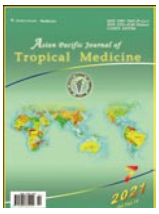

Original Article

Asian Pacific Journal of Tropical Medicine

apjtm.org

doi: 10.4103/1995-7645.320523

5-Year Impact Factor: 2.285

Phylogeny of *Brucella abortus* strains isolated in the Russian Federation

Dmitry A. Kovalev✉, Dmitriy G. Ponomarenko, Sergey V. Pisarenko, Nikolay A. Shapakov, Anna A. Khachaturova, Natalia S. Serdyuk, Olga V. Bobrysheva, Alexander N. Kulichenko

Stavropol Research Anti-Plague Institute, 355035, Stavropol, Russian Federation

ABSTRACT

Objective: To study *Brucella (B.) abortus* strains isolated in the Russian Federation, in order to identify their detailed position in the phylogenetic structure of the species global population as well as to determine genetic relationships for isolates from different geographical areas.

Methods: Based on Bayesian method, the whole genome single-nucleotide polymorphism (SNP) analysis of 258 *B. abortus* strains from different geographical areas of the world including 20 *B. abortus* strains isolated in Russia was carried out.

Results: The core genome SNP analysis of the *B. abortus* isolates allowed describing the main genetic lineages. The Russian strains entered two separate clades, including the basal branch and the C1 branch that is widely spread in Eurasia. The data on the isolation time was used for the dating of phylogenetic tree, and also the estimated time frame for the *B. abortus* genotype diversification was determined. There were sets of specific SNPs identified, which defined each of the genotypes and sub-genotypes.

Conclusions: A significant genetic diversity of the brucellosis pathogen strains from Russia has been proven. The sets of unique specific SNPs described in our study may become one of the elements within a bio-informational analysis algorithm to be used for epidemiological study of brucellosis outbreaks, including those caused by new (atypical) genetic variants of *B. abortus*.

KEYWORDS: *Brucella abortus*; Phylogeny; Evolution; Genome; SNP

1. Introduction

Brucellosis is an extremely dangerous, zoonotic infectious disease which may entail severe social and economic effects and is caused by a bacterium belonging to the *Brucella* genus including 12 species of microorganisms. The ones that are considered the most significant epidemiologically include *Brucella (B.) melitensis*, *B. abortus* and *B. suis*[1]. The typical host for *B. abortus* is cattle[2,3] and this *Brucella* species is to be found along with livestock basically all over the world.

In 2018, multilocus sequence typing data was used to show that the *B. abortus* species falls into 4 main clades[4], which correspond to the genetic lines of A, B, C1 and C2 as described by Whatmore *et al*[5]. At the same time, the data concerning the phylogeny of the *B. abortus* strains circulating in Russia is scarce.

This study aims at carrying out a whole-genome analysis of *B. abortus* strains isolated in the Russian Federation in order to identify their detailed position in the phylogenetic structure of the species global population, and to determine genetic relationships for isolates from different geographical areas.

2. Materials and methods

2.1. Bacterial strains

A total of 20 strains of *B. abortus* used throughout the study were obtained from the pathogenic microorganism collection of the Stavropol Anti-Plague Institute. Table 1 offers more detailed information regarding the biochemical properties of the isolates.

Significance: The genetic structure of the global population and the evolution of *B. abortus* have been studied insufficiently. We first time present the model of *B. abortus* evolution based on whole genome sequencing data. The place of the Russian *B. abortus* strains in the structure of the global population of the species has been determined. We proposed the hypothesis of the penetration and further spread ways of *B. abortus* on the territory of Russia.

✉To whom correspondence may be addressed. E-mail: kovalev_da.stv@list.ru

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-Non Commercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com

©2021 *Asian Pacific Journal of Tropical Medicine* Produced by Wolters Kluwer-Medknow. All rights reserved.

How to cite this article: Kovalev DA, Ponomarenko DG, Pisarenko SV, Shapakov NA, Khachaturova AA, Serdyuk NS, et al. Phylogeny of *Brucella abortus* strains isolated in the Russian Federation. Asian Pac J Trop Med 2021; 14(7): 323-329.

Article history: Received 16 March 2021 Revision 16 June 2021
Accepted 17 June 2021 Available online 12 July 2021

Table 1. Biochemical properties of *Brucella abortus* isolates.

