Variability of the Bacillus anthracis tryptophan operon

Eugene Ivan Eremenko (ejer@mail.ru)
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut https://orcid.org/0000-0002-1117-1185
Olga Bobrysheva
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut
Sergey Pisarenko
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut
Alla Ryazanova
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut
Olga Semenova
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut
Lyudmila Aksenova
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut
Dmitry Kovalev
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut
Grigorii Pechkovskii
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut
Anna Evchenko
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut
Alexander Kulichenko
Stavropol'skij naucno-issledovatel'skij protivocumnyj institut

Research article

Keywords: Bacillus anthracis, tryptophan operon, single nucleotide substitution, indel, main lineages, tryptophan nutritional requirements, evolution

Posted Date: April 21st, 2020

DOI: https://doi.org/10.21203/rs.3.rs-21866/v1

Abstract

Background

Bacillus anthracis is a causal agent of a zoonotic disease relevant for many countries, and is an agent of bioterrorism. Meanwhile, the reasons for the dependence on tryptophan of some strains with altered virulence have not been established with an almost complete absence of information on the tryptophan operon of this pathogen. In this study, we report gene variability and the structure of the tryptophan operon in *B. anthracis* strains of the three main lineages.

Results

For *in silico* analysis we used 112 *B. anthracis* genomes, including 68 of those available at the GenBank database and 44 sequenced at our institute. The *B. anthracis* tryptophan operon has an ancestral structure with a complete set of seven partially overlapping genes. The results show that the variability of all seven tryptophan operon genes is determined by the presence of single nucleotide polymorphisms and InDels. The *trpA* genes of strains of the main lineage B and *trpG* genes of strains of the C lineage are pseudogenes and the proteomes lack the corresponding enzymes of the biosynthetic pathway, which may explain the dependence of the strains of line B on tryptophan.

Conclusion

In this study, the differences in tryptophan operon genes for *B. anthracis* strains belonging to different main lineages were demonstrated for the first time. Mutation in the gene of the tryptophan synthase subunit alpha can explain the dysfunction of this enzyme and the dependence on tryptophan in strains of the main lineage B. Identified features suggest a further study of the dependence on tryptophan in *B. anthracis* strains of the main lineage B and may be of interest from the point of view of intraspecific evolution of the anthrax pathogen.

Background

The causal agent of anthrax - *Bacillus anthracis* - causes a particularly dangerous zoonotic infection with a global range and is an agent of biological terrorism of group A [1]. The ability of the spore form of this bacterium to persist in soil foci for decades and cause poorly predictable disease outbreaks among livestock, often accompanied by human infections, makes anthrax a problem for public health and veterinary medicine in many countries, including Russia [2]. The research of anthrax infection and its causal agent has been the subject of numerous works by researchers all over the world, but despite the long history of research of *B. anthracis*, some of the properties of this pathogen remain poorly understood. Among them is the dependence on tryptophan in a number of strains whose virulence is reduced [3].

The tryptophan biosynthesis pathway is one of the branches of the general branched aromatic amino acid biosynthesis pathway which starts with chorismic acid. Tryptophan operon (Trp) is responsible for tryptophan biosynthesis. The genes and operons of the tryptophan biosynthetic pathway are organized differently in different types of bacteria. These differences reflect evolutionary divergence, as well as adaptation to unique metabolic capabilities and interactions with the environment [4].

Jacques Monod described tryptophan operon of *Escherichia coli* for the first time in 1953. The *Bacillus anthracis* tryptophan operon contains genes for seven catalytic domains encoding five enzymes, including two α/β subunit complexes - tryptophan synthase and anthranilate synthase:

1. *trpA* encoding tryptophan synthase subunit alpha;
2. *trpB* coding tryptophan synthase subunit beta;
3. *trpD* encoding anthranilate phosphoribosyl transferase;
4. *trpE* encoding anthranilate synthase component I;
5. BA_1249 (*trpG*), encoding aminodeoxychorismate/anthranilate synthase component II TrpG),
6. *trpF* encoding N- (5'phosphoribosyl) anthranilate isomerase;
7. BA_1251 (*trpC*) encoding indol-3-glycerol phosphate synthase TrpC).

This is an ancestral structure of the operon, which includes a full set of specific whole-pathway operons that is widespread among prokaryotes. For some organisms, genes of biosynthetic pathways may be scattered, for others – organized in two or more “split-pathway” operons. The question is what kind of evolutionary relationships exists between these three types of pathway genes organization. Trp operon is a perfect model for studying the biosynthetic pathways [5].

Mechanism of tryptophan dependence is not quite clear, and there is very little published information on *trp* operon of *B. anthracis*. One of the possible reasons of the *trp* dependence may be the mutations in genes determining enzymes of *trp* synthesis pathway. There is an important link between the organization and genomic context of the *trp* operon genes and the mechanism that regulates its expression. The regulatory mechanisms used to control the transcription of tryptophan biosynthesis genes in *B. anthracis* are still poorly understood. It is a known fact that unlike *Bacillus subtilis, B. anthracis* lacks trpRNA binding attenuator protein (TRAP), encoded by *mtrB* gene.

