 mg/L nystatin, Becton Dickinson, Sparks, US). Later during the outbreaks, laboratory verification of tularaemia was based on microagglutination (*F. tularensis* antigen, Becton Dickinson, Sparks, United States (US)) using patient sera [15]. In total, 33 tonsil swab samples and 221 serum samples from patients were analysed. All samples and cultures were handled in a Class III safety cabinet (Labconco, US).

In addition, early during the outbreaks, 51, 34 and 55 specimens of water (0.3 to 1.5 L) were collected from unchlorinated drinking water sources or village fountains in the Çorum, Sivas and Kocaeli areas, respectively. They were transported at +4°C to Kocaeli University, filtered through cellulose acetate membranes (pore size, 0.22 µm), and filters were placed on cysteine heart agar supplemented with VCNT inhibitor for culturing. Water filtration was carried out in a Class II safety hood and cultivation was performed in a separate laboratory under Class III biological safety conditions.

All analyses, including serology, culturing and RT-PCR of clinical samples and culturing of filtered water, were performed during the outbreaks and without delay upon arrival of the samples at the laboratory.

PCR amplification, detection and identification of *Francisella tularensis* strains

For identification of colonies from clinical tonsil-swabs, DNA was extracted from the colonies and PCR using primer and probe sets targeting *ISFtu2* was done as previously described [13]. Negative and positive controls (the latter consisting of 10-fold dilutions of *F. tularensis* subsp. *holarctica* LVS; NCTC 10857) were amplified in parallel [13]. Colonies cultivated from the filters were identified as *F. tularensis* by agglutination test using specific antibody (*F. tularensis* Antisera, Becton Dickinson, Sparks, US). A PCR assay targeting *RD1* was used to identify subspecies of *F. tularensis* isolates from both the tonsil swab and water samples [16].

Whole genome sequencing

Genome sequencing on DNA extracted from the three pairs of patient and water isolates was performed using

Nextera XT DNA preparation kit (Illumina, San Diego, US) and 150 bp paired-end libraries on an Illumina MiSeq instrument. The reads were mapped against the *F. tularensis* subsp. *tularensis* FSC200 [17] genome using Bowtie2 [18] to evaluate the fraction of *Francisella* in the samples. Samples containing a mixture of other bacteria were filtered using MIRA v4.0 mirabait [19]. The filtered reads were subsequently assembled de novo using ABySS [20], and the overall genome coverage was 130-fold (95 contigs). GenBank accession number for the sequences are: FDC200:JPLL000000000, FDC201:JPMT000000000, FDC202:JPPM000000000, FDC203:JPMU000000000, FDC204:JPSV000000000 and FDC205:JPSW000000000.

In silico screening of *Francisella* strains

Genome sequences were initially screened *in silico* by CanSNPer software [21] using published canSNP markers. The markers B.4, B.6, B.10, B.11, B.12, B.13 [22], B.20, B.21 [9], B.26 [23] and B.33 [24] were needed to describe the genotypes in a schematic (canonical SNP) phylogenetic tree. Multiple genome alignment with 507 archived *F. tularensis* strains were used to identify unique SNPs and develop new specific canSNPs for the two clusters detected (B.67 and B.68, see Results), as previously described [25]. The archival genomes represent the available genetic and geographical diversity within *F. tularensis* subspecies *holarctica* from Asia, North America, and Europe.

Genome alignment and phylogenetic tree

A multiple genome alignment of the six *Francisella* genomes obtained in the outbreaks and the 74 most closely related archived *Francisella* genomes was generated by concatenation of a number of pairwise alignments where each strain was aligned against the reference strain FSC200 [17] using progressive MAUVE [26]. The neighbour-joining phylogenetic tree with complete deletion was constructed in MEGA6 [27].

Results

Epidemiology

A total of 178 patients were diagnosed with laboratory-verified tularaemia in the three outbreaks (Figure 1). Pharyngotonsillitis and cervical lymphadenopathy were the most frequent clinical findings in these patients. Epidemiological investigations indicated that drinking water was the likely source of infection. Accordingly, *F. tularensis* subsp. *holarctica* was successfully isolated from cultures of throat swabs collected from 12 patients among 33 index cases investigated. In the 140 drinking water samples, growth of *F. tularensis* subsp. *holarctica* colonies was identified in five samples. One sample pair (one patient and one water isolate) for each of the three outbreaks were selected for phylogenetic analysis.

