 doi:10.20944/preprints202209.0071.v1

Article

The First Finding of *Francisella Tularensis* Subsp. *Mediasiatica* in Krasnoyarsk Territory, Siberia and an Update of the Subspecies Genetic Diversity

Vitalii Timofeev[1], Irina Bakhteeva[1], Alexander Mokrievich[1], Galina Vakhrameeva[1], Elena Gritskova[2], Yuriy Anisimov[3], Evgeny Rozhdestvensky[4], Galina Bazarova[4], Rostislav Zhumakaev[1], Ivan Dyatlov[1] and Gilles Vergnaud[5]

1- State Research Center for Applied Microbiology and Biotechnology (SRCAMB), 142279, Obolensk, Moscow Region, Russia
2- Federal healthcare institution "Center for Hygiene and Epidemiology in the Krasnoyarsk Territory", 660100, Krasnoyarsk, st. Sopochnaya, 38, Russia
3- The Baikal State Natural Biosphere Reserve, Russia
4- Altai anti-plague station, 649002 Altai Republic Gorno-Altaysk, str.Zavodskaya, 2, Russia
5- Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), 91198, Gif-sur-Yvette, France

Abstract: Tularemia is a severe infectious disease caused by the Gram-negative bacteria *Francisella tularensis*. *F. tularensis* is currently divided into three subspecies, *holarctica, tularensis* and *mediasiatica* which differ in their virulence and geographic distribution. Subspecies *mediasiatica* is the least studied because of its very low documented virulence for humans and limited geographic distribution. It was discovered in sparsely populated regions of Central Asia. Since 2011, a new subsp. *mediasiatica* lineage was identified in Altai (Russia). In 2021, we isolated one subsp. *mediasiatica* strain in Krasnoyarsk Territory. In spite of its geographic origin, 500 km east from Altai, this strain belongs to the Altai lineage and contributes surprisingly little genetic diversity to previous knowledge. This improved knowledge of the phylogeography of subsp. *mediasiatica* led us to propose a scenario in which the two zoonotic lineages, *holarctica* and *tularensis*, independently emerged from the *mediasiatica* lineage in Siberia or Central Asia, and make predictions which will allow to challenge this hypothesis

Keywords: *Francisella tularensis*; subsp. *mediasiatica*; phylogeography; evolution

1. Introduction

Francisella tularensis is a small Gram-negative aerobic coccobacillus. It is a facultative intracellular parasite capable of infecting a wide range of animals and causing a plague-like disease called tularemia. The pathogenic *F. tularensis* is divided into two subspecies. Whereas *F. tularensis* subsp. *holartica* is present across the northern hemisphere, *F. tularensis* subsp. *tularensis* is restricted to North America [1]. *F. tularensis* subspecies *mediasiatica* is the third of the three currently known subspecies within *F. tularensis*. The virulence for humans of subspecies *mediasiatica* is unknown and it remains the least studied and understood subspecies. *Mediasiatica* was discovered in Central Asia, in some sparsely populated regions of Kazakhstan and Turkmenistan. Since 2011, a number of strains of this subspecies were isolated in the Altai region (Russia) [2]. Multiple Locus VNTR (Variable Number of Tandem Repeats) Analysis (MLVA) [3] allowed to distinguish the Altaic population of *F. tularensis* subsp. *mediasiatica* from the classical Central Asian population. According to MLVA data we

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 September 2022 doi:10.20944/preprints202209.0071.v1

proposed to divide the subspecies into three subgroups: M.I -classical Central Asian strains, M.II - Altaic strains and M.III represented by a single strain isolated in the Republic of Karakalpakstan (Uzbekistan) [2].

In 2021 we discovered subsp. mediasiatica strain K-334 in the Krasnoyarsk Territory about 500 kilometers in a straight line (and over rather rugged terrain) from the Altai focus (S1 Figure) and more than 1500 km from the Central Asian focus. We report the phylogenetic analysis based on the results of whole genome sequencing of K-334 and of a panel of Altai and Central Asian strains. We propose a new scenario about the evolution and spread of *F. tularensis* across Eurasia and North America.

2. Materials and Methods

Strains

F. tularensis strains used in this study are listed in Table 1. Sixteen strains were isolated from ticks and from one dead rodent within the Republic of Altai and Altai Territory in 2011–2015 by the Federal Healthcare Service Center for Hygiene and Epidemiology in the Altai region and by the Altai anti-plague station. Strain K-334 was isolated in the Krasnoyarsk region in 2021 from tick. Two collection strains from Kazakhstan were also included.

Table 1 List of subsp. *mediasiatica* strains used in this work.

Strain	Isolation date	Source	Isolation area
A-99	23.05.2013	ticks *H. concinna* (caught 11.04.2013)	Altai Republic, Maiminsky district, near Aleksandrovka village
A-188	26.06.2013	ticks *H. concinna* (caught 14.05.2013)	Altai Republic, Choysky district, natural landmark Ashpanak
A-82	25.05.2014	ticks *D. silvarum* (caught 20.05.2014)	Altai Republic, Choysky district, near Uba river
A-84	25.05.2014	ticks *D. silvarum* (caught 20.05.2014)	Altai Territory, Sovietskiy district, near Kolovo village
A-116	20.05.2014	ticks *D. silvarum* (caught 12.05.2014)	Altai Territory, Altaiskiy district
A-137	29.05.2014	ticks *H. concinna* (caught 16.05.2014)	Altai Republic, Choysky district, near Paspaul village
A-139	17.06.2014	ticks *H. concinna* (caught 20.05.2014)	Altai Republic, Choysky district, near Ashpanak river
A-142	10.06.2014	ticks *H. concinna* (caught 20.05.2014)	Altai Republic, Choysky district, near Uba river
A-174	23.05.2014	ticks *H. concinna* (caught 16.05.2014)	Altai Republic, Choysky district, near Uba river
A-178	26.05.2014	ticks *D. silvarum* (caught 20.05.2014)	Altai Republic, Choysky district, near Karakokshu village
A-77	23.05.2015	ticks *H. concinna* (caught 27.04.2015)	Altai Republic, Maiminsky district, near the Bakala brook
A-87	14.05.2015	ticks *H. concinna* (caught 29.04.2015)	Altai Republic, Maiminsky district
A-196	21.05.2015	ticks *H. concinna* (caught 12.05.2015)	Altai Republic, Choysky district, near Paspaul village
B-7175 (A-554)	2011	ticks *H. concinna*	Altai Republic, Eltsovskiy district, Martynovo village
B-7176 (A-678)	2011	ticks *I. persulcatus*	Altai Republic, Pervomayskiy district, Pokrovka village