Due to the low state of knowledge of the *trp* operon and the development mechanism of tryptophan auxotrophy in *B. anthracis* determines the relevance of this study. Our aim was to analyze the features of genes and structure of the *trp* operon of different *B. anthracis* strains.

Results

The comparison of nucleotide sequences of *trp* operon genes showed the following: *trpA* gene size of 33 strains of the main genetic line B is 651 b.p., for I-373 strain – 650 b.p., for Tyrol 4675 strain – 777 b.p. For 77 strains of the main lineage A gene size estimates 777 b.p., for Cvac02 strain – 834 b.p. We have

detected substitution of A490C in I-9 strain; G652T and G655T are substituted in Tyrol 4657 strain; in strain I-373 there are three substitutions of A21G, G22C, C23A, and deletion 35A; in Cvac02 strain – insertions A765^766 and of 56 b.p. 779^834.

The majority of *B. anthracis* strains, as well as *Bacillus cereus*, *Bacillus thuringiensis* and other bacilli have identical *trpA* proteins (MULTISPECIES: tryptophan synthase subunit alpha) with the size of 258 amino acid residues (a.a.). Unique *trpA* (tryptophan synthase subunit alpha) protein of the Cvac02 strain has a size of 277 a.a. because of 19 additional amino acids VSLFFLFCVINVKIYRKYI at the C-end of the polypeptide and a A258G substitution. In the strains of main lineage B from the State Collection of Pathogenic Microorganisms of Stavropol Research Anti-Plague Institute *trpA* is a pseudogene and does not encode any proteins.

Alignment of the *trpA* genes and phylogenetic analysis have shown that strains are grouped in two clades and two separate branches inside each (Fig.1). 34 strains that belong to the main lineage B of the global phylogenetic *B. anthracis* structure belong to one branch of the first clade. Tyrol 4675 that has the same gene size as strains of the line A, forms a separate branch of that clade. The other 78 strains, two of which (2000031052 and 2002013094) belong to the main lineage C, 75 – to line A, and one (I-9) forms a separate branch, compile the second clade.

Thus, observed discrepancies of the *trpA* gene sequences differentiate the strains of main genetic lineage B from the strains of lineages A and C. The presence of the defective *trpA* gene possibly explains the tryptophan dependence identified for 12-16, 12-16-1, 14-41 and 140P strains from the State Collection of Pathogenic Microorganisms of Stavropol Research Anti-Plague Institute. It is noteworthy that these strains are atypical in a number of features. Therefore, in order to verify the link between the presence of the defective *trpA* gene and tryptophan dependence, and to find out whether this mutation is enough for manifesting the feature, one should conduct a research on tryptophan dependence of other, typical *B. anthracis* strains that belong to the main lineage B.

The size of *trpB* gene for all strains is 1149 b.p., except for strain 1(CO) with insertion A1174 and the size of 1150 b.p. In the genome of the 2000031052 we have detected substitution A1017G, which does not lead to the substitution in the protein [MULTISPECIES: tryptophan synthase subunit beta [Bacillus].

The size of BA_1251(*trpC*) for all strains is 762 b.p. except for 312-163 strain with deletion AA351-352 and the size of 760 b.p. There is a substitution C762A which does not lead to substitution in the protein [MULTISPECIES: indole-3-glycerol phosphate synthase TrpC [Bacillus cereus group] for strains I-19, I-271, I-275, I-323, Ba888494-Geo, 55-VNIIViM and STI-1.

The size of *trpD* gene for all the strains is 1026 b.p. We have detected a substitution A899G for 35 strains of the main lineage B. TrpD protein with the size of 341 amino acid residues (MULTISPECIES: anthranilate phosphoribosyl transferase [Bacillus cereus group]) is found in genomes of all representatives of *Bacillus cereus* group, including *B. anthracis* strains of the main lineages A and C. TrpD protein of *B. anthracis*

strains of the main lineage B is different from strains of lineage A and C with the N300S substitution (asparagine→serine) and is characteristic only for the strains of this line.

The dendrogram (Fig.2) shows a basal division into two clades, in one of which the same 35 strains of the main lineage B are grouped as in the case of *trpA*, the second is represented by strains of lineages A and C. Thus, differences in the sequences of the *trpD* gene also differentiate strains of lineage B from strains of lineages A and C.

trpE gene size is 1419 b.p. for the majority of strains. There is an insertion of 9 b.p. GTGAAAGG at 5'-end and deletion 676-679CCG in this gene of the Han strain; its size is 1425 b.p. Besides this strain has a substitution G676C which leads to insertion 2-4KGM (lysine-glycine-methionine), substitution A225R (alanine→arginine) and G226 deletion in the protein [anthranilate synthase component I [Bacillus anthracis]. Other strains have the [MULTISPECIES: anthranilate synthase component I [Bacillus cereus group] protein. Thus, there are two types of *trpE* and correspondingly two TrpE proteins.

