Original Study Article
https://doi.org/10.36233/0372-9311-487

Development of a technique for molecular typing of *Bacillus anthracis* strains using new VNTR and INDEL markers

Grigorii A. Pechkovskii[1]✉, Evgeny I. Eremenko[1], Alla G. Ryazanova[1], Sergey V. Pisarenko[1], Nikolay A. Shapakov[1], Lyudmila Yu. Aksenova[1], Olga V. Semenova[1], Lyudmila D. Timchenko[2], Alexander N. Kulichenko[1]

[1]Stavropol Plague Control Research Institute, Stavropol, Russia;
[2]North-Caucasus Federal University, Stavropol, Russia

Abstract

Introduction. *Bacillus anthracis*, the pathogen of a particularly dangerous zoonotic disease known as anthrax, requires strict epidemiological control and is characterized by high genetic homogeneity, which necessitates the development of genotyping methods.

The aim of the study were to to find and characterize the VNTR and INDEL loci of *B. anthracis* and to develop on their basis a genotyping technique by PCR with electrophoretic detection of the results.

Materials and methods. Marker search and phylogenetic analysis were performed on a sample of 388 genomes of *B. anthracis* strains, 322 from the GenBank collection (RefSeq) and 66 from the collection of the Stavropol Anti-Plague Institute of Rospotrebnadzor. Phylogenetic analysis was performed on the basis of SNP crustal alignment using the Parsnp program. The search for markers was carried out using the Mauve program and author's scripts in Python. PCR was performed using a ScreenMix-HS kit (CJSC "Eurogen", Russia).

Results. Genomic variations of *B. anthracis* strains (SNP — 25,664, SNR — 14,387, VNTR — 693, INDEL — 14,667) were found, bioinformatic analysis of which revealed nine new VNTR and six INDEL molecular markers most suitable for genotyping. The genetic (allelic) variants of the markers are described. Primers were selected for the found markers and a PCR protocol with detection by electrophoresis in agarose gel was developed. When typing using VNTR markers was applied, the strains were divided into nine clusters: A.Br.Ames, A.Br.001/002, A.Br. Aust94, A.Br.005/006, A.Br.008/009 (Tsiankovskii), A.Br.008/009 (STI), A.Br.008/009 (A.Br.125), A.Br.008/009 (strain 228/269), B.Br.001/002. When typing using INDEL markers, the strains were divided into six clusters: A.Br.Ames, A.Br.001/002, A.Br.Aust94, A.Br.008/009(Tsiankovskii), B.Br.001/002(B.Br.014), as well as a cluster comprising several genetic lineages: A.Br.008/009 (STI), A.Br.008/009 (A.Br.125), A.Br.005/006 и B.Br.001/002.

Conclusion. The use of the developed methodology for the identification of variable VNTR and INDEL loci makes it possible to reliably determine the phylogenetic position of *B. anthracis* strains and is promising for use in the epidemiological investigation of anthrax outbreaks.

Keywords: *genotyping, VNTR, INDEL, Bacillus anthracis, whole genome sequencing, phylogenetic analysis*

Funding source. This study was not supported by any external sources of funding.

Conflict of interest. The authors declare no apparent or potential conflicts of interest related to the publication of this article.

For citation: Pechkovskii G.A., Eremenko E.I., Ryazanova A.G., Pisarenko S.V., Shapakov N.A., Aksenova L.Yu., Semenova O.V., Timchenko L.D., Kulichenko A.N. Development of a technique for molecular typing of *Bacillus anthracis* strains using new VNTR and INDEL markers *Journal of microbiology, epidemiology and immunobiology.* 2024;101(3):362–371.
DOI: https://doi.org/10.36233/0372-9311-487
EDN: https://www.elibrary.ru/kjnhyg

Оригинальное исследование
https://doi.org/10.36233/0372-9311-487

Разработка методики молекулярного типирования штаммов *Bacillus anthracis* с использованием новых VNTR- и INDEL-маркеров

Печковский Г.А.[1]✉, Еременко Е.И.[1], Рязанова А.Г.[1], Писаренко С.В.[1], Шапаков Н.А.[1], Аксенова Л.Ю.[1], Семенова О.В.[1], Тимченко Л.Д.[2], Куличенко А.Н.[1]

[1]Ставропольский противочумный институт Роспотребнадзора, Ставрополь, Россия;
[2]Северо-Кавказский федеральный университет, Ставрополь, Россия

Аннотация

Введение. *Bacillus anthracis* — возбудитель особо опасного зооноза сибирской язвы — отличается высокой генетической однородностью, что вызывает необходимость совершенствования методов генотипирования.

Целями исследования были поиск, описание VNTR- и INDEL-локусов *B. anthracis* и разработка на их основе методики генотипирования посредством полимеразной цепной реакции (ПЦР) с электрофоретической детекцией результатов.

Материалы и методы. Поиск VNTR- и INDEL-маркеров и филогенетический анализ выполняли на выборке из 388 геномов штаммов *B. anthracis*: 322 из GenBank (RefSeq) и 66 — из коллекции Ставропольского противочумного института. Филогенетический анализ проводили на основе SNP корового выравнивания с помощью программы «Parsnp». Поиск маркеров осуществляли с использованием программы «Mauve» и авторских скриптов на языке Python. ПЦР выполняли с помощью набора «ScreenMix-HS».

Результаты. Найдены геномные вариации штаммов *B. anthracis* (SNP — 25 664, SNR — 14 387, VNTR — 693, INDEL — 14 667), биоинформатический анализ которых позволил выявить 9 новых VNTR и 6 INDEL молекулярных маркеров, наиболее подходящих для генотипирования. Описаны генетические (аллельные) варианты маркеров. Для найденных маркеров подобраны праймеры и разработан протокол ПЦР с детекцией методом электрофореза в агарозном геле. В результате кластеризации при типировании с использованием VNTR-маркеров штаммы разделялись на 9 кластеров: A.Br.Ames, A.Br.001/002, A.Br.Aust94, A.Br.005/006, A.Br.008/009 (Tsiankovskii), A.Br.008/009 (STI), A.Br.008/009 (A.Br.125), A.Br.008/009 (штамм 228/269), B.Br.001/002. При типировании с применением INDEL-маркеров штаммы разделялись на 6 кластеров: A.Br.Ames, A.Br.001/002, A.Br.Aust94, A.Br.008/009 (Tsiankovskii), B.Br.001/002(B.Br.014), а также кластер, включающий представителей нескольких генетических групп: A.Br.008/009 (STI), A.Br.008/009 (A.Br.125), A.Br.005/006 и B.Br.001/002.

