Insights from *Bacillus anthracis* strains isolated from permafrost in the tundra zone of Russia

A short title: *Bacillus anthracis* strains from the tundra zone of Russia

Timofeev V[1*]., Bahtejeva I[1]., Mironova R[1]., Titareva G[1]., Lev I[1]., Christiany, D[2]., Borzilov A[1]., Bogun A[1]., Vergnaud G[2**].

1 - State Research Center for Applied Microbiology & Biotechnology» (FBIS SRCAMB), Russia

2 - Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

* corresponding author

E-mail: timofeev_vitalii@mail.ru

** corresponding author

E-mail: gilles.vergnaud@u-psud.fr

Keywords: *Bacillus anthracis*; tundra region; MLVA-genotyping; whole genome SNP analysis

Abstract

This article describes *Bacillus anthracis* strains isolated during an outbreak of anthrax on the Yamal Peninsula in the summer of 2016 and independently in Yakutia in 2015. A common feature of these strains is their conservation in permafrost, from which they were extracted either due to the thawing of permafrost (Yamal strains) or as the result of paleontological excavations (Yakut strains). All strains isolated on the Yamal share an identical genotype belonging to lineage

B.Br.001/002, pointing to a common source of infection in a territory over 250 km in length. In contrast, during the excavations in Yakutia, three genetically different strains were recovered from a single pit. One strain belongs to B.Br.001/002, as the Yamal strains. Despite the remoteness of Yamal from Yakutia, whole genome sequence analysis showed that the B.Br.001/002 strains are very closely related. The two other strains contribute to two different branches of A.Br.008/011, one of the remarkable polytomies described so far in *B. anthracis* population. The geographic distribution of the strains belonging to this polytomy is suggesting that this polytomy emerged in the thirteenth century, in combination with the constitution of a unified Mongol empire extending from China to Eastern Europe. We propose an evolutionary model for *B. anthracis* recent evolution in which the B lineage spread throughout Eurasia and was subsequently replaced by the A lineage except in some geographically isolated areas.

Funding. This work is supported by the Sectoral Scientific Program «Problem-oriented research in the epidemiological surveillance of infectious and parasitic diseases (2016-2020)» of the Russian Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing (Rospotrebnadzor) (http://www.rospotrebnadzor.ru/) and by the French Agence Nationale de la Recherche grant ANR-14-ASMA-0002-02 “Microtype”.

Introduction

The etiological agent of the anthrax disease is the gram-positive bacterium *Bacillus anthracis*. A key feature of this microorganism, which largely determines its epidemiological potential and population structure, is the ability to form endospores, extremely resistant to adverse environmental factors and able to remain viable for a long time [1-12]. High preservation of spores explains that even in regions where this disease has not been observed for decades, disease outbreaks are possible, leading to significant economic damage, the mass mortality of livestock, and human victims. In addition, due to the high virulence of *B. anthracis*, the stability of endospores in the environment and the simplicity of cultivation, this bacterium is considered a

potential biological weapon or tool for bioterrorism [13, 14], as illustrated by the anthrax contaminated letters in 2001 in the USA [15].

Anthrax is now very rare in most European countries [16] but remains a significant problem mainly in sub-Saharan Africa and in some regions of Asia [17-20]. Anthrax is endemic in Russia, where the disease manifests itself as sporadic cases among animals and rare cases of the disease among the population [21]. The presence of large territories hosting populations of wild and domestic ungulates creates a favorable context for disease outbreaks of epizootics, and the low human population density in most parts of the country makes it difficult to conduct anti-epidemic measures, and to correctly account for anthrax animal burial sites. Past animal burial sites are often not documented and occasionally corpses were not buried. These burial grounds, and entire territories, where previously epizootics took place, may become involved in increased economic activities, which, given the potentially high preservation of *B. anthracis* spores in a cold environment, could lead to new outbreaks of the disease [22]. Of particular interest in this regard is the tundra zone of Russia, located between 55 and 68 degrees north latitude.

One of the features of this climatic zone is the presence of permafrost. Permafrost is defined as lithosphere material (soil and sediment) permanently exposed to temperatures ≤0 °C and remaining frozen for at least two consecutive years. Permafrost can extend down to more than 1000 m in depth and remain frozen for thousands of years [23, 24]. In permafrost conditions, the preservation of microorganisms can significantly increase, and thus permafrost is a peculiar accumulator of microbiota [25, 26]. The preservation of spores of bacilli, including *B. anthracis*, in the permafrost theoretically should significantly exceed the preservation of microorganisms in the vegetative form. Consequently, permafrost might allow the discovery of archaic forms of this microorganism, which could supplement our knowledge of the evolution of the anthrax microbe. We investigated strains of *B. anthracis* isolated in two tundra zones - during the outbreak of anthrax in Yamal in the summer of 2016 and during extraction of paleontological material from the permafrost in Yakutia in 2015.

Materials and Methods

Yamal samples

In the summer of 2016 an outbreak of anthrax occurred on the Yamal Peninsula. The previous outbreak of anthrax was registered in 1941 and the district was officially declared "anthrax-free" territory of the USSR since 1968. In 2007, the compulsory vaccination of reindeer was abandoned. On July 16th 2016, the United Duty Control Service of the Yamal District was informed of deer death by private reindeer herders. The deer's deaths began at the estuary of the Nerosaveyakha River near Lake Pisyoto. Reindeer herders reported that sick animals became sluggish, began to move slowly, then fell and quickly died. No ulcers or skin lesions could be detected. On July 17th and 18th, the Veterinary Service of the Yamal-Nenets Autonomous Area arrived in the area for clinical examination of animals and autopsy. Pathological material was sent to the Tyumen Regional Veterinary Laboratory. An autopsy showed cardiac and pulmonary insufficiency and the preliminary diagnosis was death from a heat stroke as July 2016 had anomalously hot weather, with temperatures above 35°C. Complementary investigations by veterinarians sent to reindeer herders camps on July 19-29 led the Tyumen Regional Veterinary Laboratory to report a suspicion of *B. anthracis* on July 24th and to take prophylactic measures (vaccination and chemotherapy, restrictions on animal movements). Additional samples were sent to the All-Russian Scientific Research Institute of Veterinary Virology and Microbiology (ARSRIVVM) in Pokrov (Moscow region).

By this time, the disease was observed in three focuses: Lake Pisyoto area, Novoportovskaya tundra, Evayakha River area. The outbreak sites were separated by distances up to 250 km, including two water barriers - the Gulf of Ob (width from 30 to 80 km) and the Taz Estuary (average width is about 25 kilometers).

On July 25th, a complete laboratory confirmation of presence of *B. anthracis* in samples taken from dead deer was obtained by ARSRIVVM. A pure culture of *B. anthracis* was isolated

from one sample, this strain was called 5875. The Governor of Yamal-Nenets Autonomous District introduced a quarantine regime in the Yamal district. SRCAMB's and ARSRIVVM's employees went to Salekhard to sample and to consult local sanitary and medical institutions. Specialists of the Stavropol Anti-Plague Institute arrived on July 26th.

Starting on July 26, arrived experts organized a diagnostic laboratory in the Center for Hygiene and Epidemiology in the Yamal-Nenets Autonomous District. During the outbreak, samples from people potentially infected were investigated by this diagnostic laboratory. Medical authorities decided to hospitalize to Salekhard all children from the outbreak areas even without visible signs of the disease. People evacuation to temporary camps equipped by that time were started and preventive antibiotic therapy was applied. SRCAMB's experts flew from Salekhard to the disease focus in the Lake Pisyoto area to survey the local population and to collect samples. Until this time, both the local population and veterinarians working in this outbreak area were skeptical about the possibility of anthrax, and favored the heat shock hypothesis as a number of other infections harmless for humans could have caused the death of animals weakened by heat. Furthermore simultaneously with the beginning of vaccination and antibiotic therapy, the temperature of the infection focus decreased sharply, so there were reasons to believe that the cessation of new cases of the disease resulted from the lowering of temperature.

The typical development of the disease was as follows: a seemingly healthy deer would become suddenly weak, unable to walk and forced to lay down a few hours later, and would die after a few additional hours. In most cases the nose was bleeding (sometimes the anus too), rigor mortis developed at the usual time.

SRCAMB's experts took samples of soil, water, blood samples, ears, and lymph nodes of dead deer. The samples were delivered to Salekhard on July 27th, and eventually to SRCAMB on July 28th. On July 29th, strain 5875 was sent from ARSRIVVM to SRCAMB. On the same day, SRCAMB received a strain isolated from a sick person (washed off from skin infection) subsequently called Yamal_12 and insects caught by veterinarians working in the outbreak area:

nine *Scopeuma stercorarium* and four *Hydrotaca dentipes* from Salekhard. On August 13th, SRCAMB received strain 6063 isolated in epidemic area by ARSRIVVM.

Yakutia samples

On August 12, 2015, miners extracting mammoth tusks from the permafrost on the bank of the river Uyandina in the Abyisk ulus (district) of Yakutia 57 km from the district center Belaya Gora ("White Mountain") (latitude N (68.564567), longitude E (144.769827)) found two kittens of the cave lion *Panthera leo spelaea* frozen in ice. The discovery was remarkable by its unprecedented degree of preservation - the animals preserved wool and soft tissues. The bodies of the kittens were transferred to paleontologists. Some samples were taken for microbiological examination to Institute of Oil and Gas of the Siberian Branch of the Russian Academy of Sciences in Yakutsk (IPMR SB RAS), the nearest scientific institution. June 01, 2016 an unknown bacillus-like strain was isolated in the laboratory of geochemistry of caustobioliths of IPMR SB RAS and sent to the Institute of Genetics and Selections of Industrial Microorganisms (GosNIIgenetika) in Moscow for identification. September, 01 this strain was identified as *B. anthracis* according to PCR results. Since this institute is not equipped to carry out work with pathogenic microorganisms and has no appropriate license, this initial culture was destroyed. September, 21 according to order of the Chief State Sanitary Doctor of Yakutia, soil samples were collected in paleontological discovery point. Six separately packed glass jars with soil samples (200 g each), taken from a depth of 1 to 6 meters with an interval of 1 meter were received by SRCAMB in December 2016.

Animal experiments

Ethics statement

All protocols for animal experiments were approved by the State Research Center for Applied Microbiology and Biotechnology Bioethics Committee (Permits No: VP-2016/4 and VP-2016/5). They were performed in compliance with the NIH Animal Welfare Insurance #A5476-

01 issued on 02/07/2007 and the European Union guidelines and regulations on handling, care, and protection of laboratory animals (http://ec.europa.eu/environment/chemicals/lab_animals/home_en.htm).

All used animals were purchased from Laboratory Animals Breeding Center, Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russia. They were housed in polycarbonate cages with space for comfortable movement (5 mice per 484 cm^2 cage or 2-3 guinea pigs per 864 cm^2 cage) and easy access to food and water, under constant temperature (22°C ± 2°C) and humidity conditions (50% ± 10%) and a 12-hour light/12-hour dark cycle.