The size of *trpF* gene is 615 b.p. There is a G3A substitution in the strains I-19, I-217, I-275, I-323, Ba888494-Geo, Ba767389-Geo, 55-VNIIViM and STI-1. This substitution does not lead to the substitution in the protein [MULTISPECIES: N-(5'phosphoribosyl) anthranilate isomerase [Bacillus]. For strains A1144, Carbosap and ANSES_86 we have detected a C259T substitution which leads to the H87Y substitution (histidine→tyrosine) in the [N-(5'phosphoribosyl) anthranilate isomerase [Bacillus anthracis] protein. There are two types of *trpF* genes and corresponding TrpF proteins.

The size of the BA_1249 (*trpG*) gene for the strains of the lineages A and B is 588 b.p., and 593 b.p. for the lineage C. Gene product for the strains of the lineage A is [aminodeoxychorismate/anthranilate synthase component II] protein, which is characteristic for these strains only. C320T substitution in the strains of lineage B leads to the significant substitution A107V (alanine→valine) in the [(MULTISPECIES: aminodeoxychorismate/anthranilate synthase component II)] protein, presented in the proteome of other representatives of *Bacillus cereus* group. In the strains of the lineage C *trpG* gene is a pseudogene due to the frameshift mutation (insertion 231^232 TAGGT) and C320T substitution.

In addition to the *trpG* gene encoding the aminodeoxychorismate/anthranilate synthase component II in the tryptophan operon, the *B. anthracis* genome of all three main lineage contains the *pabA* gene encoding (MULTISPECIES: aminodeoxychorismate / anthranilate synthase component II). This protein is also synthesized by many strains of *Bacillus cereus*, *Bacillus thuringiensis* and other bacilli. The *trpG* and *pabA* genes are distinguished by multiple substitutions, InDels, as well as the proteins encoded by them.

Discussion

Reconstruction of the structure of the tryptophan operon of different *B. anthracis* strains showed that there are differences between its structure in strains of the main lineages (Fig 3). In strains of lineage B, due to a mutation in the *trpA* gene, which turns it into a pseudogene, the last step of the tryptophan

biosynthesis pathway should be blocked, since the tryptophan synthase subunit alpha is absent. This circumstance may explain the dependence on tryptophan in strains of the main lineage B.

At the second stage of tryptophan synthesis in the strains of lineage C, the existing frameshift mutation in the *trpG* gene of anthranilate synthase component II turns it into a pseudogene, which theoretically should turn off the entire subsequent pathway of tryptophan biosynthesis. However, in addition to the *trpG* gene in the tryptophan operon encoding the aminodeoxychorismate/anthranilate synthase component II, the *pabA* gene encoding the protein (MULTISPECIES: aminodeoxychorismate/anthranilate synthase component II) is represented in the *B. anthracis* genome of all three main lineages.

This protein is also synthesized by many strains of *Bacillus cereus, Bacillus thuringiensis* and other bacilli. The *trpG* and *pabA* genes are distinguished by multiple substitutions, InDels, as well as the proteins encoded by them. The presence of the *pabA* gene outside the tryptophan operon because of its glutamine amido transferase activity may possibly complement the no functionality of the *trpG* pseudogene.

Partial overlapping of the sequences of most *trp* operon genes is noteworthy. Mutations in the gene sequences of the tryptophan synthase subunit alpha of strains *B. anthracis* of the main lineage B and genes of the anthranilate synthase component II of strains of the main lineage C, converting them into pseudogenes, distinguishes these strains from representatives of the lineage A, which have a full set of functional genes of the tryptophan operon. These features should be correlated with data on the intraspecific evolution of *B. anthracis*. Studies indicate that lineage C separated from the line containing clades A and B according to various sources 9,600 – 26,000 years ago, and the next divergence with the separation of line A from line B occurred 6,500 – 17,000 years ago. Further, lineage A and B evolved in parallel [6; 7; 8; 9; 10; 11]. The high-resolution reference phylogeny, based on 11989 SNPs of genomes of 193 strains from the global collection, reveals that the next event after the separation of lineage C from A/B was the divergence of lineage C into sub clusters, then the separation of lines A and B [12].

Clade A divides into four main monophyletic subclades, from which, earlier than other subclades, formed the “Ancient A” clade, being the base for other subclades of this line. The base subclade of clade B may be subclade B.Br003, including subclade B.Br004 with strains from Europe, formed at about the same time as subclade A.Br.002, other subclades of line B include strain HYU01 from South Korea (subclade B.Br. 002), which appeared later, and finally, the strains of the clade B.Br.008 isolated in South Africa and Sweden [12].

According to our data, the subclade B.Br.002 contains, along with the isolate from Korea, strains isolated in Western Siberia (a separate cluster “Siberia”) and Finland [13], although in an earlier work the strain from Finland was described as constituting a separate clade line B.Br.002 with the nearest strains HYU01 from South Korea and BF1 from Germany [14].

