Phylogeographical pattern of *Francisella tularensis* in a nationwide outbreak of tularaemia in Norway, 2011

J E Afset1,2, K W Larssen1, K Bergh1,2, A Lärkeryd3, A Sjödin3, A Johansson4, M Forsman (mats.forsman@foi.se)3
1. Department of Laboratory Medicine, Children’s and Women’s Health, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway
2. Department of Medical Microbiology, St. Olavs Hospital, University Hospital, Trondheim, Norway
3. Division of CBRN Security and Defence, FOI - Swedish Defence Research Agency, Umeå, Sweden
4. Department of Clinical Microbiology, The Laboratory for Molecular Infection Medicine Sweden (MIMS), Umeå University, Umeå, Sweden

In 2011, a nationwide outbreak of tularaemia occurred in Norway with 180 recorded cases. It was associated with the largest peak in lemming density seen in 40 years. *Francisella tularensis* was isolated from 18 patients. To study the geographical distribution of *F. tularensis* genotypes in Norway and correlate genotype with epidemiology and clinical presentation, we performed whole genome sequencing of patient isolates. All 18 genomes from the outbreak carried genetic signatures of *F. tularensis* subsp. *holarctica* and were assigned to genetic clades using canonical single nucleotide polymorphisms. Ten isolates were assigned to major genetic clade B.6 (subclade B.7), seven to clade B.12, and one to clade B.4. The B.6 subclade B.7 was most common in southern and central Norway, while clade B.12 was evenly distributed between the southern, central and northern parts of the country. There was no association between genotype and clinical presentation of tularaemia, time of year or specimen type. We found extensive sequence similarity with *F. tularensis* subsp. *holarctica* genomes from high-endemic tularaemia areas in Sweden. Finding nearly identical genomes across large geographical distances in Norway and Sweden imply a life cycle of the bacterium without replication between the outbreaks and raise new questions about long-range migration mechanisms.

Introduction

Tularaemia is a zoonotic infection caused by the bacterium *Francisella tularensis*, a pathogen with extreme infectivity and a potential biothreat [1]. Several subspecies have been recognised, of which *F. tularensis* subsp. *holarctica* (or type B) is present in Europe. Tularaemia is a rare disease in a global context, making the experience with outbreak investigations very limited. The Scandinavian countries, however, quite regularly experience outbreaks in humans involving tens to hundreds of patients. In Norway, several outbreaks have been attributed to *F. tularensis* contamination of water wells by dead rodents. The exposure to infection by drinking water results in an oropharyngeal clinical form of tularaemia in humans and this is the most common form reported in Norway [2-4]. For unknown reasons, the incidence rate of tularaemia in humans is lower in Norway than in the two other Nordic countries Finland and Sweden. While a mean annual number of 34 (range: 11–66) tularaemia cases were reported in Norway 2006 to 2010, Finland and Sweden reported during the same period a mean annual number of 298 and 305, respectively [5]. The mode of transmission to humans is also different because tularaemia in Finland and Sweden is generally transmitted by the bite of an arthropod taking a blood meal resulting in the ulceroglandular form of tularaemia [6,7]. The reasons for the epidemiological differences in incidence and clinical form of tularaemia between the three countries are unknown.

Distinct genetic subpopulations (major phylogenetic clades) have been identified among *F. tularensis* subsp. *holarctica* strains [8]. High-resolution molecular methods including whole genome sequencing distinguish four major genetic clades denoted B.12, B.6, B.4 and B.16 [9-11]. The clades occur with different frequency in different geographical areas [12-15]. Recent studies in Sweden have indicated phylogeographical patterns both in local outbreaks and across larger geographical distances. These studies from Sweden also identified areas where *F. tularensis* persisted over several years and spatial associations of certain genetic subpopulations [16]. Little is known about phylogeographical patterns of *F. tularensis* in Norway.