Заключение. Использование разработанной методики идентификации вариабельных VNTR- и INDEL-локусов позволяет достоверно определять филогенетическое положение штаммов *B. anthracis* и перспективно для применения в процессе эпидемиологического расследования вспышек сибирской язвы.

Ключевые слова: *генотипирование, VNTR, INDEL Bacillus anthracis, полногеномное секвенирование, филогенетический анализ*

Источник финансирования. Авторы заявляют об отсутствии внешнего финансирования при проведении исследования.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Для цитирования: Печковский Г.А., Еременко Е.И., Рязанова А.Г., Писаренко С.В., Шапаков Н.А., Аксенова Л.Ю., Семенова О.В., Тимченко Л.Д., Куличенко А.Н. Разработка методики молекулярного типирования штаммов *Bacillus anthracis* с использованием новых VNTR- и INDEL-маркеров. *Журнал микробиологии, эпидемиологии и иммунобиологии.* 2024;101(3):362–371.
DOI: https://doi.org/10.36233/0372-9311-487
EDN: https://www.elibrary.ru/kjnhyg

DOI: https://doi.org/10.36233/0372-9311-487

Introduction

Bacillus anthracis is a spore-forming Gram-positive bacillus, the pathogen of anthrax, a particularly dangerous infection with a global distribution area. In early studies, attempts to determine the genetic variability of *B. anthracis* were unsuccessful, indicating the high genetic monomorphism of this species [1]. The first genetic marker suitable for differentiation of *B. anthracis* strains was tandem repeats in the chromosomal locus *vrrA*, which are sequentially repeated identical DNA fragments (variable number tandem repeats, VNTR) [2]. Allelic variants of *vrrA* with the number of repeats from 2 to 6 allowed all strains to be divided into 5 groups [2, 3]. The marker was included in the first typing scheme by MLVA8 (Multiple loci VNTR analysis, MLVA) consisting of 6 chromosomal and 2 plasmid VNTR loci. VNTR loci in general differ from other variable regions in the fact they have a high frequency of variability and a larger number of variants, as well as the manifestation of the homoplasy effect, i.e. independent or parallel mutations in different genetic lineages [4]. Therefore, genotyping based on the analysis of VNTR loci makes it difficult to study intraspecific evolution, but this method is convenient for epidemiologic investigation of anthrax outbreaks. Active searches for *B. anthracis* loci with tandem repeats led to the discovery of 32 VNTR markers in 6 MLVA genotyping schemes [5–10].

In order to study genetic diversity, a genotyping method based on the analysis of canonical SNPs (canSNP typing) with the identification of 12 major genetic lineages was developed and tested on a large selection of strains [8]. The canonical lineages most accurately reflect the evolutionary groups of *B. anthracis*; therefore, they are best suited to describe the distribution of anthrax strains in the world. Subsequently, large-scale phylogenetic studies were performed with detailed descriptions, creation of a nomenclature of names and relationships of genetic clusters. Subclusters of canonical lineages were assigned numbers or trivial names [11, 12]. In particular, the canonical lineage A.Br.008/009 includes the Tsiankovskii and STI subclusters, which are widely represented in the Commonwealth of Independent States.

In 2019, an anthrax outbreak occurred in the Republic of Dagestan with isolates that clustered into a separate phylogenetic group A.Br.125 belonging to the STI subcluster.

The canonical lineage B.Br.001/002 contains the Siberia and Europe clusters, making up B.Br.014, as well as the Asia and B.Br.018 clusters.

Taking into account the established and newly identified designations of genetic lineages and groups, we used the following order in the subsequent description. The canonical lineage is indicated first, followed by a new subgroup or cluster within it with an established designation, if identified. For example, most strains of the main lineage A, isolated in Russia, are designated as A.Br.008/009 (Tsiankovskii) or A.Br.008/009 (STI).

Molecular markers also include INDELs (insertion/deletion), which are non-repeat variable regions that exist predominantly as two genetic variants: with deletion or with insertion.

For *Francisella tularensis*, an INDEL typing scheme was developed, including 38 INDEL loci. The study showed that the use of such markers increases the accuracy of typing [13]. Genotyping methods based on INDEL-locus analysis have also been developed for *Helicobacter pylori*, *Burkholderia pseudomallei*, *Vibrio cholerae*, *Yersinia pseudotuberculosis* and proved to be of high resolution and reliable in determining the phylogenetic position of strains [14-17]. Currently, the INDEL genotyping system for *B. anthracis* has not been developed.

The objectives of the study were to find and characterize VNTR and INDEL loci of *B. anthracis* and to develop on their basis a genotyping technique by PCR with electrophoretic detection of the results.

Materials and methods

Marker search and phylogenetic analysis were performed on a sample of 388 genomes of *B. anthracis* strains: 322 from the GenBank collection (RefSeq), 66 from the collection of pathogenic microorganism genomes of the Stavropol Anti-Plague Institute, described earlier [12]. Genome numbers are given in Appendix 1 on the journal website (https://doi.org/10.36233/0372-9311-487-s1). Genomic sequences of *B. anthracis* strains from the collection of the Stavropol Anti-Plague Institute have been deposited in the "National Interactive Catalog of Pathogenic Microorganisms and Biotoxins" (State Scientific Center for Applied Microbiology and Biotechnology).

Markers were searched with the help of an algorithm (Pipeline) consisting of pairwise alignment of complete genomes to a reference sequence using the Mauve program and subsequently, with the help of the author's scripts in the Python language, extraction of genetic variants from the alignments, merging and their analysis.

Marker verification and determination of marker lengths were performed in the BLASTn program using flanking sequences or specific primers.

In order to compare phylogenetic groups with genetic markers, a phylogenetic tree was constructed on the basis of core SNP alignment using the Parsnp program from the Harvest suit package with the reference genome of *B. anthracis* Ames Ancestor (GCF_000008445.1). Positions with the unknown N nucleotide were removed from the core SNPs. The SNPs from the VCF file were then converted to a FASTA file. The phylogenetic tree was constructed in the MEGA XI program using the maximum likelihood method with the Tamura-Nei substitution model [18].

Comparison of the lengths of genetic variants of markers with phylogenetic dendrogram and data visualization were performed in R language environment with ggtree and ggplot2 libraries.

Primers were designed using the Primer-BLAST program and synthesized at the Stavropol Plague Institute.

Sample preparation of *B. anthracis* cultures was carried out according to MG 1.3.2569-09 "Organization of laboratories using nucleic acid amplification methods when working with material containing microorganisms of pathogenicity groups I-IV". DNA extraction of *B. anthracis* was performed using the DNA-Sorb-B kit (ILC).