Based on the data described, it can be assumed that clade C, the oldest, with a minimum number of isolates, has become a blind branch of the evolution, which has not received further distribution outside

the United States. Clades A and B, evolving independently, spread to varying degrees in different geographical areas, while there are local regions where strains of clades A and B exist at the same time, for example, Kruger Park in South Africa, and probably certain regions of the Russian Federation (Republic of Dagestan, Western Siberia). The fact that only 5 isolates of *B. anthracis* line C was isolated in North America only, a limited number of strains of line B and the wide distribution of strains of line A, suggests the ecological advantages of the latter, which are also associated with different functioning of tryptophan, and possibly other operons.

In the strains of *Bacillus cereus*, the structure of the tryptophan operon is not different from that in *B. anthracis*, but the genes and corresponding proteins are mainly specific for this species, although some proteins are identical in these two species and *Bacillus thuringiensis*.

Conclusion

General structure of the *B. anthracis trp* operon is conservative and is characterized by the presence of 7 partially overlapping genes. We have shown the difference in gene sequences and proteins of the biosynthetic pathway of the main lineages of the anthrax pathogen. In accordance with the nature of single nucleotide polymorphisms and InDels in the *trpA* and *trpD* genes, the studied strains are divided into two groups, one of which includes strains of the main lineages A and C, and the other - strains of lineage B. Due to a mutation in the *trpA* gene of the tryptophan synthase subunit alpha, which turns it into a pseudogene, the last step of the tryptophan biosynthesis pathway should be blocked, which may explain the dependence on tryptophan found in several *B. anthracis* strains of the main lineage B. It remains unknown whether tryptophan dependence is inherent in all strains of this line. The presence of the *trpG* pseudogene in strains of the main lineage C and the inability to synthesize of anthranilate synthase component II can be probably compensated for by expressing glutamine amido transferase activity of the functional *pabA* gene outside the tryptophan operon.

Worldwide distribution of line A strains suggest their ecological advantages, which can be associated in particular with full functioning of tryptophan operon.

The revealed features suggest a further study of tryptophan dependence in *B. anthracis* strains of the main lineage B and may be of interest from the point of view of intraspecific evolution of the anthrax causal agent.

Methods

Bacterial strains

In our study we have used 112 genomes of the *B. anthracis* strains. 44 strains of them, sequenced in our study, are from the State Collection of Pathogenic Microorganisms of Stavropol Research Anti-Plague Institute (Table 1) and 68 genomic sequences of the *B. anthracis* strains - from GenBank (Additional file 1: Table S1).

Growth of B. anthracis and extraction of DNA

B. anthracis strains were cultivated on the blood agar, then inactivated, and DNA was extracted with the use of DNA extraction kit QIAamp DNA Mini Kit (Qiagen, Germany) according to manufacturer's protocol and the requirements of biological safety rules when working with pathogens of the third group of pathogenicity. DNA concentration was quantified using the dsDNA HS Qubit assay kit (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. DNA preparations were stored at − 20 °C until further use.

Whole genome sequencing

The preparation of genomic libraries with a 400 bp read length was performed using the Ion Xpress Plus Fragment Library Kit reagent kit (Life Technologies, USA) in accordance with the manufacturer's protocol. Monoclonal amplification on microspheres was performed using Ion PGM Hi-Q View OT2 Kit reagents (Life Technologies, USA). Genome sequencing was performed using an Ion Torrent PGM sequencer and Ion 316 Chips Kit V2 chips (Life Technologies, USA).

Bioinformatics analysis

We conducted mutation search in genomes *in silico* via CLC Sequence Viewer 6 [15] and MEGA V.10.0.5 [16] programs, using genome of the *Bacillus anthracis* Ames Ancestor strain as reference and data on genes and enzymes of the *trp* operon from GenBank. Phylogenetic analysis was performed via Maximum Likelihood (bootstrap 1000) method in MEGA V.10.0.5. Visualization of the phylogenetic tree was performed in the Figtree v1.4.4 [17].

Abbreviations

a.a. – amino acid

b.p. – base pair

InDels – Insertion/Deletion

trp – tryptophan

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Generated genomic sequences of this study are available from the Russian Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing (Rospotrebnadzor) but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of the Russian Federal Service for Surveillance on Consumer Rights Protection and Human Wellbeing (Rospotrebnadzor). Other sequences used for this study, including sequence of 81-1 strain from our strains' depository, are publicly available in the National Center for Biotechnology Information in the repository under the following accession numbers:

GCA_000008445.1, GCA_000742895.1, GCA_000832965.1, GCA_000742655.1

GCA_000832565.1, GCA_000833125.1, GCA_000512835.2, GCA_000833275.1

GCA_000832465.1, GCA_003227955.1, GCA_001936375.1, GCA_000833065.1

GCA_001654475.1, GCA_001543225.1, GCA_000832505.1, GCA_000534935.2

GCA_000258885.1, GCA_000832665.1, GCA_000295695.2, GCA_000725325.1

GCA_006088855.1, GCA_000832745.1, GCA_000583105.1, GCA_000832725.1

GCA_006742565.1, GCA_000875715.1, GCA_000832585.1, GCA_000742695.1

GCA_001277955.1, GCA_000831505.1, GCA_000742315.1, GCA_000167275.1

GCA_000743805.1, GCA_000310045.1, GCA_000986915.1, GCA_001273005.1

GCA_002896655.1, GCA_003063925.1, GCA_000740925.2, GCA_000697555.2

GCA_002025335.1, GCA_003064045.1, GCA_000359465.1, GCA_000359425.1

GCA_002233635.1, GCA_000167315.1, GCA_001835485.1, GCA_001273085.1

GCA_002025455.1, GCA_002019425.1, GCA_000167295.1, GCA_000167235.1

GCA_000167255.1, GCA_900014355.1, GCA_001677295.1, GCA_008087155.1

GCA_001273145.1, GCA_000167335.1, GCA_000783215.1, GCA_000181675.2

GCA_000278385.1, GCA_000181935.1, GCA_000782995.1, GCA_000181995.1

GCA_009669005.1, GCA_000782885.1, GCA_002007035.1, GCA_009872685.1

GCA_003860145.1

Competing interests

The authors declare that they have no competing interests

Funding

Not applicable

Authors' contributions

EIE, SVP and ANK have developed a project and a research plan. EIE, SVP, DAK, and ANK compiled the manuscript. AGR, OVS, LYA, conducted bacteriological studies. SVP and OVB performed sequencing, genomes assembly and annotation. EIE, SVP, GAP and AYE conducted phylogenetic analysis.

All authors have read and approved the final manuscript.

Acknowledgements

Not applicable

References

1. Rotz L.D., Khan A.S., Lillibridge S.R., Ostroff S.M., Hughes J.M. Public health assessment of potential biological terrorism agents. Emerg Infect Dis. 2002;8(2):225–30.
2. Carlson C.J., Kracalik I.T., Ross N., Alexander K.A., Hugh-Jones M.E., Fegan M., et al. The global distribution of Bacillus anthracis and associated anthrax risk to humans, livestock and wildlife. Nat Microbiol. 2019;4(8):1337–43. doi: 10.1038/s41564-019-0435-4.
3. Eremenko E. I. Ryazanova A. G., Tsygankova O. I., Tsygankova E. A., Buravtseva N. P., Kulitchenko A. N. Genotype Diversity of Bacillus anthracis Strains Isolated from the Caucasus Region. Mol Genet Microbiol. 2012; 27(2):74–8.
4. Merino E., Jensen R.A, Yanofsky Cs. Evolution of bacterial trp operons and their regulation. Current Opinion in Microbiology. 2008;11:78–86. DOI 10.1016/j.mib.2008.02.005
5. Xie G., Keyhani N.O., Bonner C.A., Jensen R.A. Ancient Origin of the Tryptophan Operon and the Dynamics of Evolutionary Change. Microbiol Mol Biol R. 2003;67(3):303–42.
6. Pearson T., Busch J.D., Ravel J., Read T.D., Rhoton S.D., U'Ren J.M., et al. Phylogenetic discovery bias in Bacillus anthracis using single-nucleotide polymorphisms from whole-genome sequencing. 2004;101:37.
7. Van Ert M.N., Easterday W.R., Huynh L.Y, Okinaka R.T., Hugh-Jones M.E., Ravel J., et al. Global Genetic Population Structure of Bacillus anthracis. PLoS ONE. 2007;2(5): e461.
8. Pilo P., Frey J. Bacillus anthracis: Molecular taxonomy, population genetics, phylogeny and patho-evolution. Infect Genet and Evol. 2011;11(6):1218–24.

9. Pilo P., Frey J. Pathogenicity, population genetics and dissemination of Bacillus anthracis. Infect Genet and Evol. 2018; 64:115–125.
10. Eremenko E.I., Pisareenko S.V., Aksenova L.Yu., Ryazanova A.G., Zhirov A.M., Semenova O.V., Bobrisheva O.V., Kovalev D.A., Kulichenko A.N. Phylogenetics, Evolution and Phylogeography of Bacillus anthracis. Bacteriology. 2018;3(2):57–63 (Article in Russian).
11. Bruce S.A., Schiraldi N.J., Kamath P.L., Easterday W. R., Turner W.C. A classification framework for Bacillus anthracis defined by global genomic structure. Evolutionary Applications. 2020;00:1–10. wileyonlinelibrary.
12. Sahl J.W., Pearson T., Okinaka R., Schupp J.M., Gillece J.D., Heaton H., et al. A Bacillus anthracis genome sequence from the Sverdlovsk 1979 autopsy specimens. mBio. 2016;7(5):e01501-16. doi:10.1128/mBio.01501-16.
13. Pisarenko S.V., Eremenko E.I., Ryazanova A.G., Kovalev D.A., Buravtseva N.P., Aksenova L.Yu., Dugarzhapova Z.F., Evchenko A.Yu., Kravets E.V., Semenova O.V., Bobrisheva O.V., Kuznetsova I.V., Golovinskaya T.M., Volynkina A.S., Balakhonov S.V., Kulichenko A.N. Phylogenetic analysis of Bacillus anthracis strains from Western Siberia reveals a new genetic cluster in the global population of the species. BMC Genomics.2019; 20:692 https://doi.org/10.1186/s12864-019-6060-z.
14. Lienemann T., Beyer W., Pelkola K., Rossow H., Rehn A., Antwerpen M., Grass G. Genotyping and phylogenetic placement of Bacillus anthracis isolates from Finland, a country with rare anthrax cases. BMC Microbiol. 2018;18:102 https://doi.org/10.1186/s12866-018-1250-4.
15. CLC Genomics Workbench 6 (https://digitalinsights.qiagen.com)
16. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018;35:1547–1549.
17. Rambaut A. FigTree version 1.4.2. 2014

Table

Table 1 – *B. anthracis* strains, used in the research.