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 September 2022 doi:10.20944/preprints202209.0071.v1

B-7177 (A-823)	2011	Northern red-backed vole (*Clethrionomys rutilus*)	Altai Republic, Shelabolikhinskiy district, Molokovo village
K-334	30.07.2021	ticks *I. persulcatus* and *H. concinna* (caught 20.05.2021)	Krasnoyarsk region, Karatuzsky District, Western Sayan mountain-taiga zone
120	<=1968	unknown	Kazakhstan
117	1960	ticks *Hyalama* sp.	Kazakhstan

Bacterial culture

Strains were grown at 37°C on solid (FT-agar) and liquid (FT-broth) nutrient media (SRCAMB, Obolensk, Russia). The composition of the FT-agar was 3.8% erythritol-agar, 1% dried bovine blood, 1% glucose, 0.05% cysteine, and 0.0025% thiamine chloride at pH 7.2. The composition of the FT-broth was 2% casein enzymatic hydrolysate, 1% yeast extract, 1.2% KH_2PO_4, 1% glucose, 0.001% cysteine, and 0.001% $FeCl_2$ at pH 7.2.

DNA preparation and PCR analysis of species and subspecies

DNA from bacterial cultures was isolated using GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich, Moscow, Russia). PCR amplifications were run on the CFX96 ™ Real-Time PCR Detection System (Bio-Rad, Moscow, Russia).

Species determination was performed using «MULTI-FLU» Real-Time PCR-test kit (SRCAMB, Obolensk, Russia) and «OM-screen-tularemia-RT» (Syntol, Moscow, Russia). Subspecies determination was performed by single-primer genotyping with primer Chif1 (5'-CTAGGGCTGGTGGG-3') as previously described [2] and by typing of single nucleotide polymorphism (SNP) position 630822 in *F. tularensis* strain SCHU S4 genome accession GCA_000008985.1. SNP typing was performed by melt curve analysis of the PCR products as previously described [4] using 2.5× PCRmix with SYBR-GreenI (Syntol, Moscow, Russia) with minor modifications. We used primers Bla2SNPTulF1 (5'-AATAAATCAAGATGATATTGGTAAAGCCG-3'), Bla2SNPMedF1 (5'-CGGGGCGGGGCGGGGCGGGCAATAAATCAAGATGATATTGGTAAA GGCA-3'), Bla2SNPRev3 (5'-ATCTTTAGTAATAGCCTTTGGAGGTG-3'). Two allele-specific (by 3′ ultimate nucleotide) forward primers are designed to detect the two alleles at a SNP locus on DNA template, one of them is labeled with GC-clamp on the 5′ end. These primers work in concert with a third common reverse primer to generate a PCR amplicon. This reverse primer has a 3′ ultimate nucleotide specific for subsp. *tularensis* and subsp. *mediasiatica*. Depending on the template, only one of the forward primers will generate the amplicon in concert with the common primer. The ability to differentiate amplicon derived from each allele-specific primer is accomplished by the melt-curve analisys - amplicon withh GC-clamp has a higher melting temperature. To increase the specificity of this method, an artificial mismatch were added in -3 and -4 position (from the 3′ end) to the sequence of both forward primers and in -3 position to the sequence of reverse primer, according to the principle described in [4]

We used 10 pmol of each of the three primers per 25 µl of reaction mix. Thermocycling parameters were the following: 1) 50°C for 2 min; 2) 95°C for 10 min; 3) 25 cycles: 95°C for 20 sec; 60°C for 30 sec; 72°C for 30 sec; Plate Read (SYBR); 4) 95°C for 15 sec; 5) Melt Curve Read (SYBR) from 60°C to 95°C, increment 0.5 °C, time interval 5 sec. In this assay subsp. *mediasiatica* strains

 doi:10.20944/preprints202209.0071.v1

have a melting curve peak 2 degrees higher than that of strains of other subspecies. PCR primers were synthesized by Syntol

Sequencing

DNA libraries were prepared using the Nextera DNA Library Preparation Kit (Albiogen, Moscow, Russia). Whole-genome sequencing was performed using the MiSeq Illumina instrument and the corresponding reagent kit MiSeq Reagent Kit v3 (Albiogen, Moscow, Russia). WGS data from newly published genomes are available at https://www.ncbi.nlm.nih.gov/sra/PRJNA870100. Data for previously described genomes are available at the link indicated in [5]

WGS analysis

Core whole genome SNP analyses were performed essentially as previously described [6]. Assemblies and sequence reads archives (SRA) were downloaded from EBI ENA. Large-size SRAs (.gz files larger than 100 Mo) were assembled with SKESA [7] and subsequently processed as assemblies. Smaller size SRAs were directly imported in BioNumerics version 7.6.3 (Applied-Maths, Sint-Martens-Latem, Belgium). All assemblies were split in 50 bp long artificial reads with 10x coverage. SNPs were called by mapping reads on a reference genome using BioNumerics default parameters.

HyPhy [8, 9] version 2.5 method aBSREL was used to measure the ratio of non synonymous versus synonymous mutations (dN/dS) in branches of phylogenies. The input alignments were produced by mapping reads on the coding sequences (CDS) of a reference genome. The list of CDS sequences was downloaded from https://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria/ (file *_cds_from_genomic.fna.gz from ncbi ftp download site https://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria/). Partial CDSs such as pseudogenes were removed. Output files from HyPhy were visualized with HyPhy-Vision accessed at http://vision.hyphy.org/.

Homoplasic SNPs were detected and removed with SNPPar [10] version 1.1.

3. Results

K-334 strain isolation and subspecies determination.