Strain	Location, year	Lysis by phage Tb	Oxidase	Catalase	Urease	H_2S production	Thionin	Basic fuchsin	Agglutination with monospecific antisera		
									A	M	R
Brucella abortus 240	Russia, Stavropol region, 1963	+	+	+	+	+	+	+	+	-	-
Brucella abortus 264	Russia, Karachayevo-Cherkess Autonomous Region, 1965	+	+	-	-	+	+	+	+	-	-
Brucella abortus 293	Russia, Stavropol region, 1967	+	+	-	-	+	+	+	+	-	-
Brucella abortus 313	Russia, Karachayevo-Cherkess Autonomous Region, 1970	+	+	-	+	+	+	+	+	-	-
Brucella abortus 317	Russia, Stavropol region, 1970	+	+	+	-	+	+	+	+	-	-
Brucella abortus 381	Russia, Stavropol region, 1971	+	+	+	+	+	+	+	+	-	-
Brucella abortus 390	Russia, Karachayevo-Cherkess Autonomous Region, 1971	+	+	+	+	+	+	+	+	-	-
Brucella abortus 401	Russia, Stavropol region, 1971	+	+	-	+	+	+	+	+	-	-
Brucella abortus 403	Russia, Stavropol region, 1971	+	+	+	+	+	+	+	+	-	-
Brucella abortus 420	Armenia, Martunin district, 1972	+	+	+	+	+	+	+	+	+	-
Brucella abortus 1552	Russia, Rostov Region, 1962	+	+	+	+	+	+	+	-	+	-
Brucella abortus I-2	Russia, Irkutsk, 1945	+	+	+	+	-	+	+	-	+	-
Brucella abortus I-12	Russia, Irkutsk region, 1948	+	+	+	+	+	+	+	+	-	-
Brucella abortus I-29	Russia, Kyzyl, 1958	+	+	+	+	+	+	+	+	-	-
Brucella abortus I-34	Russia, Khabarovsk region, 1958	+	+	+	+	+	+	+	+	-	-
Brucella abortus I-181	Russia, Novosibirsk, 1982	+	+	+	+	+	+	+	+	-	-
Brucella abortus S-550	Russia, Republic of Kalmykia, 2012	+	+	+	+	+	+	+	+	-	-
Brucella abortus S-551	Russia, Republic of Kalmykia	+	+	+	+	+	+	+	+	-	-
Brucella abortus S-577	Russia, Republic of Dagestan	+	+	+	-	+	±	+	+	+	-
Brucella abortus S-587	Russia, Stavropol region	+	+	+	-	+	+	+	+	+	-

2.2. Cultivation conditions and DNA extraction

The bacteria were cultured on *Brucella* agar for 48 h at a temperature of 37 ℃. The microbial suspension (concentration 2×10^9 m.sub./mL) was decontaminated through adding sodium merthiolate to a final concentration of 0.01% and was further incubated at 56 ℃ for 30 min. Genomic DNA was isolated from 0.5 mL of decontaminated microbial suspension using PureLink Genomic DNA Kits (Life Technologies, USA). The concentration of the genomic DNA was determined with the Qubit 2.0 fluorimeter and the Qubit dsDNA HS Assay Kit (Invitrogen, Life Technologies, USA). The genomic DNA purity was evaluated with the NanoDrop 2000 spectrophotometer (Thermo Scientific, USA). All the manipulations with active strains were performed in a Level 3 biosafety laboratory. The obtained DNA preparations were stored at -20 ℃ until further use.

2.3. DNA sequencing

Genomic libraries with a ~400 bp read length were prepared using the Ion X PressTM Plus Fragment Library Kit (Life Technologies, USA) following the manufacturer's protocol. Genome sequencing was done with the Ion Torrent PGM sequencer and the Ion 318 Chips Kit V2 chips (Life Technologies, USA). Once the quality assessment and data filtering procedures were complete, the resulting reads were collected into incomplete genome projects using SPAdes 3.12.0[6]. The S19 *B. abortus* strain genomic sequence (GeneBank: NC_010742.1, NC_010740.1) was used as a reference to assess the accuracy and efficiency of the genomic project assembly. The genome projects were annotated using the NCBI Prokaryotic Genome Annotation Pipeline.

2.4. Whole genome SNP analysis

Multiple genome alignment of 210 strains was performed in the REALPHY 1.10 software program[7] on default settings. To construct the core genome, 20 *B. abortus* genomes that we had sequenced were used, as well as 190 openly available *B. abortus* genomic sequences including complete genomes and genomic projects. The SNP search and identification in the core genome was carried out in the Mega 10 software program[8]. The site database of *Brucella* strain nucleotide polymorphisms is presented in Supplementary Table S1.