A representative phylogenetic sample of sequenced strains was used for verification of data obtained by PCR with electrophoretic detection. PCR was performed using the ScreenMix-HS kit (Eurogen). The amount of primers in the reaction was equal to 0.3 μM. The following thermocycling mode was used: the first stage (activation) — 95ºC, 5 min — 1 cycle, the second stage — denaturation 95ºC, 20 s, annealing 60ºC, 20 s, elongation 72ºC, 60 s — 40 cycles, the third stage (final elongation) — 72ºC, 5 min — 1 cycle. Electrophoresis was performed in 2% agarose gel using a 100 bp molecular weight marker (SibEnzyme).

Clustering of data obtained from PCR results with electrophoretic detection was performed using the single (Nearest Point Algorithm) method in Python with the scipy library.

Results and discussion

The study included two main stages. The first stage involved the search and description of marker loci, and the second stage involved experimental confirmation and validation of the typing technique using the found markers.

The algorithm resulted in the following genomic variations: SNP — 25,664, SNR — 14,387, VNTR — 693, INDEL — 14,667.

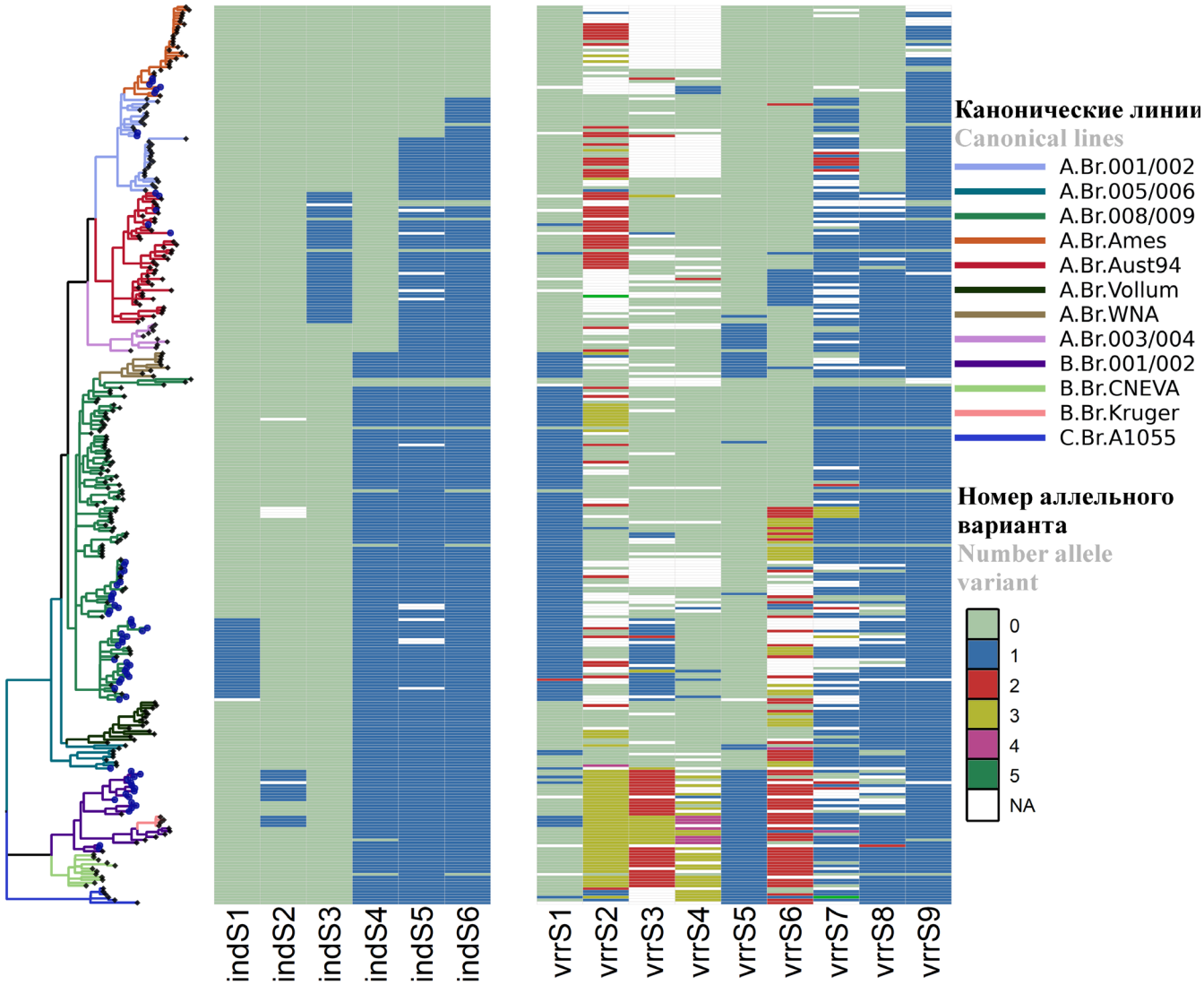

Fig. 1. Comparison of the phylogenetic tree and selected markers *B. anthracis* (some strains with repeated marker patterns were removed).

DOI: https://doi.org/10.36233/0372-9311-487

The search for markers among all found variations was performed stepwise with filtering according to a number of criteria. The difference in the size of variants of variable loci should be at least 15 bp. We selected mainly such loci with at least one genetic variant in the set of *B. anthracis* strains from the collection of pathogenic microorganisms of the Stavropol Anti-Plague Institute that was different from the variants of other strains. Variable loci already described were excluded. Using the criteria, 537 variable regions were found.

Following this, the frequency of allelic variants of loci of strains in certain genetic lineages was investigated by comparing the lengths of genetic variants of markers with the phylogenetic dendrogram based on SNPs of the bovine genome (**Fig. 1**). Most of the variant loci were found in only 1 strain or a minimal number of strains. A significant group consisted of variants sharing the major genetic lineages A, B, and C, including the previously found INDEL indE1 of 38 bp in length. [19], which is logical since these are the most evolutionarily distant genetic lineages.

Fifty-six VNTR and INDEL loci were selected (**Table 1**). The most phylogenetically significant and optimal for electrophoresis were selected from them. Thus, as a result, 9 VNTR markers and 6 INDEL markers most suitable for genotyping were selected (Fig. 1). The peculiarity of the INDELs found lies in the repeats flanking the INDELs, with one of the repeats being included in the deletion and the other not. For this reason, we can assume the formation of a complex structure between DNA strands during replication, which may cause the polymerase complex to mistakenly double the DNA strand by excising part of the sequence. In this case, reverse INDEL insertion is impossible, which likely reduces the effects of homoplasy.