Approved protocols provided scientifically validated humane endpoints, including pre-set criteria for euthanasia of moribund animals by CO_2 inhalation. Animals were euthanized when they became lethargic, dehydrated, moribund, unable to rise, or non-responsive to touch. Surviving animals were euthanized after the observation period.

Mice

Six-to-eight-weeks-old BALB/C mice of both genders, weighing 18–20g were used in all our experiments. They were fed Mouse Mixed Fodder PK-120 (Laboratorkorm, Russia) and provided tap water *ad libitum* throughout observation period.

For virulence evaluation mice were randomly divided in four groups of ten animals and infected s.c. by doses:10^0, 10^1, 10^2 and 10^3 spores/animal. They were observed for 30 days after infection.

For bioassay we used groups of three mice for each tested sample. Mice were inoculated subcutaneously with sample dispended in 0.3 ml of PBS in the inner part of the upper thigh. Animals were observed during 10 days, after which surviving mice were euthanized. Dead and euthanized mice were necropsied, and samples of their spleens and livers were inoculated on Petri dishes with selective anthrax agar (SRCAMB).

Guinea pigs

Guinea pigs were used for evaluating the virulence of strains isolated in Yamal epidemic area. We used five-to-seven-weeks-old animals of both genders, weighing 350–450g. They were fed Granuled Fodder KK-122 (Laboratorkorm, Russia) and provided tap water *ad libitum* during the entire experiment (30 days). Guinea pigs were randomly divided in three groups of five animals and infected s.c. by doses of 10^2, 10^3, 10^4 spores/animal.

Bacterial culture and DNA extraction

For bacterial cultures isolation, cultivation for DNA extraction, and for lecitinase and hemolytic activity evaluation we used GRM agar, selective anthrax agar, yolk agar and blood agar, manufactured in SRCAMB, Russia.

DNA from field and clinical samples was isolated using «Reagent kit «K-Sorb» for DNA extraction on microcolumns» (Syntol, Russia). DNA from bacterial cultures was isolated using GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich, USA).

PCR analyses

PCR amplifications were run on the CFX96 ™ Real-Time PCR Detection System (Bio-Rad Laboratories, Inc, USA). For MLVA and canSNP genotyping we used 2.5× PCRmix M-427 with SYBR-GreenI (Syntol, Moscow, Russia). Primers were synthesized by Syntol, Russia.

PCR detection of *B. anthracis* DNA was performed using Real-Time PCR-test systems «MULTI-FLU» (SRCAMB, Obolensk, Russia) and «OM-screen-anthrax-RT» (Syntol, Moscow, Russia).

MLVA was performed using primers as indicated in Thierry et al. [27], but using monoplex PCR. PCR products size was evaluated using agarose gel-electrophoresis. The PCR products and a 20 bp ladder (Bio-Rad, USA) were electrophoresed at 100 V for 240 min on a 32-cm length 3% agarose gel prepared in 0.5× TBE. The DNA fragments were visualized with ethidium bromide staining and ultraviolet (254 nm) using the Doc-Print gel documenting system and PhotoCaptMw software version 99.04 (Vilber Lourmat, Marne-la-Vallée, France). PCR products larger than 600 bp were reanalyzed on 2% agarose gel for better resolution. Also in these few cases we confirmed

the size of amplicon using Experion™ Automated Electrophoresis System (BioRad, Hercules, USA) and by sequencing the fragment.

canSNP-genotyping was performed as described in [28].

Whole Genome SNP Analysis

Yamal strains DNA was sequenced using the Ion Torrent PGM (Life Technologies, USA). Ion PGM Reagents 400 Kit (Life Technologies, USA) and Ion 318 Chip Kit (Life Technologies, USA) were used for sequencing. For each genome, reads were de novo assembled using 2.9 Newbler assembler (Roche).

Yakutia strains whole-genome sequencing was performed using the Illumina MiSeq instrument. DNA libraries were prepared using Nextera DNA Library Preparation Kit. Miseq Reagent Kit v3 was used for sequencing. For each genome, reads were assembled de novo using SPAdes v. 3.9 (http://bioinf.spbau.ru/spades).

Additional sequence read archives were recovered from the European Nucleotide Archive (ENA) using the enaBrowserTools (https://github.com/enasequence/enaBrowserTools) and the Aspera (Aspera, Inc., USA) high-speed file transfer protocol. Genome assemblies were downloaded from NCBI and *in silico* converted into 100 bp. sequence reads files with a 50x coverage using a homemade python script. Sequencing reads were mapped on reference genome *B. anthracis* Ames ancestor assembly GCA_000008445 using BioNumerics version 7.6.3. SNPs were called within BioNumerics using the strict closed dataset option. Minimum spanning trees were produced allowing the creation of hypothetical missing links.

Results

Investigation of the Yamal samples

PCR-diagnostic of soil, water and necropsy samples showed that all samples (n=23) except soil from the reindeer herding camp (n=5) contained DNA of *B. anthracis*. Set of nutrient medias - GRM agar, selective anthrax medium, yolk agar, blood agar was inoculated with materials from

investigated samples. Typical, *B. anthracis*-like colonies grew from all PCR positive samples, none from PCR-negative samples. Microscopic studies of these colonies showed the presence of chains of gram-positive bacillus coated with a capsule. All these bacillus-like colonies were PCR positive for *B. anthracis*. Some colonies of extraneous microflora grew from these samples. All insects samples were PCR-negative, and no colonies of *B. anthracis* or extraneous microflora could be recovered. This negative result is likely due to the use of ethanol for better preservation of entomological specimens. Consequently, we were unable to confirm or disprove the potential role of bloodsucking insects in the spreading of the disease over long distances.

Soil from the place of death, soil from the camp, water from a nearby lake, cervical lymph node, blood from the neck region, blood discharge from the anus, and all the clinical samples were used in a bioassay. All the mice, (except mice infected by soil from the camp) died with symptoms of anthrax on the second day after infection, their spleen and injection sites contained live anthrax bacteria as shown by inoculation on Petri dishes.

MLVA genotyping was applied in order to establish whether several genotypes circulated in the epidemic zone. We initially used loci vrrA, Bams03, Bams05, Bams22, Bams44, and vntr23 derived from the MLVA7 scheme proposed by Thierry et al. [27]. This set of loci allowed to genotype all isolates during one day with a high degree of reliability and proved to be very useful in conducting an epidemiological investigation, when it is required to minimize the time spent on analysis, down to hours. Later it turned out that the same MLVA profile was found in all strains isolated in the summer of 2016 on the Yamal, regardless of the isolation place (Lake Pisyoto area, Novoportovskaya tundra, Evayakha River area) and of the institution by which the samples were analyzed (S1 table) even when using 25 loci. Querying the *Bacillus anthracis* v4_0 MLVA database at http://microbesgenotyping.i2bc.paris-saclay.fr [27] indicates that the Yamal strains belong to the B-clade [29]. All genotypes present in the MLVA database differed at four loci or more among the 25 loci. canSNP-genotyping using melt-curves [28] assigned all Yamal strains to B.Br.001/002 lineage in agreement with the MLVA genotype.

The finding of a unique MLVA25 profile suggested that one strain circulated throughout the epidemic, possibly from a unique source of infection of humans and animals. No strains in collections of SRCAMB and Stavropol Anti-Plague Institute (Reference Center for the control of Anthrax) showed the same MLVA25 profile.

We selected four strains collected in Lake Pisyoto outbreak area for further work including whole genome sequencing (table 1).

Table 1. *B. anthracis* strains isolated from Yamal outbreak field and clinical samples

Strain name	**Source**
Yamal_2	Cervical lymph node of dead deer
Yamal_8	Water from a lake
Yamal_10	Soil near a dead deer
Yamal_12	Cutaneous carbuncle of a sick person

The strains listed in table 1 were typical for a combination of phenotypic properties, but the clinical isolate Yamal_12 initially had lecithinase activity, unlike the other isolates. After two passages on solid media, this activity was lost.

Investigation of the Yakoutia samples

Several typical *B. anthracis* colonies were cultivated from soil samples from a depth of 2, 3 and 4 meters. The colonies were confirmed as *B. anthracis* by MLVA analysis at seventeen loci (MLVA17). Three genotypes, subsequently called 3Ya, 4Ya, 5Ya could be distinguished (see S1 table). Genotype 4Ya was equally represented in samples from a depth of 2 and 3 m (Table 2). Genotypes 3Ya and 4Ya differ at four loci but querying the *B. anthracis* MLVA database indicates that both are closest to strains belonging to canSNP clade A.Br.008/009. MLVA17 genotype 5Ya is identical to the Yamal MLVA genotype.

Table 2. Yakutia soil samples investigations and strains selected for whole genome sequencing

Depth	**Detected MLVA genotypes**	**Representative Strain name_genotype**
1m	none	NA*
2m	3Ya, 4Ya	LP50_3Ya

3m	4Ya	LP51_4Ya
4m	5Ya	LP53_5Ya
5m	none	NA
6m	none	NA

*NA - not applicable

Whole genome SNP analysis

Contrary to our expectations, the recovered MLVA genotypes are very similar to already know genotypes. Furthermore, the Yamal MLVA genotype is identical to one of the three Yakut strains genotypes. Finally, we recovered three different MLVA genotypes from the same spot in Yakutia, whereas the Yamal 2016 outbreak was associated with a unique genotype in spite of its wide geographic dispersion. One simple explanation for this surprising observation from the Yakut excavation could be contamination when analyzing the samples in SRCAMB. The resistance of the endospores is a well-known cause of accidental laboratory contaminations as recently recalled [30]. In order to investigate this possibility we sequenced four Yamal (Table 1) and three Yakut (Table 2) strains, together with the four strains in the SRCAMB collection showing the closest MLVA genotype. We performed a whole genome SNP analysis comparison of these ten genomes and also included archive reads and assemblies downloaded from EBI-ENA. We used *B. cereus* strain ISSFR-23F representative of one *B. cereus* lineage closest to *B. anthracis* as outgroup to root the tree [31]. Figure 1 shows the relative position of the Yamal and Yakut strains within the global *B. anthracis* population represented by a selection of 50 strains among 650. The longest genetic distance links the MRCA of the *B. anthracis* species and the *B. cereus* outgroup. The position of the ancestor of the *B. anthracis* species along this branch is unknown. The four Yamal strains are identical and belong to B.Br001/002 in agreement with the canSNP typing. The Yakut Ya5 strain is the closest neighbor but is clearly distinct.

Fig 1. Position of the Yamal and Yakutia strains in the global *B. anthracis* phylogeny. Strains representing the main lineages including the different canSNP lineages were selected. Red star: the tree is rooted with *B. cereus* strain ISSFR-23F. Each circle is labelled with the

corresponding strain name. The color code reflects the canSNP lineage. The very rare lineages defining the most ancient currently known splits have been found only in North America. The Yamal and Yakutia strains are arrowed. The number of SNPs constituting each branch is indicated. A logarithmic scaling was used in order to visualize the shorter branches.