FIGURE 1

Location of the focal waterborne tularaemia outbreaks in the provinces of Kocaeli, Çorum and Sivas, Turkey, 2008, 2009, and 2012 (n =178 cases)

The number of human tularaemia cases (n) and date of initial case are indicated.

Whole genome sequencing

Six strains were sequenced in total, obtained from one sample pair (one patient and one *F. tularensis*-contaminated water sample) for each of the three outbreaks. Whole genome multiple alignment and CanSNPer analysis showed that all of them grouped in the B.12 clade [8] (Figure 2A). In comparative analyses with whole genome sequences of *F. tularensis* strains of global origin, an isolate from the tonsil of the patient in Kocaeli (FDC204) and the corresponding water isolate from the contaminated water source (FDC205) formed a separate new cluster with identical genome sequences (Figure 2B). The strain most closely related to this cluster was a Swedish strain isolated in 2003 (FSC374), differing at 20 SNPs from the FDC204–5 group (Figure 2B). Other close relatives have been isolated 1967 in Slovakia and 2009 in Hungary (Figure 2B). The other two patient–water pairs, from Sivas and Çorum, were assigned to another new cluster separated by 26 SNPs from its most closely related strain (FSC930, isolated in 1961 in Bulgaria). The water isolates from Sivas and Çorum were identical. However, the patient isolate from Sivas (FDC201) differed from the corresponding water isolate (FDC200) at one SNP at the whole genome level, while the patient isolate from Çorum (FDC203) differed at seven SNPs from the Çorum water isolate (FDC202).

Identification of unique canonical SNPs (canSNPs)

Nineteen and 15 SNPs, respectively, uniquely identified the FDC200–3 and FDC204–5 clusters in the multiple alignment with 507 archived *F. tularensis* genomes.

One synonymous SNP was chosen as a new canSNP for each cluster (B.67 and B.68, respectively). The genomic positions of the B.67 and B.68 SNPs, relative to the *F. tularensis* subsp. *tularensis* SCHU S4 genome (GenBank ID AJ749949.2) [28], were 975,050 and 857,235, respectively. The ancestral base for both B.67 and B.68 was C, and the derived base was A and T, respectively. It was not possible to resolve the two patient–water pairs in the FDC 200–3 cluster, representing strains associated with the outbreaks in Çorum and Sivas, because the sequenced water isolates from these areas (FDC202 and FDC200, respectively) were identical at all aligned positions.

Discussion

Oropharyngeal tularaemia is a significant waterborne disease in Turkey, where more cases were recorded in 2011 than in all European Union countries combined [5]. Better epidemiological tools can help identify the most important transmission chains resulting in human infections in the region and may ultimately aid in preventing the disease.

Because *F. tularensis* exhibits very little genetic variability, high-resolution analytical methods such as whole genome sequencing are appropriate for genetic typing in epidemiological investigations of tularaemia outbreaks and for source-tracing [29,30]. Initial canSNP analysis of the six included isolates revealed that they all belonged to the *F. tularensis* genetic clade B.12, which dominates in Europe between Scandinavia and the Black Sea [7,8,23–25]. The higher resolution

Phylogenetic trees showing genetic relationships of the sequenced *Francisella tularensis* genomes, 2008, 2009 and 2012 (n=6)

was the water sampled in Çorum, the water must have been contaminated by multiple genetically distinct *F. tularensis* possibly from different sources, e.g. several dead rodents.

Isolating *F. tularensis* from environmental specimens by cultivation is very difficult, mainly because it grows slowly and fastidiously, hence overgrowth of environmental background flora is a major problem. In this study the bacterium was isolated by filtering samples of cold spring water collected during the winter through cellulose acetate membranes before cultivation on selective media. Very low background growth was recorded, probably because the sampled water was pure spring water, which is oligotrophic and thus supports low background bacterial populations.

The isolates analysed in this study were obtained from three locations, 280 to 798 km apart, in central Anatolia and north-western Turkey (Figure 1). No tularaemia outbreaks were reported in these locations before 2004 [1,8]. Ingestion of contaminated water appears to have been the main cause of the focal outbreaks, as in previously reported Turkish outbreaks [13,14]. The patients were using water from the same reservoir (which was unchlorinated at the time of the outbreaks) for both drinking and other needs. The water inlet supplying the used reservoirs in these areas was not piped and the canals supplying drinking water were not covered, leaving the water open for contamination from environmental sources. Thus, it is reasonable to assume that upstream water became contaminated by small rodents killed by tularaemia [5].