In 2011, a large outbreak of tularaemia occurred in Norway with a total of 180 cases coinciding with the highest density of lemmings recorded in the last forty years [17]. Although the high incidence lasted
throughout the year, there were clear differences in epidemiology between seasons, both in the incidences in different geographical areas and in the clinical forms of tularemia recorded. The outbreak started in January 2011 in central Norway (Sør-Trøndelag and neighbouring counties) with mainly cases of oropharyngeal tularemia linked to the use of drinking water from private wells [18]. This part of the outbreak lasted until April. In the period from May to September, sporadic cases occurred scattered throughout the country with increasing frequency of the ulceroglandular form of tularemia linked to insect bites [17]. From October to December, many cases of tularemia were reported from the north of Norway, equally distributed between the oropharyngeal, glandular, typhoidal and respiratory forms of tularemia. For comparison, the number of tularemia cases reported in 2011 in Sweden was 349, slightly above the annual average of the period 2006 to 2010 [5,19].

The aim of this project was to use whole genome sequencing for genotyping of *F. tularensis* cultured from human specimens during the outbreak in 2011 to investigate the genotype distribution of tularemia in Norway. We also wanted to analyse associations of genotype with epidemiological characteristics and disease presentation and identify patterns of spread of the bacterium.

### Methods

#### Isolation of clinical strains

*F. tularensis* was cultured from 18 of the 180 patients diagnosed with tularemia in Norway in 2011. All these isolates were included in the study. *F. tularensis* was cultured from four of 57 tularemia cases during the period January to April, from 10 of 40 cases in the period May to September (seven of these in August), and from four of 83 cases in the period October to December. The isolates were from blood (*n* = 7), skin ulcer (*n* = 7), tissue biopsy (*n* = 2) and aspirate (*n* = 2) specimens from patients living in nine different counties (Table).

The study, which was based on informed written consent from each patient, was approved by the Regional Committee for Medical and Health Research Ethics, central Norway (project 2012/867).

#### Genome sequencing and assembly

Genome sequencing was performed using 100 bp paired-end libraries on an Illumina HiSeq 2000 machine (Illumina, San Diego, CA). The reads were assembled de novo using the ABySS [20]. The raw sequence reads were mapped back against the draft genome using Bowtie 2 [21] allowing identification of uncertain positions in the genome with VarScan 2 [22]. Positions reported having a second variant with frequency higher than 20% were subsequently masked. The overall genome coverage was ca 850-fold.