In 2021 in the Federal healthcare institution «Center for Hygiene and Epidemiology in the Krasnoyarsk Territory» one *F. tularensis* strain was isolated from homogenized mix of *Haemaphysalis concinna* and *Ixodes persulcatus* ticks (Table 1). To confirm the species and determine its subspecies, the strain was subsequently transferred to the Federal Budget Institution of Science «State research center for applied microbiology and biotechnology» (SRCAMB), where it was given the name K-334 (K - Krasnoyarsk). After species confirmation, subspecies was determined using two independent PCR assays, the first one allowing to distinguish the three subspecies as well as *F. novicida* (Figure 1), and the second one targeting a SNP position allowing to distinguish subspecies *mediasiatica* from the two others (Figure 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 September 2022 doi:10.20944/preprints202209.0071.v1

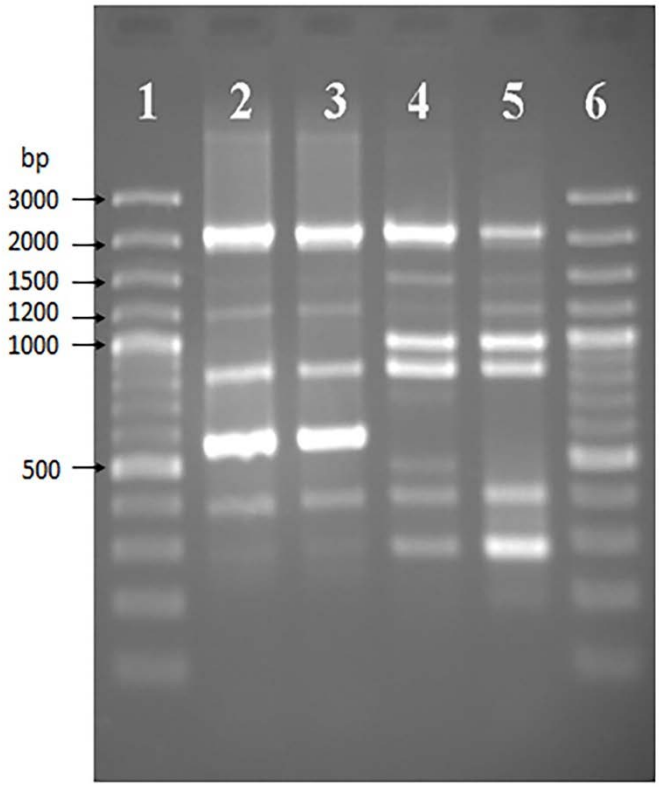


Figure 1. Electropherogram of amplicons obtained in single-primer PCR with *F. tularensis* DNA:1,6 – GeneRulerTM 100 bp Plus DNA Ladder; 2,3 –strains 15 NIIEG and 1045 (subsp. *holarctica*); 4 – strain K-334; 5 –strain A-678 (subsp. *mediasiatica*).

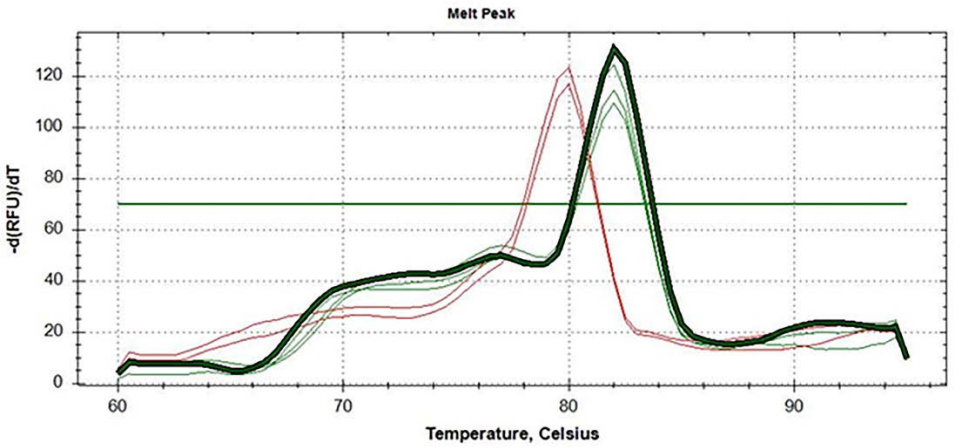


Figure 2. Melt curves analysis of amplicons obtained in allele-specific PCR with a DNA template of: subsp. *holarctica* strains 1045 and 15 NIIEG (red lines) subsp. *mediasiatica* strains A-678 (M.II lineage), 120, 117 (M.I lineage) (green lines) and strain K-334 (dark green bold line).

Whole-genome sequencing (WGS) and phylogenetic analysis.

To confirm the subspecies assignment we conducted WGS of the K-334 strain and compared its genome sequence with the genome sequences of 13 Altai and two Central Asian strains from our collection (Table 1). Nine among these strains were previously analyzed by multiple locus variable number of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 September 2022 doi:10.20944/preprints202209.0071.v1

tandem repeats (VNTR) analysis (MLVA) [2], the last four are new. WGS data from three strains were previously published [2,5]. Strains of Central Asian (M.I) and Altaic (M.II) origin form two clusters separated by 116 SNPs (Figure 3) in agreement with previous analyses [2,11]. The M.II group in turn is split in two. Five Altaic strains originating from Mayminsky, Eltsovka, or Pervomayskoye are separated by ten up to 25 SNPs from the others. Krasnoyark region strain K-334 belongs to a very tight cluster comprising a progenitor genotype represented by three strains isolated in 2014-2015, and nine radiating branches with length of one up to nine SNPs. These nine branches are represented by eight Altai strains isolated in 2011-2014 in addition to K-334. Six of the branches including the K-334 branch have a length of one or two SNPs. This suggests that essentially identical subsp. *mediasiatica* strains are circulating over a few hundred kilometers, which would constitute a single ecotype for *F. tularensis* subsp. *mediasiatica* (S1 Figure). We would like to specifically note that all manipulations with Altai and Central Asian strains were carried out in the period from 2017 to 2020. We began work with strain K-334 in the second half of 2021. In this time period no work with *Francisella* sp. was carried out in the BSL-3 block in which we worked with strain K-334, but disinfection measures were regularly carried out. K-334 was the only *Francisella* sp. strain sequenced in 2021 in SRCAMB. Therefore, we do not assume intralaboratory contamination despite K-334 is almost identical to the Altai strains.