№ п/п	*B. anthracis* strain	Date of isolation	Place of isolation	Origin of isolation
1	1	13.11.1967	Ukraine, Lviv Region, Lviv city	Carbuncle contents
2	12-16	28.06.1957	Russian Federation, Republic of Dagestan, Babaurtovskiy district	Carbuncle contents
3	12-16-1	12.11.1975	Russian Federation, Stavropol, Research Anti Plague Institute of Caucasus and Transcaucasia	Derivative of the 12-16 strain
4	14-41	27.12.1963	Russian Federation, Republic of Dagestan, Hasavurtovskiy district	Punctate of an ulcer
5	14-41-1	12.11.1975	Russian Federation, Stavropol, Research Anti Plague Institute of Caucasus and Transcaucasia	Derivative of the 14-41 strain
6	81-1	23.07.1969	Russian Federation, Stavropol Territory, Novoaleksanrovskiy district, Rashevatka village	Carbuncle contents
7	1(CO)	18.04.1968	Russian Federation, Republic of North Ossetia-Alania, Prigorodniy district	Material from cattle
8	140P	14.11.1979	Russian Federation, Tver Region, Tver city	Soil from livestock burial site
9	140P cap$^-$ wm	16.02.1992	Russian Federation, Stavropol Research Anti Plague Institute of Caucasus and Transcaucasia	Derivative of the 140P strain
10	228-269	04.02.1977	Russian Federation, Republic of North Ossetia-Alania, Digorskiy district	Soil from the cattle burial site
11	312-163	18.06.1978	Republic of Azerbaijan, Fizulinskiy district	Soil
12	850-46	07.07.1987	Russian Federation, Republic of Tatarstan, Kazan city	Material from a sick person
13	737-10	13.08.1984	Republic of Georgia, Tbilisi city	Material from a sick person
14	1051-35	1935	Russian Federation, Republic of Bashkortostan, Ufa city	Horse corpse
15	1266	07.11.2006	Russian Federation, Stavropol Territory, Kurskoy district, Sernovodskoe village	Soil from the place of slaughtering of bovine
16	1269	11.05.2007	Russian Federation, Republic of North Ossetia-Alania, Mozdokskiy region, Mozdok city	Cerebrospinal fluid from a sick person
17	1284	11.08.2010	Russian Federation, Omsk Region, Omsk city	Meat product

18	1307	05.08.2013	Russian Federation, Stavropol Territory, Izobilnenskiy district, Solnechnodolsk city	Eschar from a sick person
19	1322	13.08.2013	Russian Federation, Stavropol Territory, Izobilnenskiy district, Solnechnodolsk city	Corpse of a sheep
20	1339_24	27.07.2016	Russian Federation, Yamalo-Nenets autonomous district, in the Lake Pisieto area	Blood of a deer
21	1342_12	29.07.2016	Russian Federation, Yamalo-Nenets autonomous district, Salekhard city	Cutaneous ulcer lavage
22	1368-1	24.07.2018	Russian Federation, Republic of Tyva, Barun-Khemchinskiy district	Lung of a calf
23	1369-2	28.07.2018	Russian Federation, Republic of Tyva, Barun-Khemchinskiy district	Ear of the dead cattle
24	1370-3	28.07.2018	Russian Federation, Republic of Tyva, Barun-Khemchinskiy district	Ear of the dead cattle
25	1373-865	16.10.2019	Russian Federation, Republic of Dagestan, Novolakskiy district, Novokuli village	Cow's meat
26	1381-15	30.08.2019	Republic of Armenia	Eschar from a sick person
27	I-9	22.08.1956	Russian Federation, Republic of Tyva, Kyzyl city	Cutaneous anthrax lesion
28	I-19	13.06.1959	Russian Federation, Chita Region, Borzinskiy district, Kailastuy village	Cattle corpse
29	I-29	20.10.1961	Russian Federation, Altai Republic, Elekmonarskiy distict, Cheposh village	Carbuncle contents
30	I-35	01.09.1965	Russian Federation, Khabarovsk Territory, Viazemskiy district, Lermontovka vilage	Blood of sick person
31	I-63	29.07.1967	Russian Federation, Krasnoyarsk Territory, Krasnoturanskiy district, Baikalovo village	Soil
32	I-217	05.05.1981	Russian Federation, Tyumen Region, Tobolsk sity	Human, carbuncle contents
33	I-271-IRK	27.07.1980	Russian Federation, Yakut ASSR, Zhigan district, river Muna settlement	Soil of livestock burial site
34	I-271-OBL	No data available	Russian Federation, Yamalo-Nenets autonomous district	Corpse of a deer
35	I-275	17.03.1979	Russian Federation, Primorsk Territory, Kavalerovskiy district	Liver from the human corpse
36	I-319	28.07.1983	Russian Federation, Omsk Region, Omsk sity	Human, cutaneous ulcer lavage
37	I-323	22.07.1983	Russian Federation, Omsk Region, Omsk	Meat products