2.5. Phylogeographical and evolutionary analysis

The BEAST 2.3.0 software package was employed to study the phylogeographic taxon distribution based on whole-genome SNP analysis of *B. abortus* strains[9].We used the known *B. abortus* strain isolation dates to ensure the phylogenetic tree dating accuracy. The evolutionary model parameters were identified using the Jmodeltest 2 software[10], whereas Bayesian Markov chain Monte Carlo (MCMC) analysis was performed relying on the general time reversible +I+G (T+I+G) and the strict clock model. Three independent runs were held, with a chain length of 250 000 000 as per each run, and with

a recording rate of every 1 000 generations. The convergence of the MCMC topology and parameters was evaluated in the Tracker 1.6 software program[11]. The trees were combined through the TreeAnnotator component from the BEAST 2.3.0 software package, thus obtaining a consensus tree, while the burn-in parameter for each chain was set at 20%. The 95% HPD Interval (confidence interval) value for the clock rate parameter was $5.678\ 8\times10^{-8}$-$9.273\ 7\times10^{-8}$.

3. Results

3.1. General results

The genomic sequences of the 20 *B. abortus* strains isolated in Siberia and in South European Russia were obtained *via* high-performance sequencing using the IonTorrent PGM platform (Life Technologies, USA). The generated readings were collected based on a *de novo* approach into incomplete genomic projects. The genomic projects annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The resulting genomic projects were deposited in the GenBank database. Table 2 offers a view at the general features of the genomes.

The genomic sequences of 258 *B. abortus* strains were used here, including 20 genomes of isolates obtained through this study as well as 238 genomic sequences from the GenBank international database. We used all the complete genomes and whole genome shotgun (WGS) projects that were available at the time the study was held. It was a close phylogenetic relationship described previously between *B. abortus* and *B. melitensis* species[7], which served the determining factor for taking the genomic sequence of the *B. melitensis* 16M strain as an external group. The data on the genomic sequences of the strains used through the work can be seen from Supplementary Table S2.

We used REALPHY 1.10 with default settings to construct the multiple alignment matrix of the strain genomes[12]. The algorithms implemented in REALPHY allow obtaining a multiple genome alignment matrix that contains only orthologous nucleotide sequences (core genome). Some paralogous sequences are excluded from multiple alignment. The resulting multiple alignment matrix of complete genomes was used to construct the *B. abortus* phylogeny. The phylogenetic reconstruction was done using the BEAST 2.3.0 software package[9].

In order to increase the phylogenetic tree dating accuracy, the core genome matrix had all those strain sequences excluded, where we failed to reliably identify the isolation date. All further manipulations, including phylogenetic and evolutionary analysis, along with SNP analysis, were performed relying on an edited core genome matrix, which contained the sequences of 209 strains. A phylogenetic tree offering a description of the evolutionary relationships of the studied strains can be seen in Supplementary Figure S1. Figure 1 presents a fragment of the phylogenetic tree with evolutionary relationships of strains isolated in Russia.

3.2. Phylogenetic and evolutionary analysis

The resulting dendrogram is divided into 4 major genetic lineages in accordance with existing ideas about the structure of the global population of *B. abortus* (Figure S1). The basal branch is represented by strains from Asia. *B. abortus* strains 63/294 and 88/217, isolated

Table 2. Characteristics of *Brucella abortus* genomic projects.

Strain	GenBank accession number	No. of contigs	N50, bp	Total length, bp	GC, %	Genes (total)	CDS (total)	Proteins	Pseudo genes (total)
Brucella abortus 240	GCA_009728165.1	126	135 325	3 304 099	56.90	3 214	3 156	2 917	239
Brucella abortus 264	GCA_009733205.1	49	200 173	3 258 890	57.20	3 135	3 078	2 874	204
Brucella abortus 293	GCA_009733215.1	47	181 435	3 251 252	57.20	3 130	3 076	2 907	169
Brucella abortus 313	GCA_009733195.1	47	196 288	3 263 190	57.20	3 149	3 089	2 887	202
Brucella abortus 317	GCA_009733175.1	48	139 208	3 251 364	57.20	3 130	3 070	2 875	200
Brucella abortus 381	GCA_009733255.1	44	178 450	3 252 284	57.20	3 135	3 079	2 871	208
Brucella abortus 390	GCA_009728185.1	42	152 326	3 252 594	57.20	3 138	3 082	2 898	184
Brucella abortus 401	GCA_009733275.1	87	63 265	3 249 242	57.30	3 136	3 083	2 857	226
Brucella abortus 403	GCA_009733285.1	42	181 394	3 250 683	57.20	3 130	3 075	2 866	209
Brucella abortus 420	GCA_009733315.1	42	174 395	3 281 853	57.20	3 146	3 090	2 886	204
Brucella abortus 1552	GCA_009733335.1	60	127 514	3 250 336	57.20	3 137	3 083	2 909	174
Brucella abortus I-2	GCA_009733415.1	53	138 078	3 280 460	57.20	3 149	3 093	2 904	189
Brucella abortus I-12	GCA_009733435.1	42	189 940	3 250 398	57.20	3 132	3 076	2 868	208
Brucella abortus I-29	GCA_009733455.1	46	154 286	3 252 659	57.20	3 132	3 077	2 879	198
Brucella abortus I-34	GCA_009733465.1	40	216 664	3 251 995	57.20	3 132	3 076	2 884	192
Brucella abortus I-181	GCA_009733495.1	47	152 334	3 252 762	57.20	3 136	3 081	2 888	193
Brucella abortus S-550	GCA_009733365.1	38	216 624	3 252 314	57.20	3 130	3 074	2 857	217
Brucella abortus S-551	GCA_009733375.1	48	121 829	3 250 916	57.20	3 139	3 083	2 900	183
Brucella abortus S-577	GCA_009733355.1	63	105 242	3 279 925	57.20	3 147	3 093	2 864	229
Brucella abortus S-587	GCA_009733515.1	43	181 706	3 282 426	57.20	3 147	3 091	2 890	201