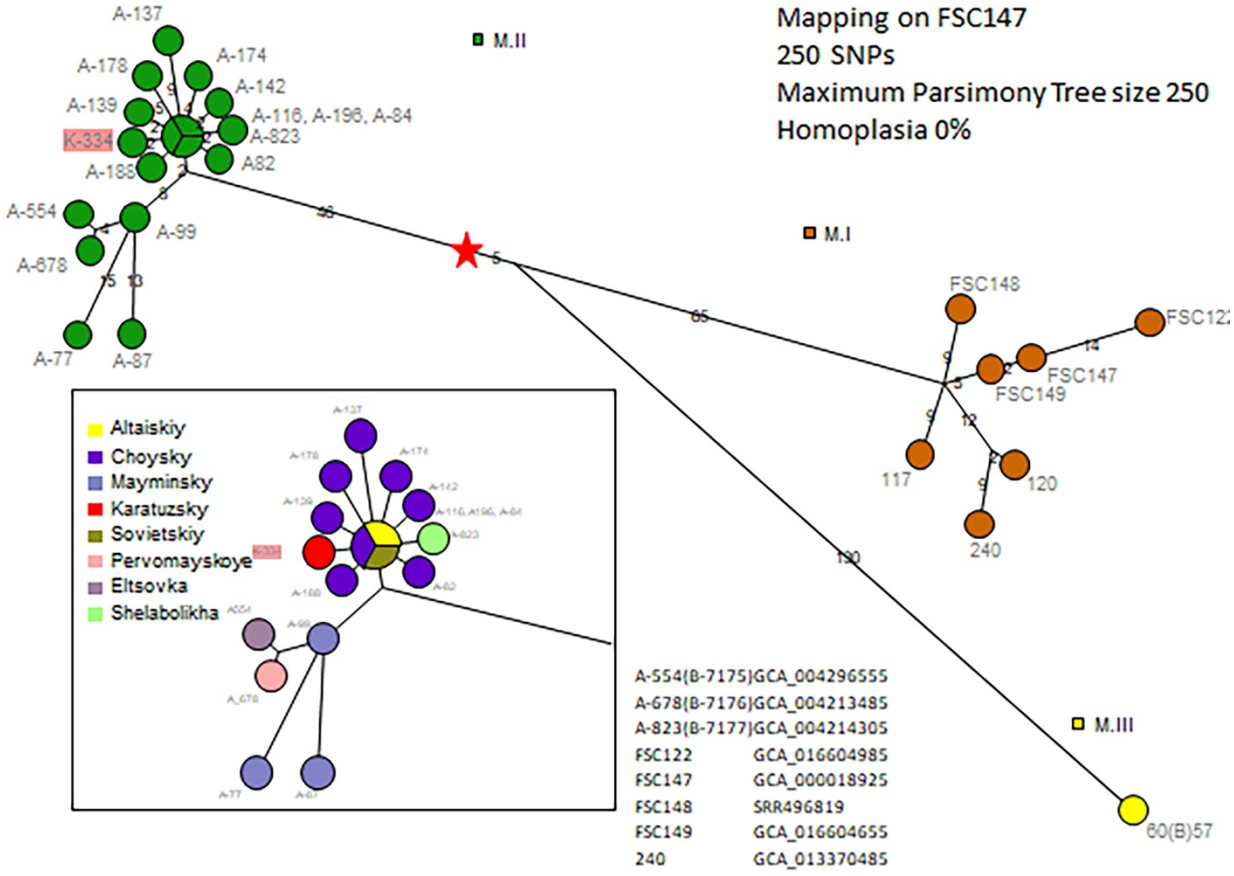


Figure 3: *Francisella tularensis* subsp. *mediasiatica*, Maximum parsimony tree based on core genome SNPs. Two hundred and fifty SNPs were called by mapping on genome accession GCA_000018925 (*mediasiatica* strain FSC147). The size of the resulting tree is 250 (no homoplasia). Branch lengths of two SNPs and more are indicated. Nodes are

 doi:10.20944/preprints202209.0071.v1

labelled with strain Ids and colored according to clade M.I or M.II as indicated. Node coloring in the inset reflects geographic origin of the M.II strains. The red star indicates the root of the M.I M.II mediasiatica lineages (branching point towards the rest of *F. tularensis*). Correspondence between strain Id and sequence data accession numbers is indicated for previously published data.

The WGS data produced in the course of this work confirms the low level of genetic diversity within subspecies *mediasiatica* of strains collected across a wide geographic area. This is in sharp contrast with the diversity observed for instance in subspecies *holarctica* over similar geographic distances [6].

The evolution of *F. tularensis* from its most recent common ancestor (MRCA) is strictly clonal and the species is highly homogeneous with pairwise average nucleotide identities (ANI) above 99.2% [12]. The clonality would imply that *F. tularensis* emerged in a single temporal and geographic point but both place and time of emergence are currently unknown. *F. tularensis* nearest neighbor is *Francisella novicida*, with ANI values above 97.7% [12,13,14]. *F. novicida* has been isolated from environmental samples and its evolution is not clonal [12, 15, 16]. The ratio of non synonymous versus synonymous genetic variations was shown to be strikingly different in *F. novicida* compared to *F. tularensis* [12]. The situation of *F. novicida* versus *F. tularensis* is reminiscent of other highly monomorphic human pathogens, including *Yersinia pestis*, *Mycobacterium tuberculosis* or *Bacillus anthracis* [17,18]. In these three instances, anthropic factors might be largely responsible for the emergence and spread of the associated diseases via the creation of new ecosystems. In these new ecosystems, the microbial environment may be strictly restricted, explaining the subsequent clonal evolution [17,19,20].