<table>
<thead>
<tr>
<th>Strain ID</th>
<th>Type of specimen</th>
<th>Clinical classification</th>
<th>Time of year</th>
<th>Geographical region</th>
<th>Genetic clade/subclade</th>
<th>NCBI accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>NO-1/2011</td>
<td>Blood</td>
<td>Typhoidal</td>
<td>January</td>
<td>Central</td>
<td>B.7</td>
<td>JPPD00000000</td>
</tr>
<tr>
<td>NO-2/2011</td>
<td>Blood</td>
<td>Respiratory</td>
<td>February</td>
<td>Central</td>
<td>B.12</td>
<td>JPMN00000000</td>
</tr>
<tr>
<td>NO-3/2011</td>
<td>Blood</td>
<td>Respiratory</td>
<td>February</td>
<td>South</td>
<td>B.7</td>
<td>JPEE00000000</td>
</tr>
<tr>
<td>NO-4/2011</td>
<td>Blood</td>
<td>Typhoidal</td>
<td>March</td>
<td>Central</td>
<td>B.7</td>
<td>JPMJ00000000</td>
</tr>
<tr>
<td>NO-5/2011</td>
<td>Aspirate</td>
<td>Ulceroglandular</td>
<td>May</td>
<td>South</td>
<td>B.12</td>
<td>JPPF00000000</td>
</tr>
<tr>
<td>NO-6/2011</td>
<td>Exudate from ulcer</td>
<td>Ulceroglandular</td>
<td>July</td>
<td>South</td>
<td>B.7</td>
<td>JPMK00000000</td>
</tr>
<tr>
<td>NO-7/2011</td>
<td>Exudate from ulcer</td>
<td>Ulceroglandular</td>
<td>August</td>
<td>North</td>
<td>B.12</td>
<td>JPML00000000</td>
</tr>
<tr>
<td>NO-8/2011</td>
<td>Exudate from ulcer</td>
<td>Ulceroglandular</td>
<td>August</td>
<td>South</td>
<td>B.12</td>
<td>JPG00000000</td>
</tr>
<tr>
<td>NO-9/2011</td>
<td>Exudate from ulcer</td>
<td>Ulceroglandular</td>
<td>August</td>
<td>South</td>
<td>B.7</td>
<td>JPI00000000</td>
</tr>
<tr>
<td>NO-10/2011</td>
<td>Tissue biopsy</td>
<td>Ulceroglandular</td>
<td>August</td>
<td>South</td>
<td>B.7</td>
<td>JPH00000000</td>
</tr>
<tr>
<td>NO-11/2011</td>
<td>Tissue biopsy</td>
<td>Respiratory</td>
<td>August</td>
<td>South</td>
<td>B.7</td>
<td>JPI00000000</td>
</tr>
<tr>
<td>NO-12/2011</td>
<td>Exudate from ulcer</td>
<td>Ulceroglandular</td>
<td>August</td>
<td>South</td>
<td>B.12</td>
<td>JPMN00000000</td>
</tr>
<tr>
<td>NO-13/2011</td>
<td>Aspirate</td>
<td>Ulceroglandular</td>
<td>August</td>
<td>South</td>
<td>B.7</td>
<td>JPO00000000</td>
</tr>
<tr>
<td>NO-14/2011</td>
<td>Blood</td>
<td>Typhoidal</td>
<td>September</td>
<td>North</td>
<td>B.7</td>
<td>JPK00000000</td>
</tr>
<tr>
<td>NO-15/2011</td>
<td>Blood</td>
<td>Respiratory</td>
<td>November</td>
<td>North</td>
<td>B.4</td>
<td>JPMO00000000</td>
</tr>
<tr>
<td>NO-16/2011</td>
<td>Exudate from ulcer</td>
<td>Ulceroglandular</td>
<td>November</td>
<td>Central</td>
<td>B.7</td>
<td>JMQ00000000</td>
</tr>
<tr>
<td>NO-17/2011</td>
<td>Exudate from ulcer</td>
<td>Ulceroglandular</td>
<td>October</td>
<td>North</td>
<td>B.12</td>
<td>JMR00000000</td>
</tr>
<tr>
<td>NO-18/2011</td>
<td>Blood</td>
<td>Glandular</td>
<td>November</td>
<td>North</td>
<td>B.12</td>
<td>JMS00000000</td>
</tr>
</tbody>
</table>

NCBI: National Center for Biotechnology Information.
Results
All 18 isolates of *F. tularensis* cultured from patients in Norway in 2011 were identified as *F. tularensis* subsp. *holarctica* based on whole genome sequence genotyping. The distribution of the isolates in genetic clades and subclades is shown in the Table and Figure 1. Seven of the isolates were assigned to the major genetic clade B.12, one to clade B.4 and 10 isolates to a subclade denoted B.7 belonging in the major clade B.6.

Type of specimen
Among the seven isolates cultured from blood, four belonged to subclade B.7, while two belonged to B.12 and one to clade B.4 (Table). Four isolates from skin ulcers belonged to subclade B.7 and three to B.12. The two isolates from aspirates belonged to B.12 and B.7, respectively, while both isolates from tissue biopsies belonged to subclade B.7.

Clinical form of tularaemia
Ten of the *F. tularensis* isolates were from cases of ulceroglandular tularaemia (Table). Five of these isolates belonged to subclade B.7; the other five strains belonged to clade B.12. Two of the four isolates from patients with a respiratory form of tularaemia belonged to subclade B.7, while the other two belonged to B.12 and B.4. The three isolates recovered from patients with typhoidal tularaemia all belonged to subclade B.7. Finally, the isolate from a patient with glandular tularaemia was of clade B.12.