The indels indS1 (FAD-*binding oxidoreductase gene*), indS2 (hypothetical protein gene *WP_000829051.1*), and indS6 (cell *surface protein gene*) localized in the protein-coding genes are genomic variations without a shift of reading frame. IndS4 localizes in the region between the *GBAA_RS02140* (ABC transporter ATP-binding protein) and *GBAA_RS02145* (ABC-F family ATP-binding cassette domain-containing protein) genes. IndS5 localizes in the region between the genes *GBAA_RS03470* (hypothetical protein) and *GBAA_RS03475* (alanine:cation symporter family protein). Indel indS3 shifts the reading frame of the gene encoding the SPFH/Band 7/PHB domain protein.

The indel indS1 genetic variant with deletion is characteristic of the A.Br.008/009 (Tsiankovskii) cluster. The indel indS2 is unusual in that there are 3 genetic variants of this locus: an insertion and 2 deletion variants. The difference between the two deletions is a shift of 9 nucleotide pairs. One variant is characteristic of the Siberia and Europe canSNP subgroups of the B.Br.001/002 group, the other is characteristic of strains of the B.Br.Kruger group. The indS3 locus variant with deletion is found in strains of the A.Br.Aust94 group, except for strain 9080-G isolated in Georgia and Kanchipuram strain from India. Variants indS4, indS5, indS6 with deletion are found in A.Br.004, A.Br.001 and A.Br.Ames clusters, respectively.

The number of allelic variants of the selected VNTR markers varies from 2 to 6 with repeat lengths ranging from 30 bp to 196 bp (Fig. 1). The vrrS1 locus has a 425-bp variant found in A.Br.008/009 and A.Br.WNA, as well as one unique 337-bp variant specific to strain 228/269. Genetic variants of the vrrS2 locus are found in the A.Br.008/009 and A.Br.Aust94 groups. Two VNTR markers, vrrrS3 and vrrS4, were found on the pXO2 plasmid. Separate genetic variants of vrrS3 are found in

Table 1. Description of the identified molecular *B. anthracis* markers

Marker	Coordinates of the locus in the genome according to the Ames Ancestor reference strain (GCF_000008445.1)	Replicon	The number of the allele variant (the length of the genetic variant, bp)
indS1	1276500–1276764	Chromosome	1 (265, 266), 2 (241)
indS2	1904893–1905267	Chromosome	1 (373–375), 2 (312)
indS3	1944246–1944531	Chromosome	1 (286), 2 (253)
indS4	402388–402715	Chromosome	1 (328), 2 (423–424)
indS5	655408–655662	Chromosome	1 (255), 2 (272, 284–285)
indS6	4691499–4691775	Chromosome	1 (277), 2 (388–389)
vrrS1	1721221–1721733	Chromosome	1 (513), 2 (425)
vrrS2	4489063–4489484	Chromosome	1 (422), 2 (381), 3 (299,307), 4 (217), 5 (258) 6 (338–340)
vrrS3	8316–8860	pXO2	1 (544–546), 2 (301–302), 3 (464), 4 (383)
vrrS4	8916–9269	pXO2	1 (354–355), 2 (263–264), 3 (444), 4 (534), 5 (174)
vrrS5	3155556–3155727	Chromosome	1 (172), 2 (142)
vrrS6	1092722–1092959	Chromosome	1 (238), 2 (198), 3 (318–319), 4 (398), 5 (278)
vrrS7	5088417–5088723	Chromosome	1 (306–307), 2 (190), 3 (229), 4 (385), 5 (346), 6 (268) (385)

End of the Table 1

Marker	Coordinates of the locus in the genome according to the Ames Ancestor reference strain (GCF_000008445.1)	Replicon	The number of the allele variant (the length of the genetic variant, bp)
vrrS8	5031546–5031803	Chromosome	1 (258, 263–265), 2 (354, 359–366)
vrrS9	3742896–3743541	Chromosome	1 (646), 2 (450)
indNS1	130607–131099	pXO1	1 (454, 456), 2 (494–495)
indNS2	596340–596832	Chromosome	1 (352), 2 (492–493)
indNS3	122138–122690	pXO1	1 (551–555), 2 (485, 487)
indNS4	77192–77540	pXO1	1 (330), 2 (349), 3 (619)
indNS5	482012–482157	Chromosome	1 (146,149), 2 (504)
indNS6	385564–385837	Chromosome	1 (271–276), 2 (305–308)
indNS7	1372136–1372298	Chromosome	1 (163), 2 (181)
indNS8	2559203–2559485	Chromosome	1 (282–284), 2 (335–336)
indNS9	3855034–3855252	Chromosome	1 (219), 2 (231), 3 (239–241)
indNS10	4303573–4303825	Chromosome	1 (253), 2 (310–311)
indNS11	4965875–4966088	Chromosome	1 (214), 2 (321)
indNS12	1209302–1209701	Chromosome	1 (253), 2 (399–401)
indNS13	2728738–2729257	Chromosome	1 (229), 2 (519–520)
indNS14	486258–486638	Chromosome	1 (285), 2 (381)
indNS15	1287411–1287701	Chromosome	1 (201), 2 (291)
indNS16	910496–910796	Chromosome	1 (301), 2 (490,491)
indNS17	2533966–2534193	Chromosome	1 (228), 2 (634–636)
indNS18	2593388–2593616	Chromosome	1 (228–230), 2 (283)
indNS19	3352013–3354229	Chromosome	1 (193,194), 2 (2124), 3 (2207, 2215–2218, 2223)
indNS20	3829833–3830053	Chromosome	1 (220–221), 2 (251)
indNS21	4811428–4811664	Chromosome	1 (236–237), 2 (600, 602)
indNS22	29253–29436	pXO1	1 (184), 2 (269)
indNS24	1146673–1147101	Chromosome	1 (256), 2 (270–272), 3 (427–430)
indNS25	2224848–2225376	Chromosome	1 (270), 2 (418), 3 (529–530, 537)
indNS26	2687438–2687847	Chromosome	1 (240–241), 2 (410,408–410), 3 (429) 4 (580)
indNS27	3304833–3305473	pXO1	1 (245, 257), 2 (640–641)
vrrNS1	226241–226786	Chromosome	1 (545–547), 2 (694, 697–699), 3 (845–847), 4 (997–998), 5 (1146), 6 (1296–1298)
vrrNS2	1333990–1334961	Chromosome	1 (343), 2 (554, 552), 3 (700), 4 (758, 762–763), 5 (779), 6 (971–974), 7 (1182–1183), 8 (1393)
vrrNS3	2014690–2015095	Chromosome	1 (277), 2 (364), 3 (406, 409), 4 (535)
vrrNS4	4233686–4234066	Chromosome	1 (237), 2 (273, 279), 3 (306, 309, 322), 4 (381), 5 (417), 6 (345)
vrrNS5	4351696–4351908	Chromosome	1 (213), 2 (231)
vrrNS6	4598742–4598948	Chromosome	1 (195, 207), 2 (171, 183)
vrrNS7	811781–812154	Chromosome	1 (284), 2 (302), 3 (320), 4 (374), 5 (428), 6 (482)
vrrNS8	1395847–1396186	Chromosome	1 (340), 2 (385)
vrrNS9	1238148–1238579	Chromosome	1 (361, 366), 2 (398), 3 (430–433), 4 (465), 5 (498)
vrrNS10	2264930–2265251	Chromosome	1 (244), 2 (283), 3 (322), 4 (361), 5 (439), 6 (517)
vrrNS11	4352078–4352327	Chromosome	1 (220), 2 (235), 3 (250, 251), 4 (264–266), 5 (295), 6 (310)
vrrNS12	4927425–4927645	Chromosome	1 (181), 2 (221)
vrrNS13	4769700–4770199	Chromosome	1 (499–501), 2 (352–353)
vrrNS15	1151194–1151463	Chromosome	1 (148), 2 (269–270), 3 (291), 4 (392–393, 396), 5 (514, 520)
vrrNS16	2006677–2007157	Chromosome	1 (481), 2 (433, 435–436), 3 (526), 4 (301), 5 (345–347), 6 (390–391), 7 (255–257)