Figure 2 is a close-up on the group including all available genome sequences assigned to B.Br.001/002 and sub-lineage B.Br.Kruger. B.Br.001/002 is split in two parts, one part including the B.Br.Kruger sublineage predominant in South Africa and the other part including strains from Northern Russia, Estonia and Korea. The "Kruger" group is characterized by relatively long branches. From the split to the tips of the tree, the "Kruger" group expansion ranges from 256 to 368 SNPs. In contrast, in the same timespan the "Eurasia" group expanded by 60 up to 86 SNPs, i.e. a ratio of 4.25 between the two groups. This observation may suggest that the "Kruger" clade is the result of a secondary introduction to South Africa. The Yakut Ya5 and Yamal strains belong to the "Eurasia" group and are separated by 54 SNPs.

Fig 2. Focus on the B.Br001/002 and B.Br.Kruger lineages. The Ames ancestor reference genome is used to root the minimum spanning tree. Branch length (number of SNPs) is indicated. Geographic origin of each strain is displayed when known.

The two other Yakut strains, Ya3 and Ya4, belong to the Transeurasian radiation TEA 008/011 [32]. TEA 008/011 is a remarkable polytomy currently including seven lineages. Figure 3 displays a minimum spanning tree based on the whole genome SNPs detected among the 75 strains assigned to TEA 008/011. The branches are named as previously proposed [32]. Distances from root to tips differ by a ratio of 11 among the different lineages. The shortest branch with 17 SNPs represented by a strain from China, is observed in lineage L2_STI. The longest, with 194 SNPs, is observed in L1_Heroin. The most represented lineages are L1_Heroin, L2_STI and

L3_Tsiankowskii. The L1_Heroin lineage was previously investigated in detail [33-35]. L1_Heroin contains one early split, three SNPs from the core of the polytomy. One branch with 75 SNPs is present in China, whereas the other branch is more complex in terms of geographic origin as it was isolated in many countries. The shortest branch (57 SNPs from red-star root to tip in Figure 3) within this lineage is defined by a cluster of strains recovered in different European countries from human patients infected via drug usage. The longest branch is the outcome of a recent split along the shortest branch. After this split, the relative rate of expansion was 28. The difference in length is not the result of horizontal gene transfer events, as the associated SNPs do not show evidence for clustering [36]. The geographic location of the reservoir is uncertain. Afghanistan is a likely option if heroin contamination occurred as part of the drug production or initial packaging processes. Other candidate countries are Turkey and Pakistan represented by strains defined by short lineages, or additional neighbor countries not represented in current *B. anthracis* databases. The L2_STI lineage contains three deep-branching sublineages. The first is defined by the Yakut 4Ya strain, the second is present mostly in China, and the third corresponds to the STI vaccine strain from Russia. The Tsiankowskii vaccine strain, the Sverdlosk 1979 strain [34], and the Yakutia 3Ya strain which is closest to a strain from Norway belong to lineage L3_Tsiankowskii lineage containing a single deep-branching lineage widely spreading in Russia and Eastern Europe, including Greece, Albania, Bulgaria, Poland. Lineage L4_Pasteur contains two deep-branching sublineages, one found in Bulgaria and the other in Italy in addition to the Pasteur I vaccine strain. Lineages L5 and L6 each contain a single deep branching lineage, observed in Turkey. The last lineage leads to TEA Br011 corresponding to the A.Br.011/009 polytomy [37, 38] found in France. In summary, eleven deep-branching lineages are detected within the seven-branch TEA 008/011 polytomy, nine of which with a strong geographic assignment, to Turkey (2), China (2), Russia (2, including Northern Yakutia, Siberia), Italy, Bulgaria and France. The root (red star) defined by the branching point of the Ames ancestor reference strain is located on L3_Tsiankowskii at a distance of one SNP from the center of the

polytomy connecting the six other branches. This might suggest that Europe is the geographic origin of the A.Br.008/011 polytomy. However this argument is weak and extensive sequencing of additional A.Br.008/011 strains will be needed to establish the geographic origin of the polytomy.

Fig 3. A.Br008/011 polytomy, minimum spanning tree, logarithmic scale. The color coding reflects the lineages within the A.Br.008/011 polytomy using the previously defined classification [32]. Geographic origin is indicated when known. The root (red star) is defined using the Ames ancestor reference genome. Two strains representing the A.Br011/009 lineage are included. The tree is based upon 1844 SNPs, and the level of homoplasia is 0.7% (the size of the tree is 1858).

Two representatives from A.Br.011/009 were included in Figure 3, in order to show the position of the root of the A.Br.011/009 polytomy located at a distance of six SNPs from the root of the A.Br.008/011 polytomy. Figure 4 shows the structure of the A.Br.011/009 polytomy, using the fifty-four read archives or genome assemblies assigned to A.Br.011/009. The polytomy comprises six branches numbered L1 to L6 in agreement with previous reports [37, 38]. The monophyletic A.Br.WNA lineage exclusively present and predominant in North-America emerged from A.Br.011/009_L2 and is represented by four strains in Figure 4 [28, 38-40]. Interestingly, the WNA lineage is branching from the Canadian strain A0303 which shows the ancestral state for the A.Br.WNA canSNP, i.e. still belongs to A.Br.011/009 [39]. This is in agreement with the report by Kenefic et al. demonstrating an introduction of WNA in the USA from Canada, and a progressive evolution from north to south. The West Africa clade from Guinea and Côte d'Ivoire [41] is branching out from lineage L2 at the same position as the WNA and Senegal-Gambia [42] clade. Apart from two Argentinian strains belonging to the L3_Pasteur II vaccine sublineage and one strain from the USA defining a recent split with a long branch within lineage L4, all the other

A.Br.011/009 strains are from Italy and France. Italian strains define two deep lineages within lineages L1 and L2, which split at a distance of one SNP from the root of the A.Br.011/009 polytomy. After the split, the length of the expansion is very similar along the French and Italian sublineages (Fig 4). For instance the total length of the French L2 lineage is 26-33 SNPs, compared to 24-28 SNPs in the Italian L2 lineage.

Fig 4. A.Br011/009 polytomy, minimum spanning tree, logarithmic scale. The color coding reflects the lineages within the A.Br.011/009 polytomy using the previously defined classification [38]. Within L3, strains derived from the Pasteur II vaccine are colored in pink. Geographic origin is indicated when known. The root (red star) is defined using the Ames ancestor reference genome.

Consequently, the analysis of currently available A.Br.011/009 sequencing data strengthens the previously reported view of the A.Br.011/009 polytomy, characterized by a limited geographic distribution of the "slowly evolving" lineages contrasting with the very fast expanding lineages observed in West Africa and North America [38]. After the North American-West African split shown in Figure 4, French strains from lineage L2 expanded for a further 15 to 22 SNPs when not including two strains obtained from the Collection de l'Institut Pasteur (CIP) which might have been extensively cultivated and define slightly longer branches. In contrast, the length observed towards Côte d'Ivoire-Guinea, Senegal-Gambia, and WNA are 167-177 SNPs, 323-458 SNPs and 112-142 respectively, i.e. ratios of 8 to 20. The currently available data provides evidence for two independent introductions of *B. anthracis* from the A.Br.011/009 polytomy in North America, one from the L2 lineage via Canada, and the second one from the L4 lineage. The second introduction is represented by a unique read archive SRR5811139 derived from a strain collected in New Hampshire. This second introduction occurred after L4 had already expanded for at least 2/3 of its

current length whereas the first introduction from L2 occurred after L2 had expanded for 1/3 of its current length.

Discussion

The Yamal 2016 anthrax outbreak and implications

Anthrax is endemic in most of Russia, including Yamal. At the beginning of the Yamal colonization by the Russian Empire in the 17-18 centuries, cases of this disease were reported. The first outbreak was recorded in 1760. From 1898 to 1931, 66 epizootics were described, during which more than a million deer died. In the 1940s, the whole reindeer livestock was vaccinated. In subsequent years, the number of vaccinated animals was lower, for example, in the 1960s, from 65 to 82% of the total number of deer was vaccinated. This proved sufficient to prevent epidemics. Thus, a centuries-old cycle of anthrax circulation in the tundra of the Far North was interrupted [43].

The absence of outbreaks suggested that the soil in the tundra was sanitized and no longer contained anthrax spores. In 1968, 360 soil samples from places of recorded mass death of reindeers were examined and no *B. anthracis* strains could be recovered [44]. This suggested that tundra soil conditions (pH 3-5, humus content lower than 3%) are unfavorable to maintain the viability of the spores. In 2007, deer vaccinating was stopped. During June-July of 2016, the air temperature in the Yamal epidemic area in was 5-9 degrees higher than normal, and did not fall below 18 C. The soil reached a temperature of 25 C at a depth of 10 cm and 7 C at a depth of 1 meter. This was combined with a very small amount of rain precipitation [45]. Apparently such an anomalous warm climate led to the thawing of permafrost, and viable *B. anthracis* spores could be exposed to the surface [3, 46].

According to the testimony of reindeer herders in epidemic area near Pisyoto lake, herds from all focus of infection had been driven through the same place in the tundra. The thawing of permafrost provoked a landslide of a hill on the bank of the river. This could explain the finding

426 of a single MLVA and wgSNP genotype. Unfortunately since all the helicopters in the region were
427 being used to transport people and cargo, it was not possible to visit and sample the place (in this
428 region roads are absent; movement is possible only by sledges with a team of reindeers or by
429 helicopter). Along these lines, two migration ways of the pathogen might be proposed: washing
430 out of bacterial cells from deep soil layers to the ground surface, or exposure of a deep soil layer
431 to the surface due to thawing and landslide. At the same time, the reindeers were weakened by the
432 heat, which could have increased susceptibility to infection. Observations in the focus of the
433 disease and a survey of veterinarians and reindeer herders gave grounds to assume that infection
434 could occur not only by spores but also by vegetative cells. In several cases, a sick reindeer could
435 start recovering and stand on its legs after receiving a single dose of antibiotic, subsequently licked
436 the muzzle of healthy animals, which fell ill and died within 24 hours.

437 Consequently, we cannot exclude a simultaneous spreading of infection from several
438 isolated foci triggered by the exceptionally hot weather. A preceding large-scale epizootic, spread
439 widely in the region and conserved in permafrost in multiple soil foci might explain the observed
440 genetic homogeneity of the strains collected during the present outbreak. In our opinion, this
441 alternative is not the most likely. Unfortunately, strains from previous outbreaks were not
442 preserved in collections. Therefore, there is no way to compare the strain isolated in 2016 with
443 strains previously circulating in the region, and, accordingly, to precisely estimate the length of
444 time during which the pathogen spores were stored in the soil. Given that the last outbreak of
445 anthrax in Yamal was registered in 1941, it is likely that the spores remained viable in the soil for
446 at least 75 years.