In order to better understand the emergence of *F. tularensis*, we tried to position the emergence of the clonal behavior of evolution defining the ancestor of *F. tularensis* along the path going from *F. novicida* to *F. tularensis*. We take advantage of the current knowledge of *F. tularensis* nearest neighbor, *F. novicida* and of the availability of a close outgroup for the two species, *Francisella* sp. TX07-6608 [13,19,20]. Figure S2 shows the calculation of dN/dS ratios within *F. tularensis* and *F. novicida*. Whereas dN/dS ratios vary between 0.04 and 012 along *F. novicida* branches and towards the outgroup, ratios along *F. tularensis* are much higher and range between 0.43 and 0.68, in agreement with previous investigations [12]. The branch immediately upstream of *F. tularensis* shows a value of 0.10 which is a typical *F. novicida* value. This analysis does not allow to evaluate if this upstream branch shows a hybrid behavior between *F. novicida* and *F. tularensis*. Consequently we tested an alternative approach to evaluate this behavior. Figure 4A shows a Maximum Parsimony Tree (MPT) based on SNPs called by mapping WGS data on the genome of outgroup strain TX07-6608. The tree is highly homoplasic (homoplasia level, 35%). Figure 4B shows the MPT tree after removal of homoplasic SNPs using SNPPar. Within the strictly *F. tularensis* part of the tree, branch sizes are reduced by a ratio of 0.95. This very moderate reduction is in agreement with the known clonality and high genetic homogeneity within *F. tularensis*. Within the *F. novicida* part, the ratio is 0.42. The size of the intermediate branch connecting *F. tularensis* and *F. novicida* is reduced from 2688 to 1516, i.e. an intermediate ratio of 0.56, indicating that the clonal behavior started before the split between subsp. *holarctica* and the two other subspecies. The ratio allows to estimate the proportion of the hybrid branch produced under a clonal regimen. This proportion would be approximately 25%, and the first radiation event along *F. tularensis* evolution would have occurred at one-third of *F tularensis* evolution (Figure 5).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 September 2022 doi:10.20944/preprints202209.0071.v1

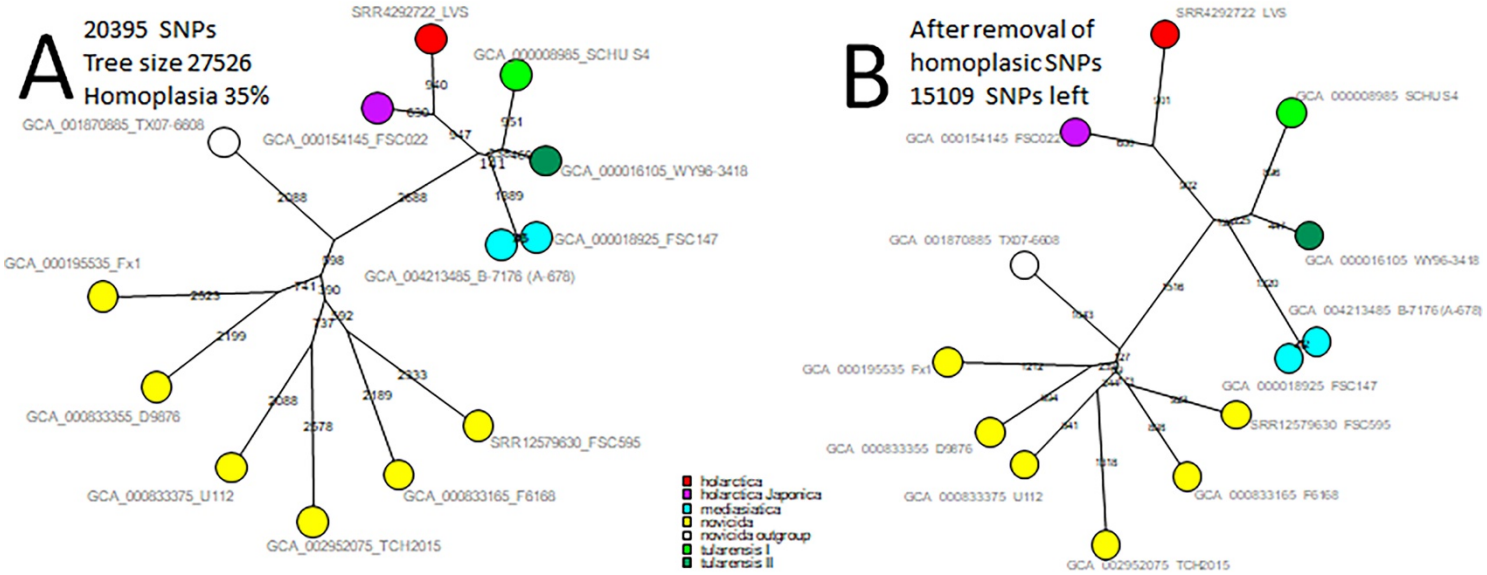
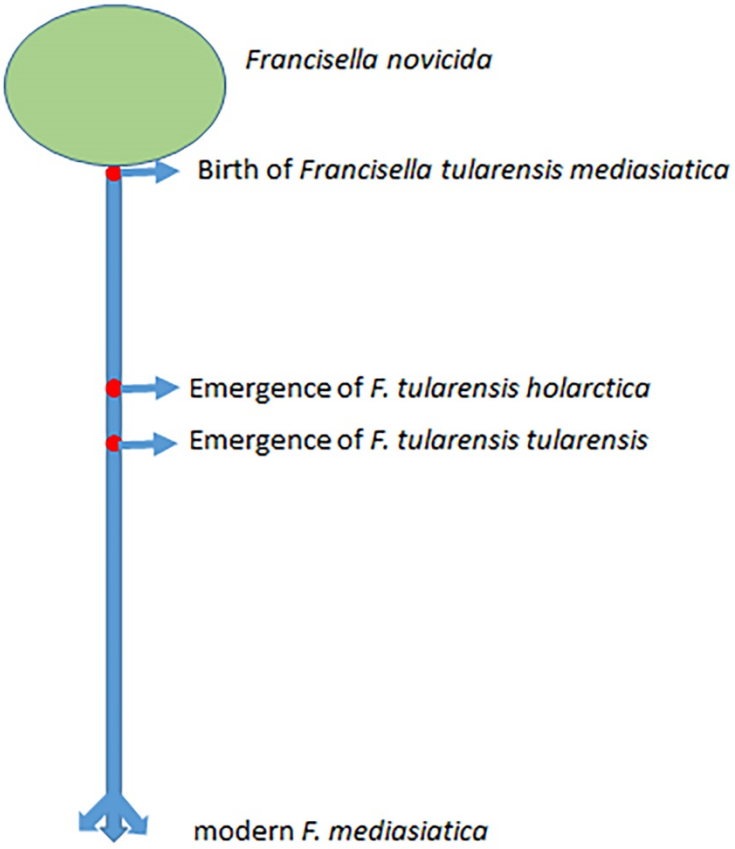