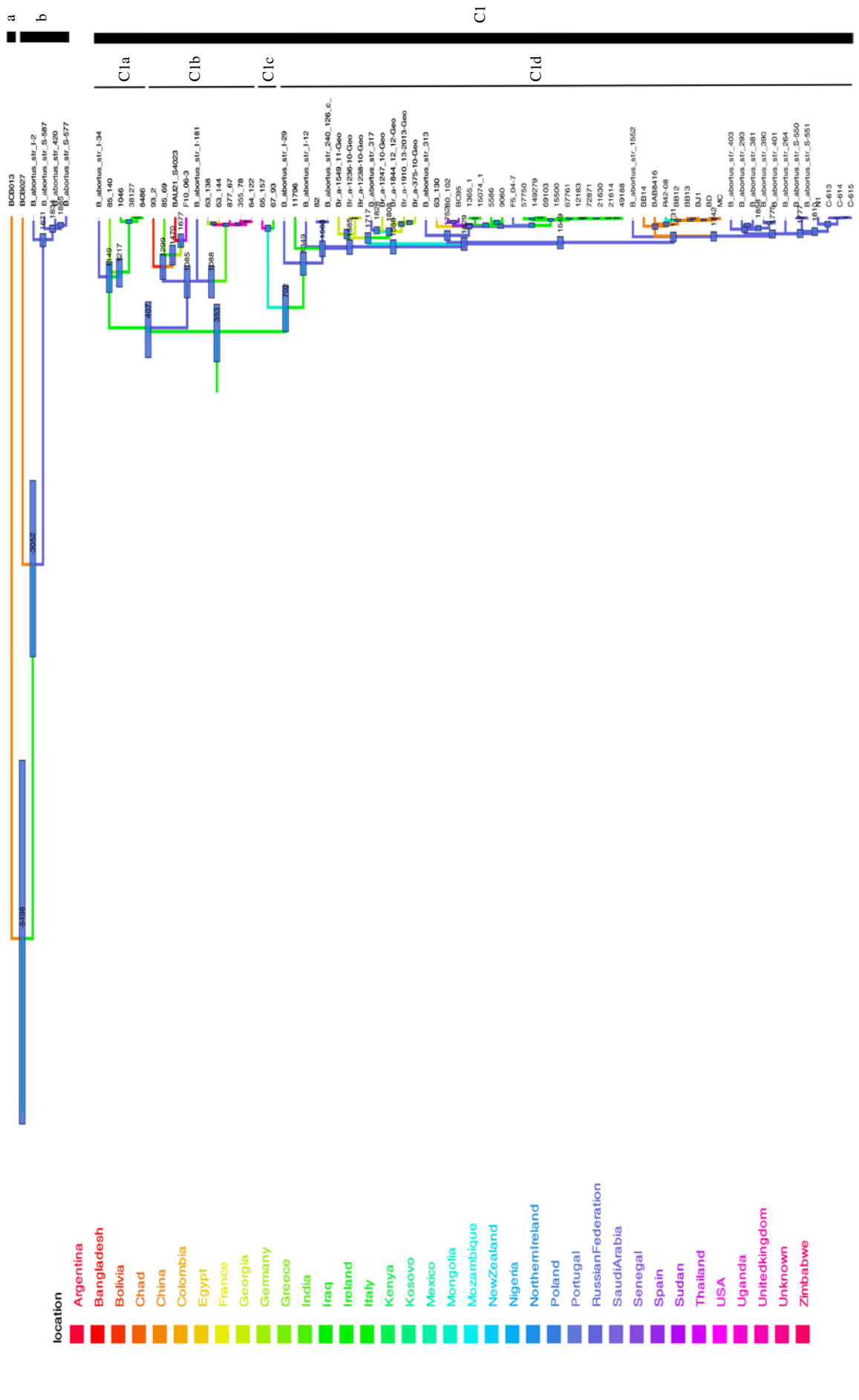

Figure 1. A fragment of the phylogenetic tree with evolutionary relationships of strains isolated in Russia. A phylogenetic tree describing the evolutionary relationships of *Brucella abortus* strains is shown in Supplementary Figure S1. The tree is reconstructed based on data of wgSNP analysis of *Brucella abortus* strains.

in Kenya and Mozambique, respectively, belong to clade A. Other African strains belong to clade B. A representative group, which includes most of the Russian strains, is assigned to clade C1. Clade C2 includes the main part of isolates from North and South America with the *B. abortus* 544 strain.

The Russian strains entered two separate clades, including the basal branch and the C1 branch, which is common in Eurasia.

The formation of the basal branch of the phylogenetic tree dates back to about 13 thousand years ago. The genotype includes strains of two sub-genotypes. Subgenotype *a* is represented by a single strain from China (BCB013, 2010), for which 1 661 specific SNPs were identified. Subgenotype *b*, differed by 579 SNP from other genotypes of the species, includes strains isolated in China (BCB027, 1983), in Armenia (420, 1972), in Russian Siberia (I-2, Irkutsk, 1945) and in South European Russia (C-587, Stavropol Territory, 2015; C-577, Republic of Dagestan, 2015). In general, 805 specific SNPs were identified for the strains of this genotype. Comparison within subgenotype *b* genomes revealed 1 331 SNPs.

Based on the results of the analysis, the C1 genotype must have developed around the second half of the 6th Century BC. The genotype includes four subgenotypes marked as C1a-C1d. We found 115 clade-specific SNPs for the specified genotype strains.

The C1a subgenotype, which separated in the middle of the 4th Century AD, includes strains isolated in Russia, Greece and Italy (96 specific SNPs). When comparing the strains within the genotype, 345 SNPs were described.

The subgenotype C1b subgroup, which includes strains isolated in Bolivia, India, Bangladesh, and Thailand, demonstrates the phylogenetic tree topology described earlier[13]. Besides, the subgenotype includes a strain from Russia (I-181, Novosibirsk, 1982), as well as strains from Germany, France and the Great Britain. The C1b subgenotype differs from other genotypes and subgenotypes by a set of 96 SNPs. While comparing strains within the subgenotype, 622 SNPs were identified.