447 Considering the scale of the epidemic and the costs of countering it, the time interval
448 between the beginning of reindeer disease and the beginning of antiepidemic measures may seem
449 too large. In the case of a faster response, the epidemic could be less severe. Several factors played
450 a role in this situation:

- Presence of a large number of unvaccinated reindeer (Yamal reindeers population reaches 800 thousand), susceptible to infection and weakened by heat).

- Lack of experience in anthrax diagnosing - the last outbreak of anthrax was recorded in this region in 1941, and local veterinarians had no experience of this infection before 2016.

- Nomadic mode of cattle breeding - even a small herd can eat the limited tundra vegetation very quickly, thus breeders have to drive the herd to another place. Nomadic cattle breeding cover much larger territories than settled cattle breeding. The migration routes of different herds may cross each other. In the case of anthrax, this can rapidly increase the epidemic area, and contribute the infection to herds that migrate through the territories where sick animals were driven.

Also the establishment of the scale of the accident and the timely examination of the corpses was hampered by the fact that reindeer herders migrated and could not always bring dead deer for inspection. The delay in identifying patients with anthrax was also due to lack of awareness of the local population about the dangers of anthrax and about its symptoms. Local populations, even with symptoms of the cutaneous form of this disease, did not pay attention to them and considered themselves healthy. This situation arose from a set of reasons. The traditional way of life of reindeer breeders, transport and communication isolation of nomadic reindeer herders from cities is associated with a more limited access to medical care. The usual high prevalence of furunculosis prevented timely detection of anthrax affections on the skin.

All these factors favored the spreading of the infection before medical and veterinary organizations were alerted. During a hidden period, the infection may have been carried by transport (helicopters and ships), along with the goods and people who visited the outbreak, where the animals infection happened first. Unfortunately, it is not possible to retrace retrospectively the movement of people and transport between the foci of infection.

The occurrence of anthrax outbreaks after a 75-year break demonstrates that the decision to stop the vaccination was premature, in line with similar independent observations in Georgia [47]. Permafrost turned out capable to conserve viable microbial spores for a long period and

figuratively speaking to be a reservoir of infection. Under favorable climatic conditions, these spores proved able to migrate to the surface of the soil and initiate new infection cycles.

If the Yamal strains were isolated during a large-scale epidemic, then the finding of the Yakut strains in a seemingly random place is very noteworthy. In the absence of historic record of anthrax outbreak in this area, there was no reason to expect the finding of *B. anthracis* spores in the soil. The microbiological investigation was prompted by the paleontological finding of cave lions leading to the serendipitous finding of *B. anthracis*. Three different genotypes from soil samples taken from a depth corresponding to the upper layers of permafrost down to four meters were characterised. One genotype was very close to the Yamal genotype and belongs to the B cluster of *B. anthracis* whereas the two others belong to the A.Br.011/008 polytomy.

Tentatively dating the emergence of the A.Br.011/008 polytomy.

The monophyletic, strictly clonal evolution of *B. anthracis* implies that the whole species derives from a unique progenitor [48]. Africa is sometimes proposed as being the "cradle" of *B. anthracis* [49]. The strongest argument in favor of such an origin may be the discovery in Africa of additional *B. cereus* lineages causing an anthrax-like disease [48, 50, 51]. However, this does not help dating or locating the origin of modern lineages, even when these lineages display a strong geographic preference. A tentative dating of the Most Recent Common Ancestor (MRCA) of the *B. anthracis* species to 13,000-27,0000 years was previously proposed based on average mutation rate and estimates of infection cycles per year [28]. The *B. anthracis* species may be much older than the MRCA defined by known lineages.

The dating of nodes along the *B. anthracis* phylogeny is particularly difficult and challenging because of the irregularity of its evolution due to its ecology [28, 32, 39]. In contrast with other pathogens such as *Mycobacterium tuberculosis* [36, 52, 53], branch length does not correlate with time, as investigated and discussed in detail by [32]. Rather branch length most likely reflects the number of infection opportunities per year [28] or more rarely a mutator phenotype [54]. This number is expected to increase when *B. anthracis* encounters a favorable

ecological niche. Each split in the tree reflects the colonization of a new ecotype allowing the fixation of an additional, independent lineage. Usually such a split will result from geographic spreading, and the new ecotype may be the one with the longest, faster evolving branch, at least when a significant difference is observed, reflecting the arrival in a new, naive environment. In this context, polytomies constitute exceptional opportunities to try and propose dating points. Polytomies result from the sudden colonization of multiple new ecotypes which may reflect exceptional environmental changes. Such changes may have an anthropic origin and it may be easier to associate a polytomy with major historical events. Two large polytomies have been described so far in *B. anthracis* the A.Br.008/011 [32] and the derived A.Br.011/009 polytomies [37]. They constitute the "TransEurAsia" (TEA) subclade [32].

The branching of fast-evolving lineages from the same position within a unique sublineage of the A.Br.011/009 polytomy towards both West Africa and North America indicated that the contamination was exported from a geographically limited region having exchanges at the same time with both continents (Canada, Senegal-Gambia). France, seventeenth century was proposed as the most likely spatiotemporal candidate [38]. Based on the proposed dating, a tentative dating of the most recent ancestor of the A.Br.011/009 polytomy to the One Hundred Years War between France and England, AD 1350-1450 was further hypothesized [38].

However, this proposition does not fit well with the present report showing a remarkable intermingling of Italian and French strains early during the emergence of the A.Br.011/009 polytomy. Rather this intermingling would indicate that the A.Br.011/009 emerged in a time of conflicts between Italy and France, rather than France and England, before the seventeenth century. Years 1250-1300 and years 1450-1550 are two candidate periods. Battles at that time involved thousands of horses. We speculate that such events provided major opportunities for contaminations of both cavalries with the same population of strains. These large "flocks" will then spread the contamination on their way back.

In contrast to the A.Br.011/009 polytomy, the A.Br.008/011 polytomy is characterized by a remarkable geographic spreading. Rare, deep branching lineages are observed in Europe (Italy, Bulgaria, France) as well as Turkey, China and Yakutia (the permafrost strains, this report). There is one major event in human history prior to the sixteenth century which could explain such a distribution, the Mongol conquests during the thirteenth century [55, 56]. Chinghis Khan assembled an empire extending from Northern China to the East side of the Caspian sea. The death of Chingis-khan in 1227 triggered the gathering of the Chingizids Armies for the election of Ögedei as new khan, and the death of Ögedei in 1241 eventually triggered a new gathering in 1246. The conquests, powered by horses, involved long-distance displacement of tens of thousands of horses. The immediate successors of Chingis-khan invaded Europe as far as Hungaria in 1237-1242 and Anatolia (Turkey) in 1241-1243.

The Mongol Empire began to disintegrate in the second half of the 13th century. But despite this, the territory from China to Eastern Europe remained under the rule of Chingizids, and Eastern Europe was constantly subjected to raids from the Golden Horde for the purpose of plundering, or simply the presence of military contingents participating in wars. Thus, a common political space was established, ensuring relatively large movements of people and animals between Asia and Europe, which could further contribute to the relatively rapid and unhindered transfer of pathogens of infectious diseases between these regions. In addition to military operations, the Mongols organized a Yam - chain of relay stations at certain distances to each other, allowing to replace horses for messengers or messengers themselves, and making possible to deliver cargo and documents within weeks over long distances. This postal system also could promote rapid spread of infections, but to a much lesser extent than the massive movements of armies driven by horses.

The presence of Chingizids in the European region ceased when Russian Tsar Ivan the fourth (Ivan the Terrible) conquered Western states, which were formed after the split of the Golden Horde - the Kazan Khanate in 1552, the Astrakhan Khanate in 1556 and the Siberian Khanate in 1582-1598. After that, the only post-Mongolian state remained the Crimean Khanate,

554 which regularly carried out raids on Russia and Poland (the territories of modern Russia, Ukraine,
555 Belarus, Lithuania, Latvia, Estonia and Moldova) until it was conquered by Catherine the Great in
556 1783. Despite the active raid policy of the Crimean Khanate and the constant use of the Tatar
557 contingents by the Moscow tsars in the course of constant wars in eastern and northern Europe,
558 the continuity of migration routes of nomads from Asia to Europe was broken.

559 Consequently we propose here that the root of the A.Br.008/011 polytomy corresponds to
560 a *B. anthracis* ecotype present in the Mongolian armies between the first half of the 13th century
561 and the middle of the 16th century. We particularly favor the first half of the 13th century as the
562 period when *B. anthracis* could have been transported in a short time-frame by the animals
563 associated with the Mongolian armies, particularly war and led horses in all geographic areas
564 covered by the deep branching lineages of the A.Br.008/011 polytomy, i.e. from China to
565 Hungaria. After that time, we speculate that the split of the Mongolian empire would have
566 significantly hindered the spread of the infection.

567 In contrast with this relatively precise dating hypothesis, we see at present no clue
568 regarding the geographic origin of the ancestor of the A.Br.008/011 polytomy. The contamination
569 of the Mongolian army might have occurred in many locations, including Eastern Europe.

570 Apparent contradiction between the dating of the emergence of
571 the A.Br.011/008 polytomy and of the permafrost samples

572 Under the proposed hypothesis, the A.Br.011/008 polytomy can be conservatively dated
573 from the early 13th century to the middle of the 16th century. The deposition of the A.Br.011/008
574 strains including the one represented by 4Ya recovered from permafrost at a depth of two and three
575 meters would be posterior to this period. The 3Ya sublineage is even more recent, the relative
576 length of the branch is less than half the total length of lineage L6_Tsiantowskii to which it
577 belongs.

578 The Yakoutia permafrost soil samples were taken from alluvial (river) sediments - there is
579 a wide flat valley with bayou lakes. Probably this corresponds to holocene sediments (age about

10,000 years), frozen as they accumulated. A layer of permafrost formed simultaneously in Yakutia and Yamal 20-40 thousand years ago. Radiocarbon analysis of other sediments sampled in Uyandina riverside and other places of Abyisk district at the same depth, indicated that they were 3-10 thousand years old [57]. On average, in Yakutia, the depth of seasonal thawing does not exceed 2 - 2.5 meters. Therefore we expected that strains of *B. anthracis*, isolated from the permafrost, would be significantly older. However, the accidental nature of the exceptional paleontological finding in Yakutia, and the extraction of paleontological material and soil samples by prospectors rather than professional paleontologists or geologists, may be responsible for a weak geological information about the soil layers from which the strains were extracted. The finding of *B. anthracis* was quite unexpected and was triggered by the cave lions investigation, and some time passed between the excavation of the cave lions and the soil sampling.

We also carefully examined the extent to which the present findings could be the result, or be affected, by contaminations of different kinds. However, the possibility that the strains were introduced into the soil samples under study as a result of contamination during work is most unlikely. Drift of spores from the surface is unlikely, because of the absence of reported cases at the sampling site recently and in view of the absence of *B. anthracis* spores in the upper (minus 1 meter) sample. Most importantly the initial identification of *B. anthracis* was made in a laboratory which does not work with pathogenic microorganisms and does not maintain such strain collections. The contamination in SRCAMB is also very unlikely as it would require a simultaneous contamination with three different strains in only three samples. In addition and most convincingly, whole genome sequence analysis of strains from the SRCAMB collection showing a similar MLVA genotype demonstrated that these are definitely different strains.