			sity	
38	I-360	11.08.2006	Russian Federation, Altai Territory, Krasnoshekovskiy district, Maralikha village	Human cutaneous ulcer lavage
39	I-361	03.09.2007	Russian Federation, Republic of Buryatia, Tunkinskiy district, Kyren village	Cattle corpse
40	I-362	06.07.2008	Russian Federation, Republic of Buryatia, Barguzinskiy district	Meat of a slaughtered calf
41	I-367	03.08.2008	Russian Federation, Republic of Buryatia, Barguzinskiy district	Soil from livestock burial site
42	I-368	26.08.2010	Russian Federation, Omsk Region, Tyukalinskiy district	Blood of the sick human
43	I-370	19.09.2012	Russian Federation, Altai Territory, Celinnyy district, Druzba village	Soil from the place of slaughtering of bovine
44	I-373	06.11.2012	Russian Federation, Altai Territory, Bystroistokskiy district, Bystriy Istok village	Cattle corpse

Figures

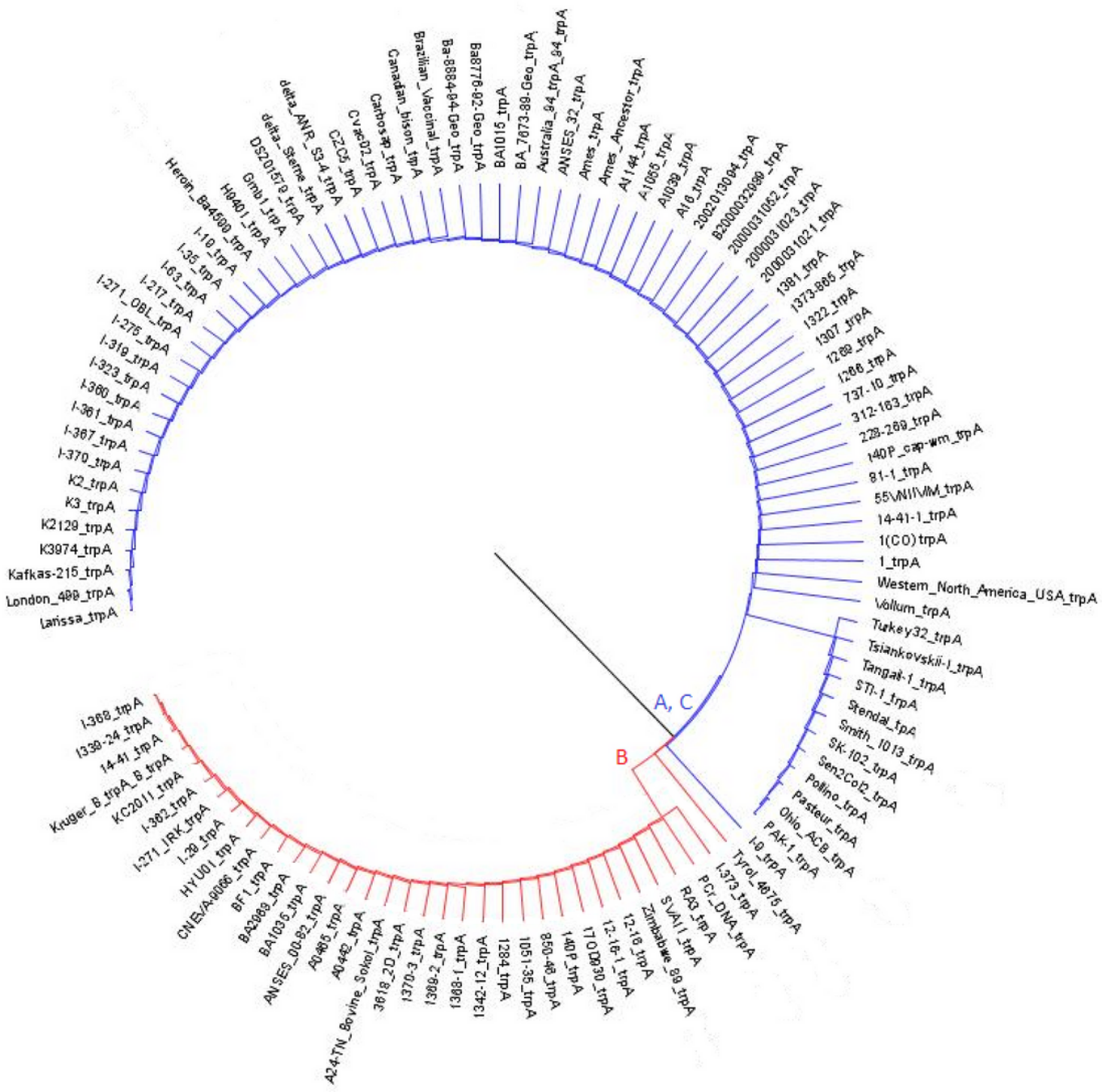

Figure 1

Dendrogram of cluster analysis of the trpA gene sequences of B. anthracis strains