The extensive subgenotype C1d includes strains from Russia (I-29, I-12, 82, 240) and Italy (11796). The strains from Georgia made up a common subclade with the strain from Russia (317). A separate group was made up by strains from Russia (313), Egypt, Spain, Italy and Portugal. It is notable that most of the studied strains isolated in the south of Russia's European part (Stavropol Territory, Republic of Kalmykia, Rostov Region) made up one subclade with strains from China and Mongolia. A set of 100 specific SNPs was detected for the C1d subgenotype. Comparison of the strains within the genotype allowed detecting 1 011 SNPs.

4. Discussion

The epidemiological situation connected with brucellosis in the Russian Federation over the past 10 years can be described as unfavorable featuring a decreasing trend in the incidence rate. Within the period of 2011-2020, there were 3 507 cases of brucellosis newly registered among people in Russia. The average long-term number of cases is 350 cases per year, including 28 cases among children aged below 17. The average long-term intensive morbidity rate as per 100 thousand people made up 0.24, children under 17 accounting for 0.1 of this value[14,15]. The highest number of brucellosis cases affecting people was observed in the North-Caucasus Federal District (NCFD)-2 291 (65.3% of the total number of brucellosis cases in Russia through 2011-2020) and the Southern Federal District (SFD)-503 cases (14.3%).

The basal branch of the resulting phylogenetic tree, which separated about 13 thousand years ago, includes two subgenotypes. Subgenotype *a* is a strain isolated in China, whereas subgenotype *b* involves strains from China, Russia, and Armenia. The subgenotype *b* strain from China (BCB027) as well as the group of strains isolated in Siberia (I-2), European Russia's southern part (S-577, S-587) and Armenia (420) are likely to share a common origin, and must have diverged around 5 thousand years ago (3 000 BC) in the early Bronze Age. Within that period, large local blocks of communities of Eurasia's population developed, which were involved into active interaction. The main two blocks of human communities, which were mostly engaged in agriculture and animal breeding, were located in the south of the central Alpine-Himalayan mountain belt: Sayans-Altai-Pamir and Tien-Shan-Caucasus-Carpathians-Alps. In the northern part of the Eurasian steppes, large nomadic tribes of cattle-raisers developed, while later on, communities of settled cattle-breeders emerged in the steppes of Eurasia[16].