Mode of infection
Three of the six *F. tularensis* isolates cultured from ulcers caused by insect bite belonged to subclade B.7 while the three others belonged to subclade B.12. The three isolates recovered from patients with typhoidal tularaemia who reported to have drunk water from a private well, belonged to subclade B.7. Inhalation (clade B.12), outdoor work in a farm (subclade B.7), handling of a dead rabbit (subclade B.7), stab wound while handling fish (clade B.12), and hunting and other outdoors activities (clade B.12) were reported as likely mode of infection for five other isolates. For the four remaining isolates, information on the likely mode of infection was not available.

Seasonal distribution
Three of the four *F. tularensis* isolates cultured in the period January to April belonged to subclade B.7 and one belonged to clade B.12 (Table). In the period May to September, 10 cases were culture-positive, eight of whom had the ulceroglandular form of tularaemia. Six of the isolates from this period belonged to subclade B.7, and four to clade B.12. Among the four isolates recovered during the period October to December, two belonged to subclade B.12, one to clade B.4 and one to subclade B.7.

Geographical distribution
Among the nine isolates recovered from patients in southern Norway, six belonged to subclade B.7 and
three to clade B.12 (Figure 2). In comparison, three of the four isolates from central Norway belonged to subclade B.7 and one belonged to clade B.12, while one of the five isolates from patients living in northern Norway belonged to subclade B.7, three belonged to B.12 and one to clade B.4. Two patients living in southern Norway provided information that they may have contracted the infection while walking in the mountains or working on a farm without specifying where. For a third patient, information on the place of infection was not available. One patient living in northern Norway reported infection after an insect bite while travelling in Sweden (strain NO-7/2011).

Comparison with Swedish strains
Among the 313 archived Swedish *F. tularensis* strains included in the analysis, the genomes of 29 strains differed at two or fewer SNPs compared with the outbreak genomes from Norway (Figure 1). All those 29 strains had been isolated during the period 1981 to 2010 in areas highly endemic for tularemia in central Sweden (Figure 2). Since culture for *F. tularensis* was not routinely performed in Norwegian medical microbiology laboratories before 2011, archived genome sequences from Norwegian *F. tularensis* isolates from that period were not available for this study.

**Discussion**
In this study we have shown that 10 of 18 *F. tularensis* subsp. *holarctica* isolates from human tularemia cases from the outbreak in Norway in 2011 belonged to the genetic subclade B.7, while seven isolates belonged to clade B.12 and one to clade B.4. Isolates from the clade B.12 showed higher genomic variability than those belonging to the subclade B.7. Comparison of the isolates from the outbreak in Norway with archived genome sequences from Sweden since 1981 revealed strong sequence identity between some Norwegian and Swedish isolates from high-endemic areas. Twenty-nine isolates from Sweden differed at two or fewer SNPs from one or several isolates from Norway (Figure 1). No information is available on the mutation rate for *F. tularensis* in general or during replication in hosts. Thus, the meaning of near-identical strains (a maximum of two SNPs across the genome) cannot be put into context of mutation rate. Because both countries are located on the Scandinavian Peninsula and share a long border, our results support a common ecology of *F. tularensis* on the Peninsula, although the incidence and mode of transmission of tularemia differ between Norway and Sweden. Reasons for differences in the epidemiology of tularemia between the two countries are not known, but one possible explanation of the higher proportion of oropharyngeal cases in Norway could be that private wells for drinking water might be more common in Norway [18]. Our data suggest that several of the genetic clades have moved long geographical distances and that near-identical genetic clones of the bacterium were found to be far apart both geographically and temporally. An alternative explanation may be that identical genotypes of *F. tularensis* have evolved by parallel evolution in many geographical regions of the Scandinavian Peninsula. Our results may be consistent with a life cycle of the bacterium that includes a state of quiescence (absence or very low level of replication) for long periods of time between outbreaks [30,31].