Figure 4: core genome SNP phylogeny of *F. tularensis*, *F. novicida* and outgroup. Nodes are labelled with accession number and strain Id and colored according to species and subspecies as indicated. Core genome SNPs were called by mapping on outgroup genome GCA_001870885 (open circle). Branch lengths (number of SNPs) are indicated. **Part A**: unfiltered. 20395 SNPs are called, tree size 27526, homoplasia 35%. **Part B**: filtered. Homoplasic SNPs were removed using SNPPar. 15109 SNPs were kept.



 doi:10.20944/preprints202209.0071.v1

Figure 5: Model for *F. tularensis* evolution. In the proposed scenario, *F. tularensis* subsp. *mediasiatica* emerged from a *F. novicida* progenitor. The main argument in favor of this scenario is the absence of significant documented virulence of *mediasiatica* and *F. novicida*, in contrast with the two other subspecies. At about one-third along its evolutionary path, subsp. *holarctica* successfully emerged, followed by subsp. *tularensis*. No sign of previous or later emergences have been discovered so far. Modern subsp. *mediasiatica* behaves like a single ecotype able to maintain only minor genetic diversity.

4. Discussion

In this report, we first confirmed by sequencing 14 *F. tularensis* subsp. *mediasiatica* strains that the subspecies displays remarkably low genetic diversity. Currently available subsp. *mediasiatica* strains from the M.II lineage have been collected in a time range of ten years (2011-2021) in a geographic area spanning approximately 600 km. The maximum distance observed between two strains is 34 SNPs. Remarkably, the strain recovered from the Krasnoyarsk region is only two SNPs away from strains recovered from the Altai area indicating that the geographic distribution of the M.II lineage is not restricted to Altai. Determining the geographic distribution of such bacteria is difficult because of the extremely low population density in the region, the resulting underdeveloped infrastructure outside the big cities, the mountainous landscape, and the absence of human infections. The vast majority of M.II strains were isolated from ticks. A single strain was isolated from a dead rodent out of the many rodents collected over a decade, and none was isolated from water or aquatic invertebrates. In addition, most of the local laboratories are not licensed to handle pathogens. Therefore, when the presence of a pathogen is suspected, this fact is often simply recorded, after which the clinical or field sample is usually eliminated, since there are no opportunities for its storage and detailed study. The Altaic mediasiatica strains were isolated owing to the direct collaboration between the Altai Anti-Plague Station and SRCAMB, where subspecies was identified by genotyping. Since then, targeted collection of strains and shipment to reference centers for detailed study led to the identification of several strains belonging to subsp. *mediasiatica*. This illustrates the importance for future investigations to implement first-line genotyping assays which might be applied in local laboratories. Such enhanced local capacities might allow to more precisely delineate the presence of *mediasiatica* in Siberia. The single PCR assay used here is a good candidate for such a first line assay as it also allows resolving *F. novicida*.

Isolating *F. novicida*, *F. tularensis* nearest neighbor, in Central Asia and Siberia would be a significant step towards the better understanding of the emergence of *F. tularensis*. Until now, *F. novicida* was recovered from salt water [14, 20]. The scenario proposed in Figure 5 is implying that the MRCA of the three species was *mediasiatica*-like in terms of virulence for humans, and that subspecies *holarctica* and *tularensis* emerged from a Central Asian or Siberian cradle. The distance between the Altai focus and the place of isolation of the almost identical K-334 strain exceeds 500 km over very rough terrain. A previous investigation pointed to migratory birds as potential long-distance carriers for the spreading of *F. tularensis* [6]. Almost all strains of subsp. *mediasiatica* M.II were isolated from ixodid (hard) ticks *H. concinna* and *D. silvarum*. Accordingly, the area of distribution of these strains falls on the region of Russia in which the areas of distribution of these two species of ticks intersect [24,25]. Ixodid ticks including *H. concinna* and *D. silvarum* can parasitize birds, which could spread both the ticks themselves and the infectious agents, including *F. tularensis* [24-26]. *H. concinna* is the second most abundant tick species collected from birds [24]. Our planet is covered by nine major flyways. Southern Siberia, Central Asia and Iran could be

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 September 2022 doi:10.20944/preprints202209.0071.v1

connected by birds migrating along the Central Asian and West Asian-East African flyways (S3Figure). These bird species winter in Africa, India, the Arabian Peninsula, or Central Asia. During the autumn and spring migrations, birds fly to South Siberia and back through Central Asia, stopping there for migratory stops. Data from satellite monitoring of migration of birds tagged in southern Siberia showed that regions where subsp. *mediasiatica* strains have been found can be linked by bird migration routes (S3 Figure B-E). In addition to the presence of bird migration routes, several more conditions must be met. The season of migration should coincide with the season of tick activity, April to early June in southern Siberia and Altai. A second important point is that during the period of time between the attachment of a tick to the body of a bird and its detachment, the bird must fly from one area suitable for ticks to another such area [25]. Both *H. concinna* and *D. silvarum* ticks prefer boreal climate conditions with precipitation all year round or with dry winter [23,24]. Therefore, large birds making relatively short flights along the boreal humid zone, which extends, among other things, between Altai and the south of the Krasnoyarsk Territory, might be candidates for transfer of subsp. *mediasiatica*. Interestingly, as these routes cover Iran (S3 Figure D), one might expect to find *mediasiatica* in Iran. Interestingly, the presence of *F. tularensis* in mammals of Iran was recently suspected, however the samples were not characterized at subspecies level [27].

However it is important to recall that the vast majority of subsp. *mediasiatica* M.II strains were isolated from ticks in spite of our demonstration in laboratory rodents that subsp. *mediasiatica* including M.II is virulent and capable to effectively overcome post-vaccinal immunity [2,28]. Thus, subsp. *mediasiatica* strains are highly virulent in laboratory experiments, but do not appear to cause significant morbidity in natural conditions. One explanation of this phenomenon might be that *mediasiatica* is not efficiently transferred from ticks to warm-blooded animals and instead circulates as an endosymbiont via both horizontal and vertical transmission among ticks [25,29,30].