To confirm the existence of the described genetic line *B. abortus* from Asia requires additional verification of the species affiliation of strains, as well as a study of a more representative sample of isolates from this region.

The obtained data suggests that the "African" genotypes A and B have not actually reached global spread and, as before, circulate mainly on one continent, which is consistent with the data from previous studies[5]. In recent decades, Africa has demonstrated a persistent trend to spread brucellosis (the pathogen migration) from the north of the Mediterranean Sea southwards the south into the continent[17]. There has been an increase in the number of human brucellosis cases registered in the countries located in the central and eastern parts of the continent. The disease cases in humans are mostly associated with the consumption of raw milk and unpasteurized dairy products from cattle[18]. At the same time, nearly each year the EU countries register brucellosis cases among refugees coming from Africa and the Middle East[19].

The C1 genotype isolates, on the other hand, are extremely common from Portugal in Western Europe to Thailand in Southeast Asia. The deviation of the C1 genotype tree branch may have occurred around the second part of the 6th Century BC. The strains of this genotype can be assumed to have originated from a common ancestor of Mediterranean origin, while their spread across Europe, North Africa and the Middle East may have occurred during the Roman conquests. Proof to this hypothesis can be seen from the geographical distribution of the C1a-C1c subgenotype strains isolated in Russia, Italy, Greece, Germany, France, Great Britain, as well as in Uganda, Iraq and a number of Asian countries.

The *B. abortus* I-34 strain isolated from an aborted cow fetus in 1958 in the Khabarovsk Territory of Russia, definitely belongs to a separate branch of the C1a subgenotype, which also includes isolates from Greece and Italy. The availability (circulation) of closely

related brucellosis pathogen strains can be explained through their penetration from Asia together with brucellosis-contaminated cattle and further advance in the areas during the period of active trade going on between East and West, including that along the Great Silk Road from China to Rome through Persia, Parthia and the Middle East, which is a connection between East Asia and the Mediterranean (China and the Far East to the Middle East and Europe)[20,21].

The *B. abortus* I-181 strain isolated from human blood in 1982 in Novosibirsk (Russia) formed an individual C1b subgenotype branch. Apart from this strain, the subgenotype in question includes two branches: "South Asian" (India, Bangladesh, Thailand) and "European" (Germany, France, Great Britain). The significant genetic heterogeneity of this group of strains can be seen from 622 differentiating SNPs. One of the possible reasons behind the isolation of the C1a and C1b subgenotype strains in Siberia in the second part of the 20th Century could be the post-war practice of forced migration from rural areas together with livestock, including to the Krasnoyarsk Territory, Kemerovo and Irkutsk Regions, from the Baltic Republics, Western Ukraine as well as from the Caucasus[22]. Another factor that cannot be excluded is that *Brucella* strains may have been brought to Siberia, and then further to the East European plain, in the 1st Century BC-the time when trade between Asia and Europe was emerging, and the caravan routes ran from China to Siberia, including through the area currently known as the Trans-Baikal Territory, the Republic of Buryatia, the Irkutsk Region, the Krasnoyarsk Territory and the Novosibirsk Region[20,21].

The Eastern Mediterranean countries established active trade with Western Europe including that involving livestock products. Following the collapse of the Western Roman Empire, Mediterranean goods turned to be coming in specifically large supplies from the southwestern parts of Asia (mostly from the territory of present-day Turkey), Syria, and Egypt, while the areas in question featured poor status in terms of brucellosis contamination. At the same time, as feudalism was at its peak, from the 11th Century on, Western Europe was most active purchasing "eastern goods", including high-value animals and raw products[21].

The time of the extensive C1d subgenotype separation dates back to the early 8th Century. Within our study, this subgenotype is represented by 53 strains, 19 of them isolated in Russia. On the phylogenetic tree, the C1d subgenotype is represented by three large clades. The first clade, which separated from the main branch around the year 1630 during the military conflicts between Eastern Georgia, seeking its independence, and Turkey and Iran[23], embraces 7 strains isolated in Georgia along with a strain from Russia (Stavropol Territory, 1970). The second clade consists of strains from Italy, Spain, Portugal and Egypt, as well as a separate strain isolated in Russia (Karachay-Cherkess Republic, 1970). The third clade, which emerged in the middle of the 17th Century, includes a group of 7 strains of Mongolian and Chinese origin, as well as a group of strains isolated in the south of the European part of Russia from 1962 to 2018. The period of intensive strain diversification in the North Caucasus coincides with the time when Russian was involved in the lengthy and bloody war in the Caucasus (Circassian war; 1763-1864)[24].