Both subclade B.7 and clade B.12 of *F. tularensis* subsp. *holarctica* were detected in tularemia cases from all three regions of Norway. However, while subclade B.7 were mostly found in the southern and central regions of the country, the clade B.12 isolates were more evenly distributed between the regions (Figure 2).

Before the outbreak in 2011 and since the disease became a notifiable disease in Norway in 1977, only five cases of human tularemia have been reported from the northernmost county Finnmark. Altogether 62 cases of tularemia were reported from this county.
during this outbreak which in Finnmark lasted until 2012. Considering the high sequence identity between the B.7 isolate from Finnmark (strain NO-14/2011) and other B.7 isolates both from southern Norway and central Sweden, the emergence of tularemia in Finnmark was most probably caused by bacteria already present in the environment. Their activation and amplification may have required additional factors such as the large outbreak in lemmings that occurred in Finnmark in 2011. This raises questions about what mechanisms allowed these near-identical strains to spread across such large geographical distances. Birds have been implicated in the transportation of *F. tularensis* [32-33]. However, while occasional transportation by birds cannot be excluded, the general geographical pattern of near-identical strains found in this study is not characteristic of the north-south migration routes of migratory birds on the Scandinavian Peninsula. Many different mechanisms for dissemination could be envisaged, such as long-range aerosol transport by wind, carriage by arthropod vectors and/or infected migratory wild animals, or a combination of several mechanisms. Further research into this phenomenon is needed.

In this study we found no statistical association between the genotype of *F. tularensis* and type of specimen, clinical presentation, mode of transmission or time of the year when the specimen was collected, although the low number of bacterial strains may have obscured weak associations (Table). Our findings are in analogy with those reported in a recent genomic study of a respiratory tularemia outbreak in Sweden [30]. In the latter study it was shown that the respiratory form of tularemia was not tied to specific genotypes of *F. tularensis* and that outbreak genomes shared high sequence similarity with archived isolates originating from patients from distant geographical regions and collected up to 10 years apart [30]. Despite the mentioned lack of significant association in our study, it is worth noting that *F. tularensis* of clade B.12 was found mainly in patients with ulceroglandular tularemia, and that all three isolates from patients with typhoidal tularemia belonged to subclade B.7 (Table).

We were able to culture *F. tularensis* from only 10% of the tularemia cases in 2011. Few reports are available on the sensitivity of culture in the diagnosis of tularemia, but a low sensitivity has been reported in an outbreak of ulceroglandular tularemia [34]. This, as well as the overrepresentation of the ulceroglandular form of the disease among culture-positive cases, make it difficult to assess the representativeness of the isolates for the whole outbreak. Another limitation is that for three of the isolates, we did not have data on where the patients had been infected.

In conclusion, 18 isolates of *F. tularensis* subsp. *holarctica* from a nationwide outbreak of tularemia in Norway were genotyped by whole genome sequencing: among those, subclade B.7 was most frequent (10 isolates), followed by clade B.12 (seven isolates) and B.4 (one isolate). We found no association between genotype and clinical presentation of tularemia, time of year of disease, or specimen type. Subclade B.7 was most common in southern and central Norway, while three of the five isolates from patients from northern Norway belonged to clade B.12. The isolates from this study showed near-identity with archived genomes from high-endemic areas in Sweden.

Acknowledgments

We thank Norwegian microbiology laboratories where *F. tularensis* were isolated from patient samples for submitting the isolates to the national reference laboratory, and laboratory technicians at Department of Medical Microbiology, St Olavs Hospital for technical support. Sequencing was mainly performed by the SNP&SEQ Technology Platform which is supported by Uppsala University, Uppsala University Hospital, Science for Life Laboratory in Uppsala, Sweden and the Swedish Research Council. 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. JEA, MF, AS, AL, AJ defined the research theme. KWL, JEA and KB did the primary characterisation of the bacterial strains, applied for ethical approval of the study and collected patient data. JEA, MF, AS AL analysed data, interpreted results and wrote the draft manuscript. All authors have contributed to and approved the manuscript.

References