The results presented here suggest that although the required infectious dose of *F. tularensis* in humans by the oral route may be high, contaminated drinking water poses a substantial risk to human health. Thus, the often cited bioterrorism potential of *F. tularensis* may not be solely restricted to aerosol distribution, and intentional contamination of drinking water may be an underestimated risk.

In conclusion, whole genome sequencing of outbreak strains confirmed an epidemiological link between drinking water and three outbreaks of human tularemia in Turkey. The genomic epidemiology approach is particularly powerful for genetically monomorphic bacterial pathogens such as *F. tularensis*.

Acknowledgments

We thank all the health workers in the epidemic areas for their professional help in collecting the samples. This project was supported in part by the Swedish Ministry of Defence (No. A404014).

Conflict of interest

None declared.

Authors' contributions

All of the authors collaborated in the presented work. AK, MF, AS, MT, HŞ defined the research theme. AK, MF, AS AL, CÖ, KM, LÖ analysed data, interpreted results and wrote the draft manuscript. AJ gave conceptual advice and participated in manuscript writing. All authors have contributed to, seen and approved the manuscript.

References

- Hornick RB, Dawkins AT, Eigelsbach HT, Tulis JJ. Oral tularemia vaccine in man. *Antimicrob Agents Chemother* (Bethesda). 1966;6:11-4. PMID:4860768
- Quan SF, McManus AG, Von Fintel H. Infectivity of tularemia applied to intact skin and ingested in drinking water. *Science*. 1956;123(3204):942-3. <http://dx.doi.org/10.1126/science.123.3204.942-a> PMID:13324122
- Karadenizli A, Gurcan S, Kolayli F, Vahaboglu H. Outbreak of tularaemia in Golcuk, Turkey in 2005: report of 5 cases and an overview of the literature from Turkey. *Scand J Infect Dis*. 2005;37(10):712-6. <http://dx.doi.org/10.1080/00365540510012125> PMID:16191887
- Ulu-Kilic A, Gulen G, Sezen F, Kilic S, Sencan I. Tularemia in central Anatolia. *Infection*. 2013;41(2):391-9. <http://dx.doi.org/10.1007/s15010-012-0355-1> PMID:23104256
- Gurcan S. Epidemiology of tularemia. *Balkan Med J*. 2014;31(1):3-10. <http://dx.doi.org/10.5152/balkanmedj.2014.13117> PMID:25207161
- Ellis J, Oyston PCF, Green M, Titball RW. Tularemia. *Clin Microbiol Rev*. 2002;15(4):631-46. <http://dx.doi.org/10.1128/CMR.15.4.631-646.2002> PMID:12364373
- Penn R. *Francisella tularensis* (Tularemia). In: Mandell GL, Bennet JE, Dolin RC, editors. *Mandell, Douglas, and Bennett's principles and practice of infectious diseases*. Livingstone, Philadelphia: Elsevier; 2010:2927-2937.
- 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(8):2474-84. <http://dx.doi.org/10.1128/JB.01786-08> PMID:19251856
- 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(12):e8360. <http://dx.doi.org/10.1371/journal.pone.0008360> PMID:20027310
- Özsüreki Y, Birdsell DN, Çelik M, Karadağ-Öncel E, Johansson A, Forsman M, et al. Diverse *Francisella tularensis* strains and oropharyngeal tularemia, Turkey. *Emerg Infect Dis*. 2015;21(1):173-5. <http://dx.doi.org/10.3201/eid2101.141087> PMID:25531237
- Gurcan S, Eskioçak M, Varol G, Uzun C, Tatman-Otkun M, Sakru N, et al. Tularemia re-emerging in European part of Turkey after 60 years. *Jpn J Infect Dis*. 2006;59(6):391-3. PMID:17186960
- Akalin H, Helvacı S, Gedikoğlu S. Re-emergence of tularemia in Turkey. *Int J Infect Dis*. 2009;13(5):547-51. <http://dx.doi.org/10.1016/j.ijid.2008.09.020> PMID:19119037
- Simşek H, Taner M, Karadenizli A, Ertek M, Vahaboğlu H. Identification of *Francisella tularensis* by both culture and real-time TaqMan PCR methods from environmental water specimens in outbreak areas where tularemia cases were not previously reported. *Eur J Clin Microbiol Infect Dis*. 2012;31(9):2353-7. <http://dx.doi.org/10.1007/s10096-012-1576-z> PMID:22382819