Strains isolated in different parts of Russia (I-12, I-29, 82, 240), as well as strain 11796 from Italy, too, belong to the C1d subgenotype.

Unfortunately, we could not obtain data regarding the biovar-affiliation of all the studied strains. It is obvious, though, that a particular strain belonging to one of the genotypes/subgenotypes based on wgSNP offers no clear connection with its biovar. The members of biovars 1 and 3, for instance, are assigned to different subgenotypes of the main genotypes. The discrepancy in the biovar affiliation and the position on the phylogenetic tree is accounted for by the fact that the biovar classification of *B. abortus* strains does not offer a full reflection of their genetic relationship or by possible inaccuracy in the biochemical determination of biovars[25].

In view of the above, the study can be seen as the first description for genomes of *B. abortus* strains circulating in Russia, to the best of our knowledge. Results of this study revealed a high degree of similarity among the whole-genome SNP profiles of *B. abortus* strains circulating within the same area, which allows employing the whole-genome SNP analysis as an effective tool used to identify the origin of certain individual isolates through epidemiological research.

The expansion of knowledge concerning the unique features of *B. abortus* isolates obtained in different countries, as well as the integration of the respective data into international databases would allow better use of whole-genome sequencing data used to identify the genetic link degree between different strains; to identify the geographical region of the infection origin; to detect the causes, the conditions of brucellosis occurrence in animals and humans, as well as to localize the infection focus.

Conflict of interest statement

The authors declare that there are no conflicts of interest.

Acknowledgements

We thank all departments and institutes for the *Brucella* genome sequences that were used in this study: National Center for Biotechnology RSE, Beijing Institute of Biotechnology, Beijing Institute of Disease Control and Prevention, Broad Institute, Central India Institute of Medical Sciences, China Agricultural University, China Animal Disease Control Center, China Institute of Veterinary Drug Control, the Colombian Corporation for Agricultural Research, Inner Mongolia Agricultural University, Indian Veterinary Research Institute, Istituto Zooprofilattico Sperimentale del Mezzogiorno, Madurai Kamaraj University, National Institute for Communicable Disease Control & Prevention, Chinese Center for Disease Control and Prevention, National Center for Disease Control and Public Health, Scientific Centre for Expert Evaluation of Medicinal Products of the Ministry of Health of the Russian Federation, Sardarkrushinagar Dantiwada Agricultural University,

Sardarkrushinagar Dantiwada Agricultural University, Shanghai JiaoTong University School of Medicine, University of New Hampshire, University of Sharjah, Virginia Bioinformatics Institute, the Wellcome Trust Sanger Institute.

Authors' contributions

DAK has developed a project and a research plan. DAK and ANK compiled the manuscript. DGP, AAK and NSS conducted bacteriological studies. SVP and OVB performed sequencing, genomes assembly and annotation. DAK, SVP and NAS conducted phylogenetic analysis. All authors have read and approved the final manuscript.

References

[1] Rajendhran J. Genomic insights into *Brucella*. *Infect Genet Evol* 2021; **87**: 104635.

[2] Plumb GE, Olsen SC, Buttke D. Brucellosis: 'One Health' challenges and opportunities. *Rev Sci Tech Off Int Epiz* 2013; **32**: 271-278.

[3] Frank KA, Krecek RC, Häsler BN, Arenas-Gamboa AM. Brucellosis remains a neglected disease in the developing world: A call for interdisciplinary action. *BMC Public Health* 2018; **18**(1): 125.

[4] Vergnaud G, Hauck Y, Christiany D, Daoud B, Pourcel C, Jacques I, et al. Genotypic expansion within the population structure of classical *Brucella* species revealed by MLVA16 typing of 1404 *Brucella* isolates from different animal and geographic origins, 1974-2006. *Front Microbiol* 2018; **9**: 1545.

[5] Whatmore AM, Koylass MS, Muchowski J, Edwards-Smallbone J, Gopaul KK, Perrett LL. Extended multilocus sequence analysis to describe the global population structure of the genus *Brucella*: Phylogeography and relationship to biovars. *Front Microbiol* 2016; **7**: 2049.

[6] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**(5): 455-477.