A number of conclusions can thus be robustly drawn. We looked for the presence of *B. anthracis* spores over a depth of six meters, going from one meter below surface down to the cave lion kittens. The state of preservation of the kittens indicated that the bodies, and the corresponding permafrost layer, remained frozen for thousands of years. Under these conditions, the most

parsimonious explanation for the finding of *B. anthracis* strains only in the upper layers is that *B. anthracis* was not present in the region at the time of the death of the kittens, 5000 to 10,000 thousand years ago. Rather *B. anthracis* arrived relatively recently in three occasions represented by three distinct lineages. This diversity cannot be the result of pathogen evolution in the soil. The three genotypes are clearly positioned in different places of the *B. anthracis* evolutionary tree. Thus, all three strains most likely hit the ground and were conserved at different times, during various epizootics. The lack of a clear stratigraphy, that is, the isolation of isolates that represent a single genotype from soil samples from different depths, and more importantly the finding of spores lower than expected from the proposed dating would imply that microorganisms are able to migrate in permafrost. This would be compatible with current knowledge on permafrost [58, 59].

Explaining the close relationship between the Yamal and Yakoutia *B. anthracis* strains

One surprising observation in our study is the fairly close genetic relationship of the Yamal isolates with the Yakut strain 5Ya. Between the regions where these strains were isolated lies a distance of about 2 thousand kilometers. However, despite their remoteness from each other, they are located at similar longitude, and the ecosystems located in them are almost identical. In this connection it can be assumed that these strains are representatives of a certain "tundra" population of *B. anthracis*, spread on a significant territory of Northern Eurasia, at least in the past. In this case, the circulation of strains could be ensured by migration of ungulate populations, primarily reindeer.

The territory of Yakutia was inhabited by modern humans at least from the Mesolithic, but before the beginning of the 2nd millennium AD it was inhabited exclusively by tribes of hunters and fisher-men. The first population of herders, the ancestors of modern Yakuts, migrated here only at the beginning of the second millennium from the Baikal region. They were livestock

631 breeders, bred cows and horses (in the Baikal region they also bred sheep and camels, but it is
632 impossible to breed them in the territory of Yakutia due to the severe climate).

633 The conquest by the Mongols of southern Siberia, slightly increased the migration
634 movement from Baikal region. Tribes with domestic reindeer that came from the south of Siberia
635 (Baikal region), by the middle of the second millennium AD reached the territory of western
636 Siberia, which includes the Yamal Peninsula, and the north of Eastern Siberia - Yakutia. In western
637 Siberia, these were the ancestors of the Nenets, in the eastern of the Evenks. Theoretically, it can
638 be assumed that in the middle of the 2nd millennium AD those and others could contact somewhere
639 in the area of the Yenisei basin, south of Taimyr, for example, in the Turukhansk district, from
640 where a route to the Yamal is possible. At that time reindeers were mainly used for transportation
641 and the number of domestic deers was low. Until the 17th century there were no large herds, the
642 maximum livestock in one farm could reach one hundred heads of deer. Large-herd reindeer
643 herding developed only after the colonization of Siberia by Russians, beginning from the 17th and
644 18th centuries. Currently, the number of herds in one farm reaches several thousand heads. More
645 knowledge on *B. anthracis* strains present in Northern Europe and Siberia will help better
646 understand the history of *B. anthracis* spreading among reindeers.

647 Only in the middle of the 20th century, in connection with the beginning of mass
648 vaccination of reindeer and the introduction of veterinary and sanitary control, obstacles arose for
649 the free distribution of *B. anthracis* in the tundra zone. However, considering the ability of *B.*
650 *anthracis* to form endospores and the presence of permafrost, capable of preserving these spores,
651 further increasing the period of their viability, multiple soil foci could have formed by this time
652 scattered over a vast territory. This territory is very little involved in economic activity and is
653 extremely poorly populated (for example, the average population density in Yakutia is 0.3 people
654 per square kilometer; since two thirds of the population lives in cities, the density in rural areas is
655 only about 0.1 person per square kilometer). This makes not only sanitation, but even detection
656 and recording of such foci an impossible task. The events of the summer of 2016 in Yamal have

shown that such foci retain their epidemiological potential for a long time, and under favorable conditions, primarily in the thawing of permafrost due to local or global warming, they can become a source of infection, causing large-scale epidemics, resulting in casualties and significant economic costs. A reactivated focus may remain active for some years. In 2017 32 samples from the Yamal 2016 epidemic area were tested. Two samples - ash from the place of a dead deer burning (lat. N68.24010, lon. E71.01.200) and the biological material from not completely burnt deer (lat. N68.24989, lon. E70.44435) – containing PCR detectable *B. anthracis* DNA detected in PCR. Virulent bacteria could be cultivated from both samples. The MLVA genotype was identical to the strains previously isolated in the epidemic area in 2016 [60].

If spores were able to keep viability during a year on the soil surface, then there is little doubt that during seasonal snow melting they could be spread over a large area and penetrate into the depths of the soil. Thus, in the tundra areas, where at least once an anthrax outbreak was recorded, it is necessary to conduct continuously anti-epidemic measures, such as vaccination of livestock and herdsmen, as well as maintain the readiness of medical and veterinary institutions for the diagnosis of anthrax and emergency measures for detecting the disease.

Conclusion

In summary, we have detected three independent events of *B. anthracis* introduction in Northern Yakutia, stored at different depth in the permafrost. In the proposed model, the third and most recent introduction, detected at minus 2 meters, occurred as a side effect of Russian conquests and development of agriculture in the 17th-18th century. The second introduction detected at minus 2 and minus 3 meters, would be the byproduct of Yakut's population migration from Lake Baikal area in the 14th-15th century. The first introduction detected at minus 4 meters, cannot be dated precisely but the location in the permafrost may indicate that it is not more than a few centuries older than the second introduction. We propose to date the emergence of the A.Br.008/011 polytomy to the first half of the thirteenth century, in relation with the Mongolian conquests.

Acknowledgements

We wish to express our gratitude to Vitalii Kharkov, the Chief doctor of Center for Hygiene and Epidemiology in the Yamalo-Nenets Autonomous District for his help in organizing of sampling, Dr. Alla Ryazanova from Stavropol Anti-Plague Institute for her role in the clinical and laboratory diagnosis of anthrax during the Yamal anthrax epidemic and Dr. Oktyabrina Sofronova, the head of the Laboratory of especially dangerous infections of Center for Hygiene and Epidemiology in the Republic of Sakha (Yakutia) for samples from paleontological discovery point sent to SRCAMB. Also we show our appreciation to Dr. Victor Karlov from Department of Ethnology, Faculty of History, Moscow State University for kindly provided historical advice and to Dr. Anatoli Brouchkov, the Head of Geocryology Department, Faculty of Geology, Lomonosov Moscow State University, for his geology advice.

References

1. Saile E, Koehler TM. *Bacillus anthracis* multiplication, persistence, and genetic exchange in the rhizosphere of grass plants. Appl Environ Microbiol. 2006;72(5):3168-74. Epub 2006/05/05. doi: 10.1128/AEM.72.5.3168-3174.2006. PubMed PMID: 16672454; PubMed Central PMCID: PMCPMC1472387.

2. Lembo T, Hampson K, Auty H, Beesley CA, Bessell P, Packer C, et al. Serologic surveillance of anthrax in the Serengeti ecosystem, Tanzania, 1996-2009. Emerg Infect Dis. 2011;17(3):387-94. Epub 2011/03/12. doi: 10.3201/eid1703.101290. PubMed PMID: 21392428; PubMed Central PMCID: PMCPMC3166018.

3. Turner WC, Kausrud KL, Beyer W, Easterday WR, Barandongo ZR, Blaschke E, et al. Lethal exposure: An integrated approach to pathogen transmission via environmental reservoirs. Sci Rep. 2016;6:27311. Epub 2016/06/07. doi: 10.1038/srep27311. PubMed PMID: 27265371; PubMed Central PMCID: PMCPMC4893621.

706 4. Ahsan MM, Khan MFR, Rahman MB, Hassan J, Chowdhury SMZH, Parvej MS,
707 et al. Investigation into *Bacillus anthracis* Spore in Soil and Analysis of Environmental Parameters
708 Related to Repeated Anthrax Outbreak in Sirajganj, Bangladesh. Thai J Vet Med. 2013;43(3):449-
709 54. PubMed PMID: WOS:000339784200018.
710 5. Van Ness GB. Ecology of anthrax. Science. 1971;172:1303-7. doi: 10.1126 /
711 science.172.3990.1303
712 6. Dragon DC, Rennie RP. The ecology of anthrax spores: tough but not invincible.
713 Can Vet J. 1995;36(5):295-301. Epub 1995/05/01. PubMed PMID: 7773917; PubMed Central
714 PMCID: PMCPMC1686874.
715 7. Gates CC, Elkin BT, Dragon DC. Investigation, control and epizootiology of
716 anthrax in a geographically isolated, free-roaming bison population in northern Canada. Can J Vet
717 Res. 1995;59(4):256-64. Epub 1995/10/01. PubMed PMID: 8548686; PubMed Central PMCID:
718 PMCPMC1263779.
719 8. Elvander M, Persson B, Sternberg Lewerin S. Historical cases of anthrax in Sweden
720 1916-1961. Transbound Emerg Dis. 2017;64(3):892-8. Epub 2015/12/17. doi:
721 10.1111/tbed.12456. PubMed PMID: 26671241.
722 9. Maksimović Z, Cornwell MS, Semren O, Rifatbegović M. The apparent role of
723 climate change in a recent anthrax outbreak in cattle. Rev Sci Tech Off Int Epiz. 2017;36(3):959-
724 63. doi: doi/org/10.20506/rst.36.3.2727.
725 10. Griffin DW, Petrosky T, Morman SA, Luna VA. A survey of the occurrence of
726 *Bacillus anthracis* in North American soils over two long-range transects and within post-Katrina
727 New Orleans. Applied Geochemistry. 2009;24:1464–71. doi: 10.1016/j.apgeochem.2009.04.016.
728 11. Hugh-Jones ME, de Vos V. Anthrax and wildlife. Rev Sci Tech. 2002;21(2):359-
729 83. Epub 2002/04/27. PubMed PMID: 11974621.
730 12. Blackburn JK, McNyset KM, Curtis A, Hugh-Jones ME. Modeling the geographic
731 distribution of *Bacillus anthracis*, the causative agent of anthrax disease, for the contiguous United

732 States using predictive ecological [corrected] niche modeling. Am J Trop Med Hyg.
733 2007;77(6):1103-10. Epub 2008/01/01. PubMed PMID: 18165531.