In summary, our limited current knowledge of subsp. *mediasiatica* may suggest that it represents a transition between pathogenic and non-pathogenic *Francisella*. This subspecies seems to have acquired a full range of traits allowing it to effectively colonize the body of a warm-blooded animal and multiply in it, causing severe disease, but in contrast with subspp. *tularensis* and *holarctica* it has not yet acquired the mechanisms allowing it to effectively penetrate into the body of a warm-blooded host.

The proposed scenario reflects this view. We showed that the phylogenetic branch prior to the MRCA of extant *F. tularensis* lineages displays a "hybrid" behavior between *novicida* and *tularensis*. The proportion of each behavior along the branch allowed to propose a first estimate for the position of the *F. tularensis* ancestor and this position opens the possibility that more closely related *novicida* lineages may be found in the future. The lack of virulence of *F. tularensis* subsp. *mediasiatica* is an argument in favor of this subspecies to be the progenitor from which the two zoonotic subspecies emerged. We speculate that long-distance transportation to North America or across Eurasia and Japan where spp. *tularensis* and *holarctica* triggered the independent selection of variants with zoonotic potential. The finding of subsp. *mediasiatica* only in Central Asia and Siberia and of *F. novicida* predominantly in salt water [21] may point to salt water in these areas as being the cradle of *F. tularensis*. The Eastern coast of Siberia, the Caspian Sea and the Aral Sea are candidates.

Author Contributions: The authors contributed as follows: conceptualization, G.V. and V.T.; methodology, G.V., A.M.; validation, G.V.; formal analysis, G.V. and V.T.; investigation, I.B., G.V.,E.G., R.Zh., Yu.A., G.B., .; resources, A.M., E.R. and E.G.; data curation, V.T. and G.V.; writing—original draft preparation, V.T. and G.V.; visualization, G.V. and Yu.A.; supervision, I.D.; project administration, A.M., and I.D.; funding acquisition, I.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Sectoral Scientific Program of the. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Supplementary Materials:

Figure S1: geographic origins of *F. tularensis mediasiatica* M.II strains

Figure S2: dN/dS ratios along the *Francisella* phylogeny. Genome accession numbers and strain Ids are indicated. "novicida": *Francisella novicida*; "medi", "tula", "hola": *F. tularensis* subsp. *mediasiatica, tularensis, holarctica* respectively.

Figure S3: The main bird flyways and bird migration tracks linking the regions in which subsp. *mediasiatica* was found

References

1. Telford, S.R., 3rd; Goethert, H.K. Ecology of *Francisella tularensis*. *Annu Rev Entomol* **2020**, *65*, 351-372, doi:10.1146/annurev-ento-011019-025134.
2. Timofeev V, Bakhteeva I, Titareva G, Kopylov P, Christiany D, Mokrievich A, et al. Russian isolates enlarge the known geographic diversity of *Francisella tularensis* subsp. mediasiatica. *PLoS One* **2017**, *12*(9):e0183714.
3. Johansson A, Farlow J, Larsson P, Dukerich M, Chambers E, Bystrom M, et al. Worldwide genetic relationships among *Francisella tularensis* isolates determined by multiple-locus variable-number tandem repeat analysis. *J Bacteriol* **2004**, *186*(17):5808-18
4. Bakhteeva, I.V.; Kravchenko, T.B.; Ryabko, A.K.; Titareva, G.M.; Lev, I.O.; Mokrievich, A.N.; Timofeev, V.S. Features of beta-lactamase activity in *Francisella tularensis* subsp. *mediasiatica*. *iimm* **2018**, *8*, 33-42, doi:10.15789/2220-7619-2018-1-33-42.
5. Mokrievich, A.N.; Kislichkina, A.A.; Kudryavtseva, T.Y.; Mironova, R.I.; Vakhrameeva, G.M.; Shishkova, N.A.; Timofeev, V.S.; Bogun, A.G.; Pavlov, V.M.; Dyatlov, I.A. Draft Genome Sequences of Three *Francisella tularensis* subsp. *mediasiatica* Strains Isolated in the Altai Territory, Russian Federation. *Microbiol Resour Announc* **2020**, *9*, doi:10.1128/MRA.01202-19.
6. Shevtsov, V.; Kairzhanova, A.; Shevtsov, A.; Shustov, A.; Kalendar, R.; Abdrakhmanov, S.; Lukhnova, L.; Izbanova, U.; Ramankulov, Y.; Vergnaud, G. Genetic diversity of *Francisella tularensis* subsp. *holarctica* in Kazakhstan. *PLoS Negl Trop Dis* **2021**, *15*, e0009419, doi:10.1371/journal.pntd.0009419.
7. Souvorov, A.; Agarwala, R.; Lipman, D.J. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol* **2018**, *19*, 153, doi:10.1186/s13059-018-1540-z.
8. Pond, S.L.; Frost, S.D.; Muse, S.V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **2005**, *21*, 676-679, doi:10.1093/bioinformatics/bti079.
9. Smith, M.D.; Wertheim, J.O.; Weaver, S.; Murrell, B.; Scheffler, K.; Kosakovsky Pond, S.L. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Mol Biol Evol* **2015**, *32*, 1342-1353, doi:10.1093/molbev/msv022.
10. Edwards, D.J.; Duchene, S.; Pope, B.; Holt, K.E. SNPPar: identifying convergent evolution and other homoplasies from microbial whole-genome alignments. *Microb Genom* **2021**, *7*, doi:10.1099/mgen.0.000694.
11. Kuibagarov, M.; Kairzhanova, A.; Vergnaud, G.; Amirgazin, A.; Lukhnova, L.; Izbanova, U.; Ramankulov, Y.; Shevtsov, A. Draft Genome Sequence of the Strain *Francisella tularensis* subsp. *mediasiatica* 240, Isolated in Kazakhstan. *Microbiol Resour Announc* **2020**, *9*, doi:10.1128/MRA.00766-20.
12. Larsson, P.; Elfsmark, D.; Svensson, K.; Wikstrom, P.; Forsman, M.; Brettin, T.; Keim, P.; Johansson, A. Molecular evolutionary consequences of niche restriction in *Francisella tularensis*, a facultative intracellular pathogen. *PLoS Pathog* **2009**, *5*, e1000472, doi:10.1371/journal.ppat.1000472.
13. Ohrman, C.; Sahl, J.W.; Sjodin, A.; Uneklint, I.; Ballard, R.; Karlsson, L.; McDonough, R.F.; Sundell, D.; Soria, K.; Backman, S.; et al. Reorganized Genomic Taxonomy of *Francisellaceae* Enables Design of Robust Environmental PCR Assays for Detection of *Francisella tularensis*. *Microorganisms* **2021**, *9*, doi:10.3390/microorganisms9010146.
14. Kingry, L.C.; Petersen, J.M. Comparative review of *Francisella tularensis* and *Francisella novicida*. *Front Cell Infect Microbiol* **2014**, *4*, 35, doi:10.3389/fcimb.2014.00035.
15 . Matz, L.M.; Kamdar, K.Y.; Holder, M.E.; Metcalf, G.A.; Weissenberger, G.M.; Meng, Q. et al. Challenges of Francisella classification exemplified by an atypical clinical isolate. *Diagn Microbiol Infect Dis*. **2018**, *90*, 241-7.