DOI: https://doi.org/10.36233/0372-9311-487

A.Br.008/009 (Tsiankovskii), B.Br.KrugerB and main lineage B, respectively. The vrrS4 locus separates the strains into lineages A and B. The 142 bp allelic variant vrrS5 is found simultaneously in strains of lineage B, A.Br.WNA and A.Br.003/004 groups. The vrrS6 variants are characteristic of part of the strains of the A.Br.008/009 group, the A.Br.004 cluster and lineage B. The vrrrS7 307 bp genetic variant is specific for the A.Br.Ames group. The 258 bp vrrS8 variant is specific to the A.Br.Ames and A.Br.001/002 groups. The 646 bp vrrS9 genetic variant is specific for A.Br.Ames strains isolated in North America.

Variable loci can be grouped according to their belonging to certain genetic clusters. For example, variants of indels indS3, indNS27, and VNTR - vrrNS7 have similar belonging to the A.Br.Aust94 group. Indels indS4 and indNS11 are characteristic of A.Br.004, indNS17 and vrrS9 are characteristic of A.Br.Ames strains isolated in North America. Variants indNS5, indNS9, indNS10 are found in strains of the A.Br.Ames and A.Br.001/002 groups.

For the major lineage B, the characteristic loci are indNS1, indNS12, indNS19, indNS2, indNS3, indNS13, indNS14, and vrrNS12. Characteristic loci for both lineages B and C are indNS18, indNS21, indNS4, indNS6, indNS7 and indNS8.

Some of the unselected markers could also be used for typing. For example, vrrNS1 has high variability but a long repeat of 150 bp and a large length difference between the minimum and maximum genetic variant, which is difficult for electrophoretic detection by PCR. The VNTR of vrrrNS15 is variable within the A.Br.008/009 group. The tandem repeats vrrNS16, vrrNS2, vrrNS4 do not have strict specificity.

The primers were selected for the found markers (**Table 2**) and a PCR protocol with detection of the results by agarose gel electrophoresis was developed (**Fig. 2, Fig. 3**). The selected markers had a nucleotide sequence length sufficient for reliable determination of genetic variants of loci (**Table 3**).

Some strains lack the pXO2 plasmid, and accordingly, they also lack the vrrS3 and vrrS4 loci.

As a result of clustering based on the INDEL typing, strains were divided into 6 clusters: A.Br. Ames, A.Br.001/002, A.Br.Aust94, A.Br.008/009 (Tsiankovskii), B.Br.001/002 (B.Br.014), as well as a cluster including representatives of several genetic groups: A.Br.008/009 (STI), A.Br.008/009 (A.Br.125), A.Br.005/006 and B.Br.001/002. The cluster is singled out as a separate group because no specific INDEL markers have been identified for the strains of these lineages (**Fig. 4**).

As a result of clustering based on VNTR-typing, strains were divided into 9 clusters: A.Br.Ames, A.Br.001/002, A.Br.Aust94, A.Br.005/006, A.Br.008/009 (Tsiankovskii), A.Br.008/009 (STI), A.Br.008/009 (A.Br.125), A.Br.008/009 (strain 228/269), B.Br.001/002 (Fig. 4). Strain 228/269 is a member of the A.Br.008/009 (Tsiankovskii) group.

The discriminatory power determined using the Hanter–Gaston diversity index [20] was 0.7 for canSNP typing, and 0.79 and 0.84 for typing based on the analysis of new VNTR and INDEL markers, respectively.

Conclusion

Thus, as a result of genome analysis of 388 *B. anthracis* strains, variable regions were identified and characterized. New VNTR- and INDEL-markers were found and their binding to clusters of global phylogeny was studied. The developed protocol for marker iden-

Table 2. Primers to VNTR and INDEL loci *B. anthracis*

Name	Forward primer	Reverse prime
indS1	TATTGGGCAGCAGCATTTGG	ATGAGTTGTACGGGACGCAA
indS2	TGGAGGGGTTGTTCAAGCG	GCGTAACTCGGAGACCATGTA
indS3	AGCAACAGAAAAATGGGGCG	AATCGCTCTTGCTTCCCCTT
indS4	AGAAGGAACAAAAGGAAAAGTAGAG	CAACATGCTCGCCCTTCAAT
indS5	GGTCTATACGGCACACTCCA	GCTTCCAATATTCCCCCTCC
indS6	AGCCCCTTCTTTCGGTGTAT	CGATGAAGATGTAAGACAGCCC
vrrS1	TCGTCCTGGAGCATCTTTCA	CCAAATCGCCCCTAGACCAA
vrrS2	GTTGTTTCATACGTCTATCCCCTTC	GTCCTTTTGGACAGCCTCTCTT
vrrS3	ACTGTAGTTGTCCCTACCCTT	AGAAGTACAGGTGGGACAGGA
vrrS4	TTTCCTTGCGATGCTTCAGT	TGCTGGTATAGAGCCATCTGC
vrrS5	AGCAATGTTTAATTCACCATCAAGT	GTACGCTTTAGTCGGAGACGG
vrrS6	AGGAAGCAGGTTAGCGTTGT	GCGCTATGTGGCGTCTTTTC
vrrS7	AGGAACACTGGTTCAGCCTAT	AGCAGGATCGCTTGCTAGAT
vrrS8	CTGCAATTGCCTTCGCCTTT	GCGAAAAAGAGAAAGCGCTAC
vrrS9	ATGAAGGTGTGACATGCCGT	GTGAAGCTGTAATTGTGGCGT