734 13. Inglesby TV, O'Toole T, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, et al.
735 Anthrax as a biological weapon, 2002: updated recommendations for management. JAMA.
736 2002;287(17):2236-52. Epub 2002/05/01. PubMed PMID: 11980524.

737 14. Fowler RA, Shafazand S. Fowler RA, Shafazand S (2011) Anthrax Bioterrorism:
738 Prevention, Diagnosis and Management Strategies. 2:107. doi: 10.4172/2157-2526.1000107. J
739 Bioterr Biodef. 2011;2(2). doi: 10.4172/2157-2526.1000107.

740 15. Spencer RC. *Bacillus anthracis*. J Clin Pathol. 2003;56(3):182-7. Epub 2003/03/01.
741 PubMed PMID: 12610093; PubMed Central PMCID: PMCPMC1769905.

742 16. WHO. Anthrax in humans and animals. 4 ed. Geneva (CH): World Health
743 Organization Press; 2008.

744 17. Peculi A, Campese E, Serrecchia L, Marino L, Boci J, Bijo B, et al. Genotyping of
745 *Bacillus anthracis* Strains Circulating in Albania. J Bioterror Biodef. 2015;6(1). doi:
746 10.4172/2157-2526.1000131.

747 18. Okello A, Welburn S, Smith J. Crossing institutional boundaries: mapping the
748 policy process for improved control of endemic and neglected zoonoses in sub-Saharan Africa.
749 Health Policy Plan. 2015;30(6):804-12. Epub 2014/07/09. doi: 10.1093/heapol/czu059. PubMed
750 PMID: 25000963.

751 19. Chakraborty A, Khan SU, Hasnat MA, Parveen S, Islam MS, Mikolon A, et al.
752 Anthrax Outbreaks in Bangladesh, 2009-2010. Am J Trop Med Hyg. 2012;86(4):703-10. doi:
753 10.4269/ajtmh.2012.11-0234. PubMed PMID: WOS:000302519700024.

754 20. Okutani A, Tungalag H, Boldbaatar B, Yamada A, Tserennorov D, Otgonchimeg I,
755 et al. Molecular epidemiological study of *Bacillus anthracis* isolated in Mongolia by multiple-
756 locus variable-number tandem-repeat analysis for 8 loci (MLVA-8). Jpn J Infect Dis.
757 2011;64(4):345-8. Epub 2011/07/27. PubMed PMID: 21788715.

21. Cherkasskiy BL. A national register of historic and contemporary anthrax foci. J Appl Microbiol. 1999;87(2):192-5. Epub 1999/09/04. PubMed PMID: 10475946.

22. Revich BA, Podolnaya MA. Thawing of permafrost may disturb historic cattle burial grounds in East Siberia. Glob Health Action. 2011;4. Epub 2011/11/25. doi: 10.3402/gha.v4i0.8482. PubMed PMID: 22114567; PubMed Central PMCID: PMC3222928.

23. Williams PJ, Smith MW. The frozen earth: fundamentals of geocryology: Cambridge University Press; 1991 1991. 306 p. : ill. ; 25 cm p.

24. Steven B, Leveille R, Pollard WH, Whyte LG. Microbial ecology and biodiversity in permafrost. Extremophiles. 2006;10(4):259-67. Epub 2006/03/22. doi: 10.1007/s00792-006-0506-3. PubMed PMID: 16550305.

25. Jansson JK, Tas N. The microbial ecology of permafrost. Nat Rev Microbiol. 2014;12(6):414-25. Epub 2014/05/13. doi: 10.1038/nrmicro3262. PubMed PMID: 24814065.

26. Legendre M, Bartoli J, Shmakova L, Jeudy S, Labadie K, Adrait A, et al. Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology. Proc Natl Acad Sci U S A. 2014;111(11):4274-9. Epub 2014/03/05. doi: 10.1073/pnas.1320670111. PubMed PMID: 24591590; PubMed Central PMCID: PMCPMC3964051.

27. Thierry S, Tourterel C, Le Flèche P, Derzelle S, Dekhil N, Mendy C, et al. Genotyping of French *Bacillus anthracis* Strains Based on 31-Loci Multi Locus VNTR Analysis: Epidemiology, Marker Evaluation, and Update of the Internet Genotype Database. PLoS One. 2014;9(6):e95131. Epub 2014/06/06. doi: 10.1371/journal.pone.0095131

PONE-D-13-27818 [pii]. PubMed PMID: 24901417; PubMed Central PMCID: PMC4046976.

28. 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

783 SNP genotyping assay in bacterial models. PLoS One. 2012;7(3):e32866. Epub 2012/03/23. doi:
784 10.1371/journal.pone.0032866
785 PONE-D-11-21413 [pii]. PubMed PMID: 22438886; PubMed Central PMCID:
786 PMC3306377.
787 29. Van Ert MN, Easterday WR, Huynh LY, Okinaka RT, Hugh-Jones ME, Ravel J, et al.
788 Global genetic population structure of *Bacillus anthracis*. PLoS One. 2007;2(5):e461. Epub
789 2007/05/24. doi: 10.1371/journal.pone.0000461. PubMed PMID: 17520020; PubMed Central
790 PMCID: PMC1866244.
791 30. Antwerpen MH, Sahl JW, Birdsell D, Pearson T, Pearce MJ, Redmond C, et al.
792 Unexpected Relations of Historical Anthrax Strain. MBio. 2017;8(2). Epub 2017/04/27. doi:
793 10.1128/mBio.00440-17. PubMed PMID: 28442608; PubMed Central PMCID:
794 PMCPMC5405232.
795 31. Venkateswaran K, Singh NK, Checinska Sielaff A, Pope RK, Bergman NH, van
796 Tongeren SP, et al. Non-Toxin-Producing *Bacillus cereus* Strains Belonging to the *B. anthracis*
797 Clade Isolated from the International Space Station. mSystems. 2017;2(3). Epub 2017/07/07. doi:
798 10.1128/mSystems.00021-17. PubMed PMID: 28680972; PubMed Central PMCID:
799 PMCPMC5487513.
800 32. Sahl JW, Pearson T, Okinaka R, Schupp JM, Gillece JD, Heaton H, et al. A *Bacillus*
801 *anthracis* Genome Sequence from the Sverdlovsk 1979 Autopsy Specimens. MBio. 2016;7(5).
802 Epub 2016/09/30. doi: 10.1128/mBio.01501-16. PubMed PMID: 27677796; PubMed Central
803 PMCID: PMCPMC5050339.
804 33. Price EP, Seymour ML, Sarovich DS, Latham J, Wolken SR, Mason J, et al.
805 Molecular epidemiologic investigation of an anthrax outbreak among heroin users, Europe. Emerg
806 Infect Dis. 2012;18(8):1307-13. Epub 2012/07/31. doi: 10.3201/eid1808.111343. PubMed PMID:
807 22840345; PubMed Central PMCID: PMC3414016.

808 34. Keim P, Grunow R, Vipond R, Grass G, Hoffmaster A, Birdsell DN, et al. Whole
809 Genome Analysis of Injectional Anthrax Identifies Two Disease Clusters Spanning More Than 13
810 Years. EBioMedicine. 2015;2(11):1613-8. Epub 2016/02/13. doi: 10.1016/j.ebiom.2015.10.004.
811 PubMed PMID: 26870786; PubMed Central PMCID: PMCPMC4740342.
812 35. Hanczaruk M, Reischl U, Holzmann T, Frangoulidis D, Wagner DM, Keim PS, et
813 al. Injectional anthrax in heroin users, Europe, 2000-2012. Emerg Infect Dis. 2014;20(2):322-3.
814 Epub 2014/01/23. doi: 10.3201/eid2002.120921. PubMed PMID: 24447525; PubMed Central
815 PMCID: PMCPMC3901468.
816 36. Blouin Y, Cazajous G, Dehan C, Soler C, Vong R, Hassan MO, et al. Progenitor
817 "*Mycobacterium canettii*" clone responsible for lymph node tuberculosis epidemic, Djibouti.
818 Emerg Infect Dis. 2014;20(1):21-8. Epub 2014/02/13. doi: 10.3201/eid2001.130652. PubMed
819 PMID: 24520560; PubMed Central PMCID: PMC3884719.
820 37. Girault G, Blouin Y, Vergnaud G, Derzelle S. High-throughput sequencing of
821 *Bacillus anthracis* in France: investigating genome diversity and population structure using whole-
822 genome SNP discovery. BMC Genomics. 2014;15(1):288. Epub 2014/04/17. doi: 10.1186/1471-
823 2164-15-288
824 1471-2164-15-288 [pii]. PubMed PMID: 24734872; PubMed Central PMCID:
825 PMC4023602.
826 38. Vergnaud G, Girault G, Thierry S, Pourcel C, Madani N, Blouin Y. Comparison of
827 French and Worldwide *Bacillus anthracis* Strains Favors a Recent, Post-Columbian Origin of the
828 Predominant North-American Clade. PLoS One. 2016;11(2):e0146216. Epub 2016/02/24. doi:
829 10.1371/journal.pone.0146216. PubMed PMID: 26901621; PubMed Central PMCID:
830 PMCPMC4763433.
831 39. Kenefic LJ, Pearson T, Okinaka RT, Schupp JM, Wagner DM, Hoffmaster AR, et
832 al. Pre-Columbian origins for North American anthrax. PLoS One. 2009;4(3):e4813. Epub

2009/03/14. doi: 10.1371/journal.pone.0004813. PubMed PMID: 19283072; PubMed Central PMCID: PMC2653229.

40. Kenefic LJ, Okinaka RT, Keim P. Population genetics of *Bacillus*: phylogeography of anthrax in North America. In: Robinson DA, Falush D, Feil EJ, editors. Bacterial population genetics in infectious disease. Hoboken, New Jersey: Wiley-Blackwell; 2010. p. 169-79.

41. Pullan ST, Pearson TR, Latham J, Mason J, Atkinson B, Silman NJ, et al. Whole-genome sequencing investigation of animal-skin-drum-associated UK anthrax cases reveals evidence of mixed populations and relatedness to a US case. Microb Genom. 2015;1(5):e000039. Epub 2015/11/07. doi: 10.1099/mgen.0.000039. PubMed PMID: 28348823; PubMed Central PMCID: PMCPMC5320680.

42. Rouli L, MBengue M, Robert C, Ndiaye M, La Scola B, Raoult D. Genomic analysis of three African strains of *Bacillus anthracis* demonstrates that they are part of the clonal expansion of an exclusively pathogenic bacterium. New Microbe and New Infect. 2014;2:161-9.

43. Popova A. Opyt likvidatsii vspyshki sibirskoy yazvy na Yamale v 2016 godu (The experience of eliminating the outbreak of anthrax on Yamal in 2016) [Russian], Izhevsk, "Print-2", 2017

44. Cherkasskiy BL. Puteshestvie epidemiologa vo vremeni i prostranstve (Traveling of an epidemiologist in time and space). Voronezh2003. 638 p.