[7] Foster JT, Beckstrom-Sternberg SM, Pearson T, Beckstrom-Sternberg JS, Chain PS, Roberto FF, et al. Whole-genome-based phylogeny and divergence of the genus *Brucella*. *J Bacteriol* 2009; **191**(8): 2864-2870.

[8] Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 2018; **35**(6): 1547-1549.

[9] Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, et al. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput Biol* 2014; **10**(4): e1003537.

[10]Darriba D, Taboada GL, Doallo R, Posada D. jModelTest 2: More models, new heuristics and parallel computing. *Nat Methods* 2012; **9**(8): 772.

[11]Rambaut A, Suchard MA, Xie D, Drummond AJ. *Tracer v 1.6. Computer program and documentation distributed by the author. 2014.* [Online]. Available from: http://beast.bio.ed.ac.uk/software/tracer. [Accessed on 21 February 2020].

[12]Bertels F, Silander OK, Pachkov M, Rainey PB, van Nimwegen E. Automated reconstruction of whole-genome phylogenies from short-sequence reads. *Mol Biol Evol* 2014; **31**(5): 1077-1088.

[13]Islam MS, El Zowalaty ME, van Vliet AHM, Thakur S, Khatun MM, Saha S, et al. First genome sequence of *Brucella abortus* biovar 3 strain BAU21/S4023, isolated from a dairy cow in Bangladesh. *Microbiol Resour Announc* 2019; **8**(24): e00446-19.

[14]Ponomarenko DG, Rusanova DV, Khachaturova AA, Skudareva ON, Logvinenko OV, Rakitina EI, et al. Analysis of the epidemic and epizootic situation on Brucellosis around the world in 2019 and the forecast for the Russian Federation for 2020. *Probl Particularly Dangerous Infect* 2020; **2**: 48-56. (In Russ.) doi: https://doi.org/10.21055/0370-1069-2020-2-48-56.

[15]The Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing [Федеральная служба по надзору в сфере защиты прав потребителей и благополучия человека]. *Review of the epidemic and epizootic situation of brucellosis in the world in 2020 and forecast for 2021 in the Russian Federation (Cited 3 March 2021).* [Online]. Available from: https://www.snipchi.ru/updoc/2021/EPID_OBZOR_BRUZ_2020_2021.pdf. [Accessed on 21 February 2020]. (in Russ).

[16]Lagoida NG. *Ethnic History of Eurasia in the concept of ethnogenesis L.N. Gumileva. Evrazijstvo i mir. 2014; 3.* [Online]. Available from: https://cyberleninka.ru/article/n/etnicheskaya-istoriya-evrazii-v-kontseptsii-etnogeneza-l-n-gumileva. [Accessed on 21 February 2020]. (in Russ).

[17]Govindasamy K. Human brucellosis in South Africa: A review for medical practitioners. *S Afr Med J* 2020; **110**(7): 646-651.

[18]Siouane Z. *Tiaret: Plus de 150 cas de brucellose depuis janvier.* [Online]. Available from: https://www.algerie360.com/tiaret-plus-de-150-cas-de-brucellose-depuis-janvier. [Accessed on 21 February 2020].

[19]Wang XH, Jiang H. Global prevalence of human brucellosis. *Zhonghua Liu Xing Bing Xue Za Zhi* 2020; **41**(10): 1717-1722.

[20]History.com editors (eds). *Silk Road.* [Online]. Available from: https://www.history.com/topics/ancient-middle-east/silk-road. [Accessed on 21 February 2020].

[21]Liu Z, Wang C, Wei K, Zhao Z, Wang M, Li D, et al. Investigation of genetic relatedness of *Brucella* strains in countries along the Silk Road. *Front Vet Sci* 2021; **7**: 539444.

[22]Zhigunova MA, Krott II (eds). *Migration processes in Siberia: Peoples, cultures, public policy. collection of scientific papers.* Omsk: KAN Publishing Center; 2018, p. 328.

[23]Imgrunt EV. Iran-Turkish conflicts in the 16-18 centuries. *Molodoy Uchenyi* 2014; **20**(79): 519-521.

[24]Russian foreign policy in the first half of the 19th century Russian politics on the eastern issue in the 1820s of the XIX century. Russia and the peoples of the North Caucasus. In: Orlov AS, Georgiev VA, Georgieva NG, Sivohina TA. (eds.) *History of Russia*. Moscow: Prospect; 2013, p. 218-219.

[25]Garofolo G, Di Giannatale E, Platone I, Zilli K, Sacchini L, Abass A, et al. Origins and global context of *Brucella abortus* in Italy. *BMC Microbiol* 2017; **17**(1): 28.