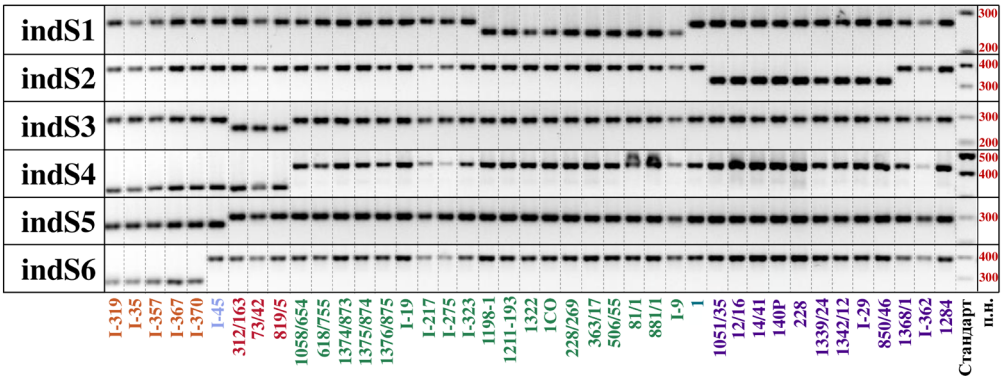

Fig. 2. Results of PCR reaction of INDEL loci with detection by electrophoresis.
The font colors of the strains correspond to the canonical lineages according to Fig. 1.

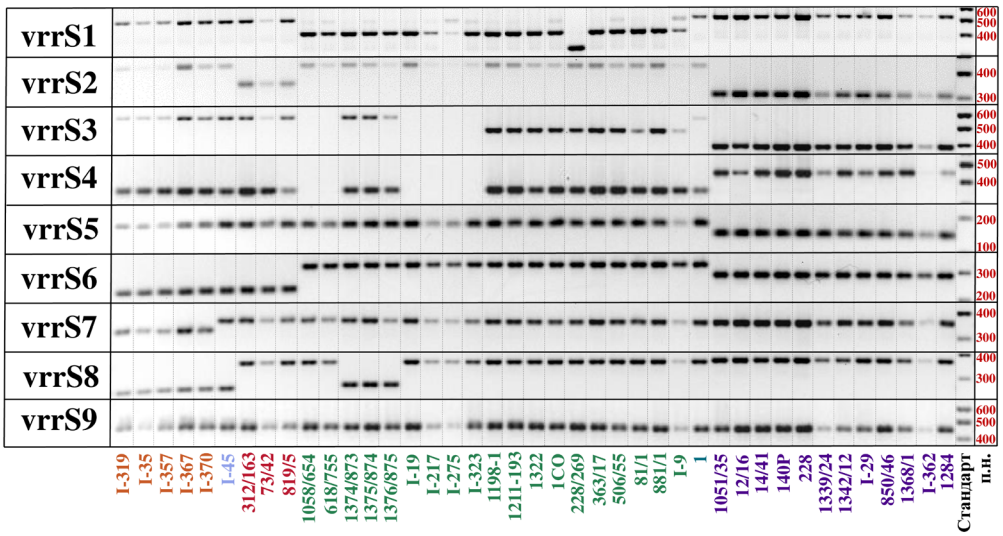

Fig. 3. Results of PCR reaction of VNTR loci with detection by electrophoresis.
The font colors of the strains correspond to the canonical lineages according to Fig. 1.

Table 3. Lengths of INDEL and VNTR loci according to phylogenetic groups

Markers	Phylogenetic groups									
	A.Br. Ames	A.Br. 001/002	A.Br. Aust94	A.Br. 005/006	A.Br. 008/00 (STI)	A.Br. 008/00 (A.Br.125)	A.Br. 008/00 Tsiankovskii	A.Br. 008/009 (228-269)	B.Br. 001/002	B.Br. 001/002 (B.Br.014)
indS1	265	265	265	265	265	241	241	265	265	265
indS2	375	375	375	375	375	375	375	375	375	312
indS3	286	286	253	286	286	286	286	286	286	286
indS4	328	328	328	424	424	424	424	424	424	424
indS5	255	255	285	285	285	285	285	285	285	285
indS6	277	388	388	388	388	388	388	388	388	388
vrrS1	513	513	513	513	425	425	425	337	513	513
vrrS2	422	422	340	422	422	422	422	422	299	299
vrrS3	545	545	545	545	545	545	464	464	383	383
vrrS4	354	354	354	354	354	354	354	354	444	444
vrrS5	172	172	172	172	172	172	172	172	142	142
vrrS6	238	238	238	318	318	318	318	318	278	278
vrrS7	307	346	346	346	346	346	346	346	346	346
vrrS8	258	258	360	360	360	264	360	360	360	360
vrrS9	450	450	450	450	450	450	450	450	450	450

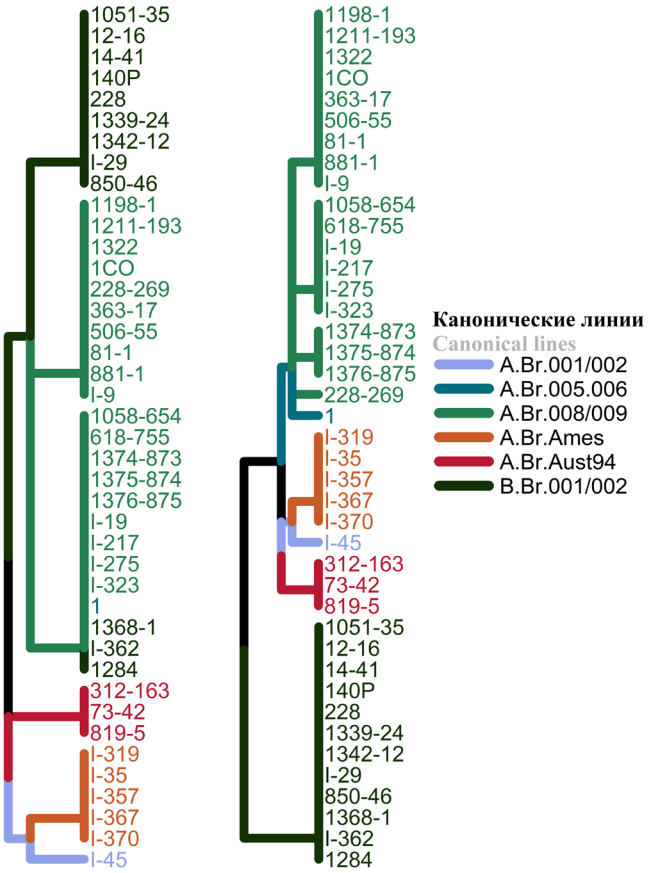

Fig. 4. Clustering of *B. anthracis* strains based on VNTR and INDEL typing.