16. Petersen, J.M.; Mead, P.S.; Schriefer, M.E. *Francisella tularensis*: an arthropod-borne pathogen. *Vet Res*. **2009**, *40*(2), 7.
17. Vergnaud, G. *Bacillus anthracis* evolution: taking advantage of the topology of the phylogenetic tree and of human history to propose dating points. *Erciyes Med J* **2020**, *42*, 362-369, doi:10.14744/etd.2020.64920.
18. Achtman, M. How old are bacterial pathogens? *Proc Biol Sci* **2016**, *283*, doi:10.1098/rspb.2016.0990.
19. Blouin, Y.; Hauck, Y.; Soler, C.; Fabre, M.; Vong, R.; Dehan, C.; Cazajous, G.; Massoure, P.L.; Kraemer, P.; Jenkins, A.; et al. Significance of the identification in the Horn of Africa of an exceptionally deep branching *Mycobacterium tuberculosis* clade. *PLoS One* **2012**, *7*, e52841, doi:10.1371/journal.pone.0052841
20. Blouin, Y.; Cazajous, G.; Dehan, C.; Soler, C.; Vong, R.; Hassan, M.O.; Hauck, Y.; Boulais, C.; Andriamanantena, D.; Martinaud, C.; et al. Progenitor "*Mycobacterium canettii*" clone responsible for lymph node tuberculosis epidemic, Djibouti. *Emerg Infect Dis* **2014**, *20*, 21-28, doi:10.3201/eid2001.130652.
21. Challacombe, J.F.; Pillai, S.; Kuske, C.R. Shared features of cryptic plasmids from environmental and pathogenic *Francisella* species. *PLoS One* **2017**, *12*, e0183554, doi:10.1371/journal.pone.0183554.
22. Petersen, J.M.; Carlson, J.; Yockey, B.; Pillai, S.; Kuske, C.; Garbalena, G.; Pottumarthy, S.; Chalcraft, L. Direct isolation of *Francisella* spp. from environmental samples. *Lett Appl Microbiol* **2009**, *48*, 663-667, doi:10.1111/j.1472-765X.2009.02589.x.
23. Rubel, F.; Brugger, K.; Belova, O.A.; Kholodilov, I.S.; Didyk, Y.M.; Kurzrock, L.; Garcia-Perez, A.L.; Kahl, O. Vectors of disease at the northern distribution limit of the genus *Dermacentor* in Eurasia: *D. reticulatus* and *D. silvarum*. *Exp Appl Acarol* **2020**, *82*, 95-123, doi:10.1007/s10493-020-00533-y.
24. Rubel, F.; Brugger, K.; Walter, M.; Vogelgesang, J.R.; Didyk, Y.M.; Fu, S.; Kahl, O. Geographical distribution, climate adaptation and vector competence of the Eurasian hard tick *Haemaphysalis concinna*. *Ticks Tick Borne Dis* **2018**, *9*, 1080-1089, doi:10.1016/j.ttbdis.2018.04.002.
25. Hasle, G. Transport of ixodid ticks and tick-borne pathogens by migratory birds. *Frontiers in Cellular and Infection Microbiology* **2013**, *3*, doi:10.3389/fcimb.2013.00048.
26. Guo, W.-B.; Shi, W.-Q.; Wang, Q.; Pan, Y.-S.; Chang, Q.-C.; Jiang, B.-G.; Cheng, J.-X.; Cui, X.-M.; Zhou, Y.-H.; Wei, J.-T.; et al. Distribution of *Dermacentor silvarum* and Associated Pathogens: Meta-Analysis of Global Published Data and a Field Survey in China. *International Journal of Environmental Research and Public Health* **2021**, *18*, 4430.
27. Mostafavi, E.; Ghasemi, A.; Rohani, M.; Molaeipoor, L.; Esmaeili, S.; Mohammadi, Z.; Mahmoudi, A.; Aliabadian, M.; Johansson, A. Molecular Survey of Tularemia and Plague in Small Mammals From Iran. *Frontiers in Cellular and Infection Microbiology* **2018**, *8*, doi:10.3389/fcimb.2018.00215.
28. Timofeev, V.; Titareva, G.; Bahtejeva, I.; Kombarova, T.; Kravchenko, T.; Mokrievich, A.; Dyatlov, I. The Comparative Virulence of *Francisella tularensis* Subsp. *mediasiatica* for Vaccinated Laboratory Animals. *Microorganisms* **2020**, *8*, 1403.
29. Buczek, A.; Bartosik, K.; Buczek, W.; Buczek, A.M.; Kulina, D.; Kulisz, J.; Tomasiewicz, K. A unique phenomenon of oral-anal contact between ticks observed in two tick species *Ixodes ricinus* and *Dermacentor reticulatus*. *Annals of Agricultural and Environmental Medicine* **2018**, *25*, 686-689, doi:10.26444/aaem/99054.
30. Buczek, A.; Bartosik, K.; Buczek, A.M.; Buczek, W.; Stanko, M. Conspecific hyperparasitism in the *Hyalomma excavatum* tick and considerations on the biological and epidemiological implications of this phenomenon. *Ann Agric Environ Med* **2019**, *26*, 548-554, doi:10.26444/aaem/110128.