- Gurcan S. [Francisella tularensis and tularemia in Turkey]. *Mikrobiyol Bul*. 2007;41(4):621-36. Turkish. PMID:18173084
- World Health Organization (WHO). WHO guidelines on tularemia. Geneva, WHO: 2007. Available from: <http://apps.who.int/iris/handle/10665/43793>
- Broekhuijsen M, Larsson P, Johansson A, Byström M, Eriksson U, Larsson E, et al. Genome-wide DNA microarray analysis of *Francisella tularensis* strains demonstrates extensive genetic conservation within the species but identifies regions that are unique to the highly virulent *F. tularensis* subsp. *tularensis*. *J Clin Microbiol*. 2003;41(7):2924-31. <http://dx.doi.org/10.1128/JCM.41.7.2924-2931.2003> PMID:12843022
- Svensson K, Sjödin A, Byström M, Granberg M, Brittnacher MJ, Rohmer L, et al. Genome sequence of *Francisella tularensis* subspecies *holarctica* strain FSC200, isolated from a child with tularemia. *J Bacteriol*. 2012;194(24):6965-6. <http://dx.doi.org/10.1128/JB.01040-12> PMID:23209222
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9(4):357-9. <http://dx.doi.org/10.1038/nmeth.1923> PMID:22388286
- Chevreur B, Pfisterer T, Drescher B, Driesel AJ, Müller WEG, Wetter T, et al. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. *Genome Res*. 2004;14(6):1147-59. <http://dx.doi.org/10.1101/gr.1917404> PMID:15140833
- Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. ABySS: a parallel assembler for short read sequence data. *Genome Res*. 2009;19(6):1117-23. <http://dx.doi.org/10.1101/gr.089532.108> PMID:19251739
- Lärkeryd A, Myrtenäs K, Karlsson E, Dwibedi CK, Forsman M, et al. CanSNPer: a hierarchical genotype classifier of clonal pathogens. *Bioinformatics*. 2014;30(12):1762-4. <http://dx.doi.org/10.1093/bioinformatics/btu113>. PMID: 24574113
- Vogler AJ, Birdsell D, Wagner DM, Keim P. An optimized, multiplexed multi-locus variable-number tandem repeat analysis system for genotyping *Francisella tularensis*. *Lett Appl Microbiol*. 2009;48(1):140-4. <http://dx.doi.org/10.1111/j.1472-765X.2008.02484.x> PMID:19018964
- Chanturia G, Birdsell DN, Kekelidze M, Zhgenti E, Babuadze G, Tsertsvadze N, et al. Phylogeography of *Francisella tularensis* subspecies *holarctica* from the country of Georgia. *BMC Microbiol*. 2011;11(1):139. <http://dx.doi.org/10.1186/1471-2180-11-139> PMID:21682874
- 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(2):290-3. <http://dx.doi.org/10.3201/eid1802.111305> PMID:22305204
- Karlsson E, Svensson K, Lindgren P, Byström M, Sjödin A, Forsman M, et al. The phylogeographic pattern of *Francisella tularensis* in Sweden indicates a Scandinavian origin of EuroSiberian tularaemia. *Environ Microbiol*. 2013;15(2):634-45. <http://dx.doi.org/10.1111/1462-2920.12052> PMID:23253075
- Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS ONE*. 2010;5(6):e11147. <http://dx.doi.org/10.1371/journal.pone.0011147> PMID:20593022
- Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Mol Biol Evol*. 2013;30(12):2725-9. <http://dx.doi.org/10.1093/molbev/mst197> PMID:24132122
- Larsson P, Oyston PCF, Chain P, Chu MC, Duffield M, Fuxelius H-H, et al. The complete genome sequence of *Francisella tularensis*, the causative agent of tularemia. *Nat Genet*. 2005;37(2):153-9. <http://dx.doi.org/10.1038/ng1499> PMID:15640799
- Larsson P, Elfsmark D, Svensson K, Wikström P, Forsman M, Brettin T, et al. Molecular evolutionary consequences of niche restriction in *Francisella tularensis*, a facultative intracellular pathogen. *PLoS Pathog*. 2009;5(6):e1000472. <http://dx.doi.org/10.1371/journal.ppat.1000472> PMID:19521508
- Sjödin A, Svensson K, Ohrman C, Ahlinder J, Lindgren P, Duodu S, et al. Genome characterisation of the genus *Francisella* reveals insight into similar evolutionary paths in pathogens of mammals and fish. *BMC Genomics*. 2012;13(1):268. <http://dx.doi.org/10.1186/1471-2164-13-268> PMID:22727144