tification by PCR with electrophoretic visualization of the results allows reliable determination of allelic variants of markers. The found 9 VNTR markers and 6 INDEL markers allow to divide *B. anthracis* strains into 6 and 9 genetic groups when typing with separate analysis of these markers and into 10 groups when analyzing them together. The genotyping technique based on the analysis of new VNTR- and INDEL-markers is recommended to be used jointly or separately as an addition to the existing genotyping schemes. The use of the developed methodology for the identification of variable VNTR- and INDEL-loci allows reliable determination of the phylogenetic position of *B. anthracis* strains and is promising for use in the epidemiologic investigation of anthrax outbreaks.

СПИСОК ИСТОЧНИКОВ | REFERENCES

1. Harrell L.J., Andersen G.L., Wilson K.H. Genetic variability of *Bacillus anthracis* and related species. *J. Clin. Microbiol.* 1995;33(7):1847–50. DOI: https://doi.org/1128/jcm.33.7.1847-1850.1995
2. Andersen G.L., Simchock J.M., Wilson K.H. Identification of a region of genetic variability among *Bacillus anthracis* strains and related species. *J. Bacteriol.* 1996;178(2):377–84. DOI: https://doi.org/1128/jb.178.2.377-384.1996
3. Jackson P.J., Walthers E.A., Kalif A.S., et al. Characterization of the variable-number tandem repeats in vrrA from different *Bacillus anthracis* isolates. *Appl. Environ. Microbiol.* 1997;63(4):1400–5. DOI: https://doi.org/1128/aem.63.4.1400-1405.1997
4. Pearson T., Busch J.D., Ravel J., et al. Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. *Proc. Natl Acad. Sci.* 2004; 101(37):13536–41. DOI: https://doi.org/1073/pnas.0403844101
5. Keim P., Price L.B., Klevytska A.M., et al. Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J. Bacteriol.* 2000;182(10):2928–36. DOI: https://doi.org/1128/jb.182.10.2928-2936.2000
6. Le Flèche P., Hauck Y., Onteniente L., et al. A tandem repeats database for bacterial genomes: application to the genotyping of *Yersinia pestis* and *Bacillus anthracis*. *BMC Microbiol.* 2001;1:2. DOI: https://doi.org/1186/1471-2180-1-2
7. Lista F., Faggioni G., Valjevac S., et al. Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. *BMC Microbiol.* 2006;6:33. DOI: https://doi.org/1186/1471-2180-6-33
8. Van Ert M.N., Easterday W.R., Huynh L.Y., et al. Global genetic population structure of *Bacillus anthracis*. *PLoS One.* 2007; 2(5):e461. DOI: https://doi.org/1371/journal.pone.0000461
9. Beyer W., Bellan S., Eberle G., et al. Distribution and molecular evolution of *Bacillus anthracis* genotypes in Namibia. *PLoS Negl. Trop. Dis.* 2012;6(3):e1534. DOI: https://doi.org/1371/journal.pntd.0001534
10. Thierry S., Tourterel C., Le Flèche P., 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. DOI: https://doi.org/1371/journal.pone.0095131
11. Sahl J.W., Pearson T., Okinaka R., et al. A *Bacillus anthracis* genome sequence from the Sverdlovsk 1979 autopsy specimens. *mBio*. 2016;7(5):e01501–16. DOI: https://doi.org/1128/mBio.01501-16
12. Eremenko E., Pechkovskii G., Pisarenko S., et al. Phylogenetics of *Bacillus anthracis* isolates from Russia and bordering countries. *Infect. Genet. Evol.* 2021;92:104890. DOI: https://doi.org/1016/j.meegid.2021.104890
13. Larsson P., Svensson K., Karlsson L., et al. Canonical insertion-deletion markers for rapid DNA typing of *Francisella tularensis*. *Emerg. Infect. Dis.* 2007;13(11):1725–32. DOI: https://doi.org/3201/eid1311.070603
14. Сорокин В.М., Водопьянов А.С., Писанов Р.В. INDEL-типирование — новый метод дифференциации штаммов *Helicobacter pylori*. *Бактериология*. 2020;5(1):8–13. Sorokin V.M., Vodop'yanov A.S., Pisanov R.V. INDEL-typing: a new method of differentiation of *Helicobacter pylori* strains. *Bacteriology*. 2020;5(1):8–13. DOI: https://doi.org/20953/2500-1027-2020-1-8-13
15. Леденева М.Л., Водопьянов А.С., Ткаченко Г.А. и др. Выявление INDEL-маркеров в геномах штаммов *Burkholderia pseudomallei* для внутривидового генотипирования. *Журнал микробиологии, эпидемиологии и иммунобиологии.* 2017;(4): 35–41. Ledeneva M.L., Vodop'yanov A.S., Tkachenko G.A., et al. Detection of INDEL-markers in genomes of *Burkholderia pseudomallei* strains for intra-species genotyping. *Journal of Microbiology, Epidemiology and Immunobiology.* 2017;(4): 35–41. DOI: https://doi.org/36233/0372-9311-2017-4-35-41 EDN: https://elibrary.ru/bucgfm
16. Водопьянов А.С., Водопьянов С.О., Олейников И.П., Мишанькин Б.Н. INDEL-типирование штаммов *Vibrio cholerae*. *Эпидемиология и инфекционные болезни*. 2017;22(4):195–200. Vodop'yanov A.S., Vodop'yanov S.O., Oleinikov I.P., Mishan'kin B.N. Indel-genotyping of *Vibrio cholerae* strains. *Epidemiology and Infectious Diseases.* 2017;22(4):195–200. DOI: https://doi.org/18821/1560-9529-2017-22-4-195-200 EDN: https://elibrary.ru/zhlhhh
17. Трухачев А.Л., Мелоян М.Г., Воскресенская Е.А. и др. INDEL-типирование штаммов *Yersinia pseudotuberculosis*.