45. [17/10/2018]. Available from: http://climatechange.igce.ru/index.php?option=com_docman&task=cat_view&gid=32&Itemid=75&lang=en.

46. Braun P, Grass G, Aceti A, Serrecchia L, Affuso A, Marino L, et al. Microevolution of Anthrax from a Young Ancestor (M.A.Y.A.) Suggests a Soil-Borne Life Cycle of Bacillus anthracis. PLoS One. 2015;10(8):e0135346. Epub 2015/08/13. doi: 10.1371/journal.pone.0135346. PubMed PMID: 26266934; PubMed Central PMCID: PMCPMC4534099.

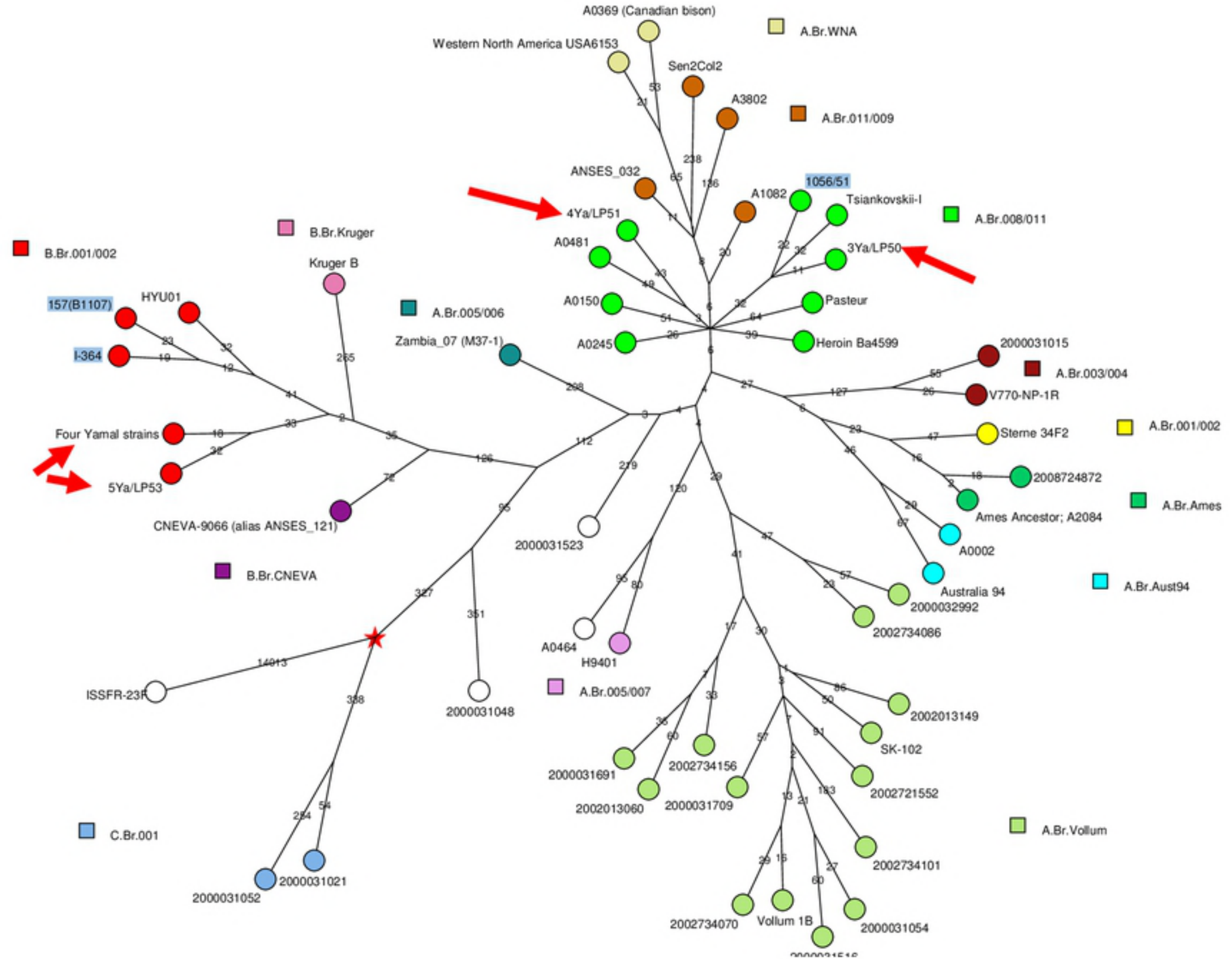
859 47. Kracalik I, Malania L, Broladze M, Navdarashvili A, Imnadze P, Ryan SJ, et al.
860 Changing livestock vaccination policy alters the epidemiology of human anthrax, Georgia, 2000-
861 2013. Vaccine. 2017;35(46):6283-9. Epub 2017/10/11. doi: 10.1016/j.vaccine.2017.09.081.
862 PubMed PMID: 28988866.
863 48. Okinaka R, Pearson T, Keim P. Anthrax, but not *Bacillus anthracis*? PLoS Pathog.
864 2006;2(11):e122. Epub 2006/11/24. doi: 10.1371/journal.ppat.0020122. PubMed PMID:
865 17121463; PubMed Central PMCID: PMCPMC1657067.
866 49. Smith KL, De Vos V, Bryden HB, Hugh-Jones ME, Klevytska A, Price LB, et al.
867 Meso-scale ecology of anthrax in southern Africa: a pilot study of diversity and clustering. J Appl
868 Microbiol. 1999;87(2):204-7. Epub 1999/09/04. PubMed PMID: 10475949.
869 50. Leendertz FH, Lankester F, Guislain P, Neel C, Drori O, Dupain J, et al. Anthrax
870 in Western and Central African great apes. Am J Primatol. 2006;68(9):928-33. Epub 2006/08/11.
871 doi: 10.1002/ajp.20298. PubMed PMID: 16900500.
872 51. Hoffmaster AR, Hill KK, Gee JE, Marston CK, De BK, Popovic T, et al.
873 Characterization of *Bacillus cereus* isolates associated with fatal pneumonias: strains are closely
874 related to *Bacillus anthracis* and harbor *B. anthracis* virulence genes. J Clin Microbiol.
875 2006;44(9):3352-60. Epub 2006/09/07. doi: 10.1128/JCM.00561-06. PubMed PMID: 16954272;
876 PubMed Central PMCID: PMCPMC1594744.
877 52. Blouin Y, Hauck Y, Soler C, Fabre M, Vong R, Dehan C, et al. Significance of the
878 identification in the Horn of Africa of an exceptionally deep branching *Mycobacterium*
879 *tuberculosis* clade. PLoS One. 2012;7(12):e52841. Epub 2013/01/10. doi:
880 10.1371/journal.pone.0052841
881 PONE-D-12-22563 [pii]. PubMed PMID: 23300794; PubMed Central PMCID:
882 PMC3531362.
883 53. Blouin Y. A new scenario for the early evolution of *Mycobacterium tuberculosis*:
884 University Paris-Saclay; 2014.

885 54. Ågren J, Finn M, Bengtsson B, Segerman B. Microevolution during an Anthrax
886 outbreak leading to clonal heterogeneity and penicillin resistance. PLoS One. 2014;9(2):e89112.
887 Epub 2014/02/20. doi: 10.1371/journal.pone.0089112. PubMed PMID: 24551231; PubMed
888 Central PMCID: PMCPMC3923885.
889 55. Saunders JJ. The History of the Mongol Conquests. Philadelphia, Pennsylvania
890 19104 4011: University of Pennsylvania Press; 2001. XIX-275 p.
891 56. Büntgen U, Di Cosmo N. Climatic and environmental aspects of the Mongol
892 withdrawal from Hungary in 1242 CE. Sci Rep. 2016;6:25606. Epub 2016/05/27. doi:
893 10.1038/srep25606. PubMed PMID: 27228400; PubMed Central PMCID: PMCPMC4881396.
894 57. Kospokevich V, Dneprovskaya O. Radio carbon data of the laboratory of
895 geochemistry of the institute of permafrost of the USSR Academy of Sciences, Message XI
896 [Russian]. Bulletin of the commission on studying of the Quarterny period. 1984;53:139-42.
897 58. Dash JG, Fu H, Wettlaufer JS. The premelting of ice and its environmental
898 consequences. Rep Prog Phys. 1995;58:115-67. doi: 10.1088/0034-4885/58/1/003.
899 59. Stähli M, Jansson P-E, Lundin L-C. Soil moisture redistribution and infiltration in
900 frozen sandy soils. Water Resour Res. 1999;35:95–103.
901 60. Savinych AV, Pavlov DL, Kuznetsovskiy AV, Fomenkov OO, Tumanov AS,
902 Susloparov AA, et al. The results of the monitoring of anthrax death fields of the 2016 epizootic
903 in the Yamal district of the Yamalo-Nenets Autonomous District [Russian]. Vestnik voysk RKhB
904 zachity (Bulletin of RCB protection forces). 2017;1(3):18-27.