Figure 2

Dendrogram of the cluster analysis of the trpD gene sequences of B. anthracis strains

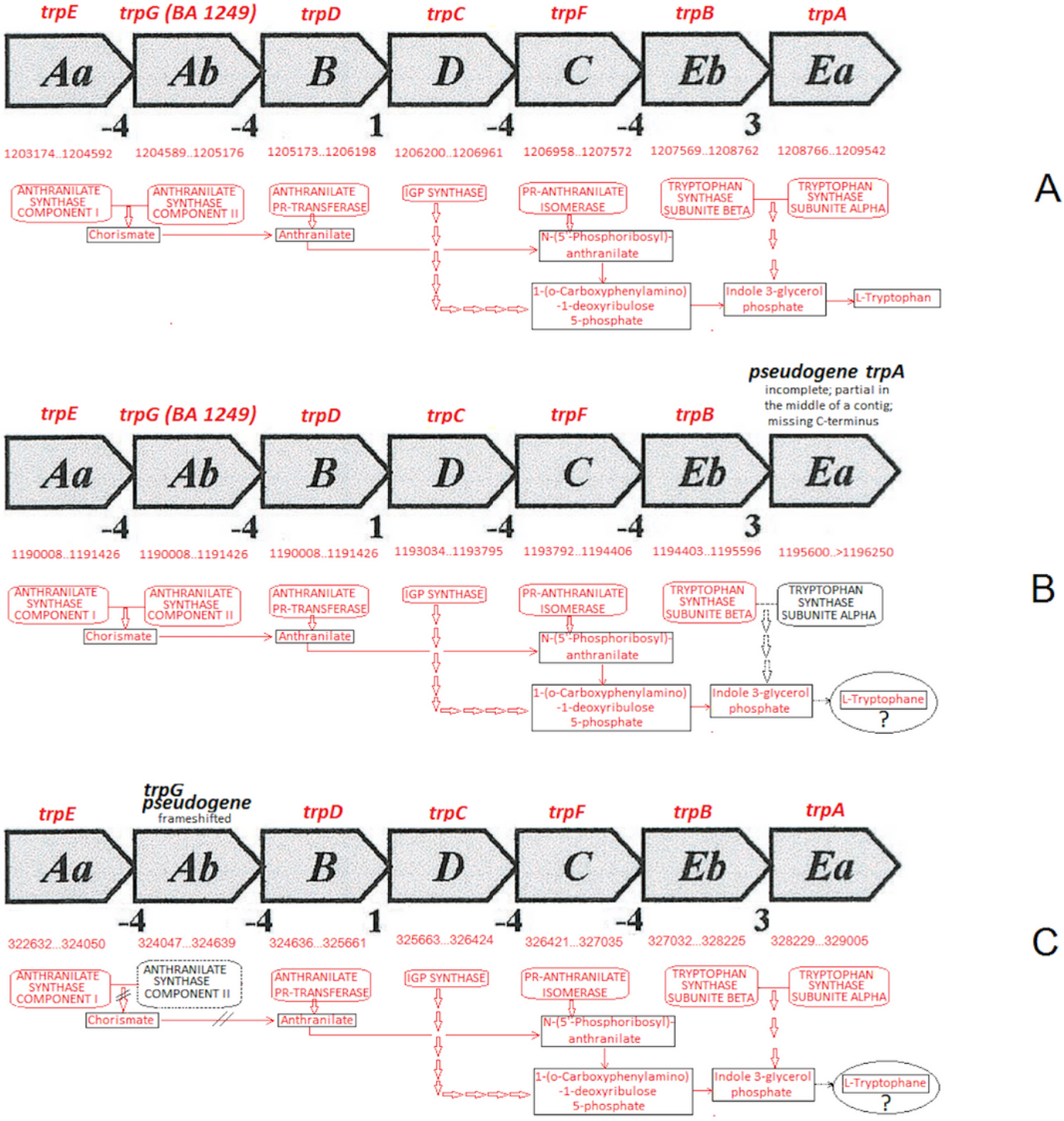

Figure 3

Schematic representation of tryptophan operon genes. A, B, C - operons of strains of the three main lineages of B. anthracis. The generally accepted designations of genes are given (above the figures depicting genes) and recommended by Xie et al. (inside the figures). The positions of the genes in the genomes of the strains Ames Ancestor (A), HYU01 (B) and 2000031052 (C), the names of the enzymes

encoded by the genes and the sequence of reactions with intermediate products of the tryptophan biosynthetic pathway (adapted from Xie et al., 2003) are indicated below.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- genomemetadate.xlsx