Проблемы особо опасных инфекций. 2022;(4):102–9. Trukhachev A.L., Meloyan M.G., Voskresenskaya E.A., et al. INDEL-typing of *Yersinia pseudotuberculosis* strains. *Problems of Particularly Dangerous Infections.*
DOI: https://doi.org/21055/0370-1069-2022-4-102-109
EDN: https://elibrary.ru/rucepz

18. Tamura K., Stecher G., Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 2021; 38(7):3022–7. DOI: https://doi.org/1093/molbev/msab120
19. Еременко Е.И., Рязанова А.Г., Писаренко С.В. и др. Новые генетические маркеры для молекулярного типирования штаммов *Bacillus anthracis*. *Проблемы особо опасных инфекций.* 2019;(3):43–50. Eremenko E.I., Ryazanova A.G., Pisarenko S.V., et al. New genetic markers for molecular typing of *Bacillus anthracis* strains. *Problems of Particularly Dangerous Infections.* 2019;(3):43–50.
DOI: https://doi.org/21055/0370-1069-2019-3-43-50
EDN: https://elibrary.ru/pgefkd
20. Hunter P.R., Gaston M.A. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J. Clin. Microbiol.* 1988;26(11):2465–6.
DOI: https://doi.org/10.1128/jcm.26.11.2465-2466.1988

Information about the authors

Grigorii A. Pechkovskii✉ — junior researcher, Anthrax laboratory, Stavropol Plague Control Research Institute, Stavropol, Russia, grigorii.pechkovskii@gmail.com, https://orcid.org/0000-0001-7033-9972

Evgeny I. Eremenko — D. Sci. (Med.), Professor, principal researcher, Anthrax laboratory, Stavropol Plague Control Research Institute, Stavropol, Russia, https://orcid.org/0000-0002-8163-1300

Alla G. Ryazanova — Cand. Sci. (Med.), Head, Anthrax laboratory, Stavropol Plague Control Research Institute, Stavropol, Russia, https://orcid.org/0000-0002-5196-784X

Sergey V. Pisarenko — Cand. Sci. (Chem.), leading researcher, Laboratory of biochemistry, Stavropol Plague Control Research Institute, Stavropol, Russia, https://orcid.org/0000-0001-6458-6790

Nikolay A. Shapakov — junior researcher, Laboratory of biochemistry, Stavropol Plague Control Research Institute, Stavropol, Russia, https://orcid.org/0000-0001-9152-4026

Lyudmila Yu. Aksenova — Cand. Sci. (Med.), senior researcher, Anthrax laboratory, Stavropol Plague Control Research Institute, Stavropol, Russia, https://orcid.org/0000-0002-7744-3112

Olga V. Semenova — Cand. Sci. (Biol.), researcher, Anthrax laboratory, Stavropol Plague Control Research Institute, Stavropol, Russia, https://orcid.org/0000-0003-0274-898X

Lyudmila D. Timchenko — D. Sci. (Vet.), Professor, principal researcher, Interdepartmental scientific and educational laboratory of experimental immunomorphology, immunopathology and immunobiotechnology, North-Caucasus Federal University, Stavropol, Russia, https://orcid.org/0000-0003-2011-880X

Alexander N. Kulichenko — D. Sci. (Med.), Professor, Academician of RAS, Director, Stavropol Research Anti-Plague Institute, Stavropol, Russia, https://orcid.org/0000-0002-9362-3949

Author contribution. *Pechkovskii G.A.* — research concept, laboratory research, bioinformatic analysis, writing an article; *Eremenko E.I.* — genomic research, phylogenetic analysis, writing an article; *Ryazanova A.G.* — phylogenetic analysis, writing an article; *Pisarenko S.V.* — genomic research, editing an article; *Shapakov N.A.* — genomic research, synthesis of primers; *Aksenova L.Yu., Semenova O.V.* — laboratory research; *Timchenko L.D., Kulichenko A.N.* — research concept, article editing. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published.

The article was submitted 15.02.2024;
accepted for publication 25.04.2024;
published 29.06.2024

Информация об авторах

Печковский Григорий Александрович✉ — м.н.с., лаб. сибирской язвы Ставропольского противочумного института, Ставрополь, Россия, grigorii.pechkovskii@gmail.com, https://orcid.org/0000-0001-7033-9972

Еременко Евгений Иванович — д.м.н., проф., г.н.с., лаб. сибирской язвы Ставропольского противочумного института, Ставрополь, Россия, https://orcid.org/0000-0002-8163-1300

Рязанова Алла Геннадьевна — к.м.н., зав. лаб. сибирской язвы Ставропольского противочумного института, Ставрополь, Россия, https://orcid.org/0000-0002-5196-784X

Писаренко Сергей Владимирович — к.х.н., в.н.с., лаб. биохимии Ставропольского противочумного института, Ставрополь, Россия, https://orcid.org/0000-0001-6458-6790

Шапаков Николай Андреевич — м.н.с., лаб. биохимии Ставропольского противочумного института, Ставрополь, Россия, https://orcid.org/0000-0001-9152-4026

Аксенова Людмила Юрьевна — к.м.н., с.н.с., лаб. сибирской язвы Ставропольского противочумного института, Ставрополь, Россия, https://orcid.org/0000-0002-7744-3112

Семенова Ольга Викторовна — к.б.н., н.с., лаб. сибирской язвы Ставропольского противочумного института, Ставрополь, Россия, https://orcid.org/0000-0003-0274-898X

Тимченко Людмила Дмитриевна — д.вет.н., проф., г.н.с. межкафедральной научно-образовательной лаборатории экспериментальной иммуноморфологии, иммунопатологии и иммунобиотехнологии Северо-Кавказского федерального университета, Ставрополь, Россия, https://orcid.org/0000-0003-2011-880X

Куличенко Александр Николаевич — д.м.н., проф., академик РАН, директор Ставропольского противочумного института, Ставрополь, Россия, https://orcid.org/0000-0002-9362-3949

Участие авторов. *Печковский Г.А.* — концепция исследования, лабораторные исследования, биоинформационный анализ, написание статьи; *Еременко Е.И.* — геномные исследования, филогенетический анализ, написание статьи; *Рязанова А.Г.* — филогенетический анализ, написание статьи; *Писаренко С.В.* — геномные исследования, редактирование статьи; *Шапаков Н.А.* — геномные исследования, синтез праймеров; *Аксенова Л.Ю., Семенова О.В.* — лабораторные исследования; *Тимченко Л.Д., Куличенко А.Н.* — концепция исследования, редактирование статьи. Все авторы внесли существенный вклад в проведение поисково-аналитической работы и подготовку статьи, прочли и одобрили финальную версию до публикации.

Статья поступила в редакцию 15.02.2024;
принята к публикации 25.04.2024;
опубликована 29.06.2024