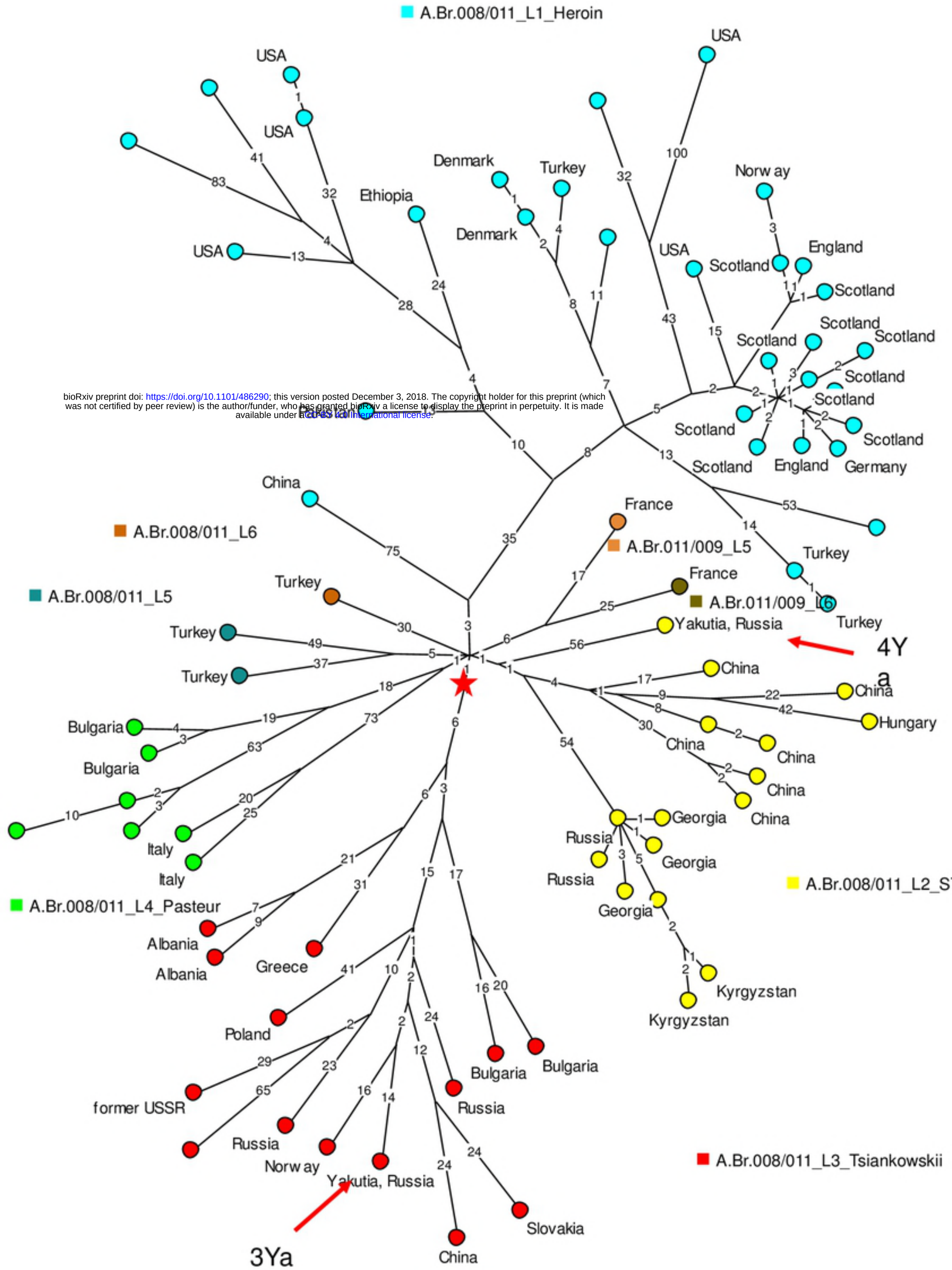
905 # Supporting information caption

906 **S1 Table. The MLVA profiles of described strains**

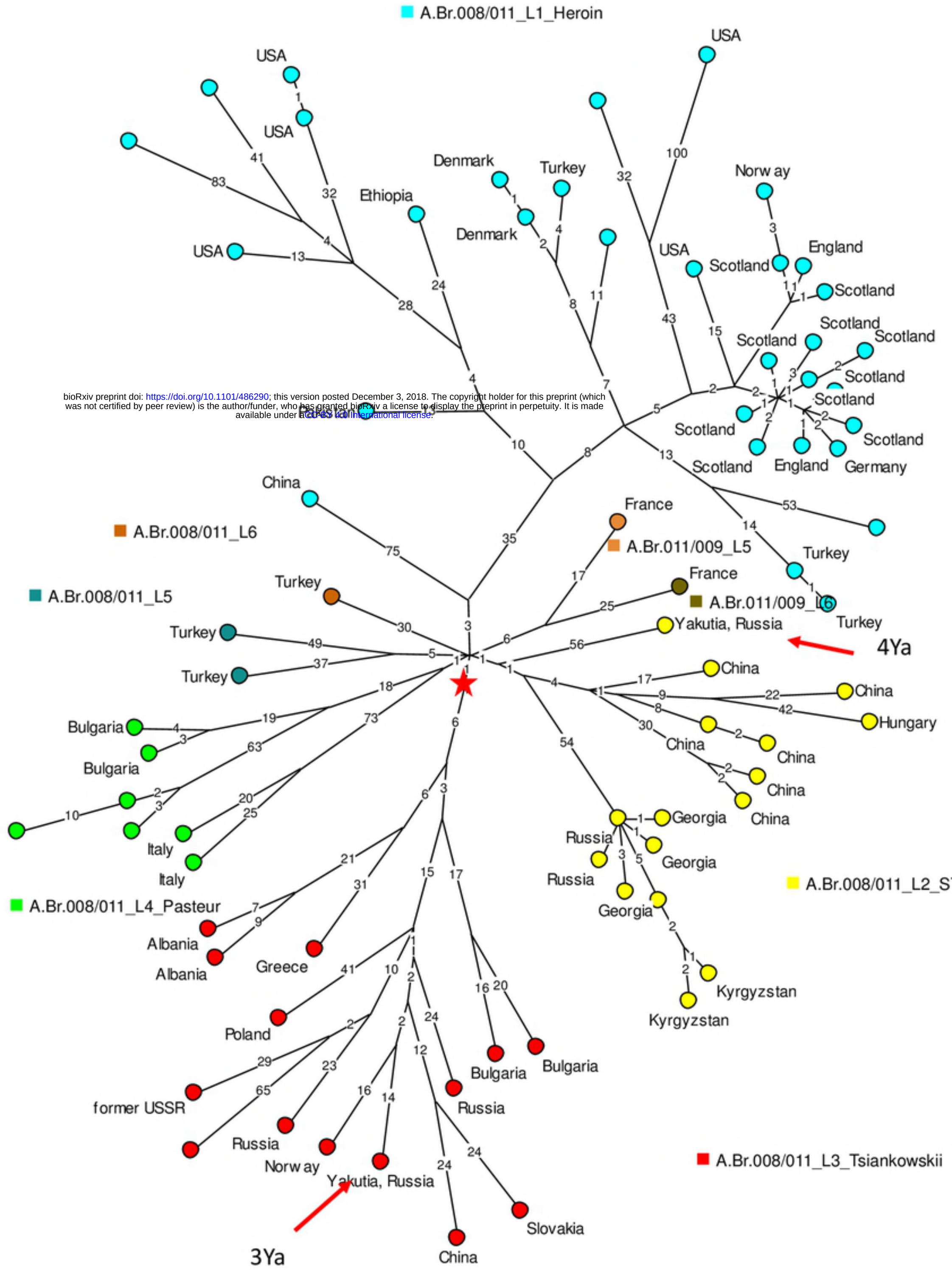
907 **S2 Table LD50 values of studied strains for mice and guinea pigs**


A0369 (Canadian bison)
A.Br.WNA
Western North America USA6153
Sen2Col2
A3802
A.Br.011/009
ANSES_032
A1082
1056/51
Tsiankovskii-I
4Ya/LP51
A.Br.008/011
3Ya/LP50
B.Br.Kruger
B.Br.001/002
A0481
Kruger B
157(B1107)
HYU01
A.Br.005/006
A0150
Pasteur
I-364
Zambia_07 (M37-1)
A0245
Heroin Ba4599
2000031015
A.Br.003/004
V770-NP-1R
A.Br.001/002
Four Yamal strains
Sterne 34F2
2008724872
A.Br.Ames
5Ya/LP53
Ames Ancestor; A2084
CNEVA-9066 (alias ANSES_121)
A0002
2000031523
B.Br.CNEVA
Australia 94
A.Br.Aust94
2000032992
2002734086
A0464
H9401
A.Br.005/007
ISSFR-23F
2000031048
2002013149
SK-102
2000031691
2002734156
2002721552
2002013060
2000031709
C.Br.001
A.Br.Vollum
2000031021
2002734101
2000031052
2002734070
Vollum 1B
2000031054

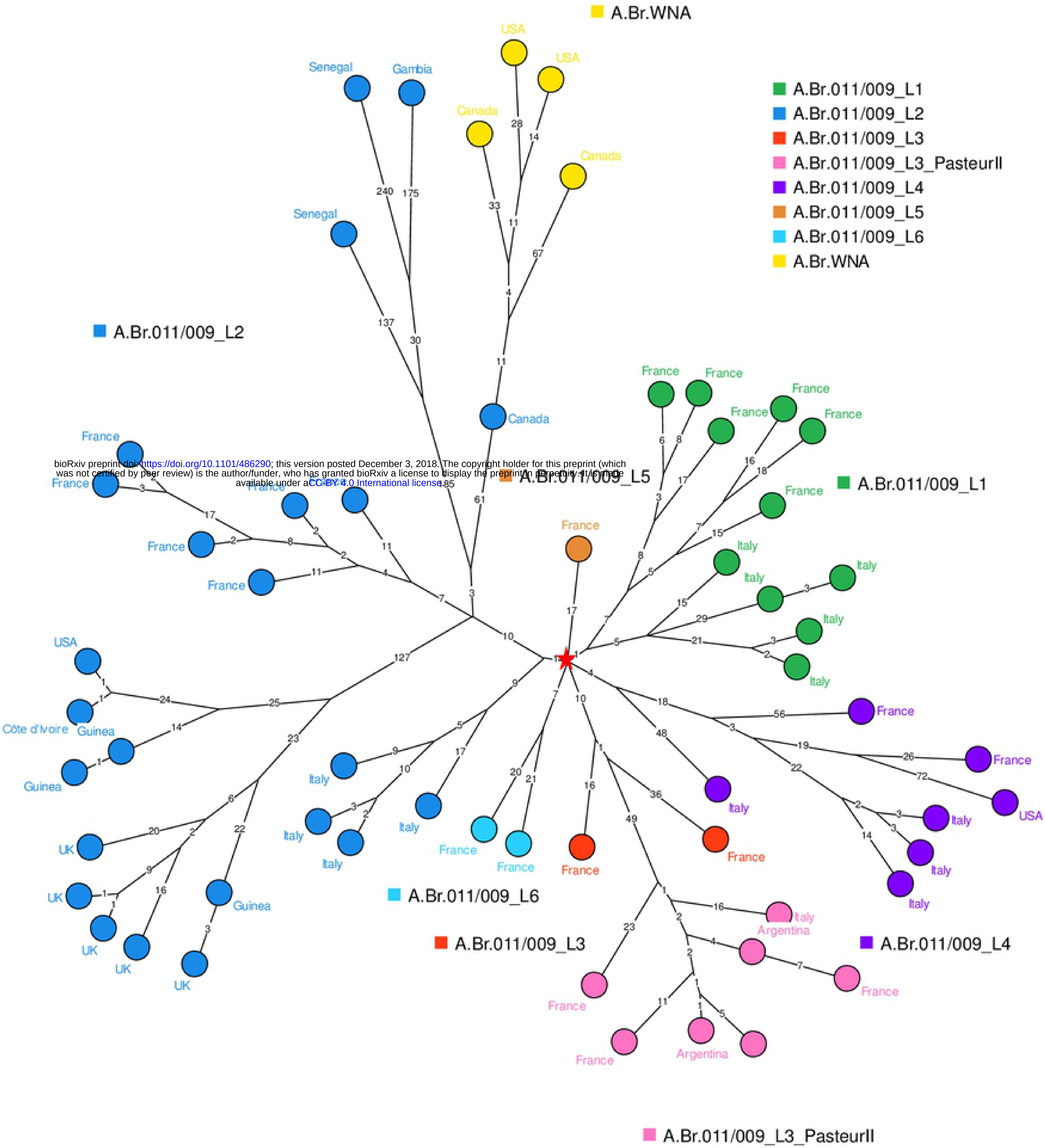
Figure



Figure



Figure



Figure