MLVA- and SNP analysis identified an unique genetic cluster in Bulgarian *Bacillus anthracis* strains

M. Antwerpen[1*], D. Ilin[2], E. Georgieva[2], H. Meyer[1], E. Savov[2] and D. Frangoulidis[1]

[1] Bundeswehr Institute of Microbiology, Munich, Germany

[2] Department of Clinical Microbiology, Military Medical Academy, Sofia, Republic of Bulgaria

*corresponding author: markusantwerpen@bundeswehr.org

Bundeswehr Institute of Microbiology

Neuherbergstr. 11

80937 Muenchen

Germany

European Journal of Clinical Microbiology & Infectious Diseases

Abstract

A collection of 40 *Bacillus (B.) anthracis* strains mostly isolated from soil in Bulgaria between 1960 and 1980 were investigated. All strains were proven to be *B. anthracis* by culture and amplification of a *B. anthracis*-specific chromosomal marker. PCR demonstrated that in 9 strains both virulence plasmids (pX01+/pX02+) and in 4 strains only one plasmid (pX02+) were present, whereas the majority of strains (n=27) lacked both plasmids (pX01-/pX02-). Multi-Locus-Variable Number of Tandem Repeat-Analysis (MLVA) using 15 markers differentiated three genotypes. Comparison with typing data of more than 1.000 different *B. anthracis* strains revealed that Bulgarian genotypes affiliated with the A1.a cluster and form an own unique cluster different from clusters containing strains isolated in geographical proximity like Turkey, Georgia, Hungary, Albania or Italy. In addition, a new allele of one marker (vrrC2) was identified. Canonical Single Nucleotide Polymorphisms analysis allocated 31 Bulgarian strains into the A.Br.008/009 and 9 strains into the A.Br.WNA group which is the first description of *B. anthracis* strains of the A.Br.WNA group on the Eurasian continent.

Introduction

Bacillus (B.) anthracis, a gram-positive, aerobe, spore-forming bacterium is found throughout the world. It is the causative agent of anthrax, a disease of wild animals and domestic livestock, predominantly cattle and goats. Anthrax is a zoonotic disease and humans can be infected by virulent strains harboring the virulence plasmids pXO1 and pXO2. Depending on the route of infection, pulmonary or inhalational anthrax, gastrointestinal or cutaneous anthrax can develop and, if left untreated, especially in case of pulmonary anthrax, death may occur [1]. Spores of *B. anthracis* may persist in the soil for decades. Recently, spores have been used in bioterrorist attacks [2-5].

DNA-based typing tools have been developed to evaluate the diversity of *B. anthracis* isolates from different geographic areas. These studies have also been undertaken with a view to trace intentionally released strains back to their origins [5-7]. Multi-Locus-Variable Number of Tandem Repeat (VNTR)-Analysis (MLVA) became one of the major assays for molecular typing of *B. anthracis* [8, 9]. Up-to-now, most strains are analyzed by MLVA using 15 loci, representing eleven chromosomal and four plasmid loci. Further progress in differentiating strains is increasingly being achieved by whole genome sequencing which efficiently identifies the rare SNPs present in *B. anthracis* [10]. Single nucleotide polymorphisms (SNPs) were identified which define twelve major clonal lineages [9]. A molecular typing strategy was presented which applies canonical-SNPs analysis and VNTR analysis using 15 markers. This strategy had been used to subtype a collection of 1,044 *B. anthracis* isolates from 42 countries and an extensive genotype data set was created. These analyses divided the isolates into three previously recognized major lineages (A, B, and C), with further subdivision into 12 clonal groups and, finally, 221 unique MLVA15 genotypes. By this, the global spread and regional diversity of *B. anthracis* became more evident [9].

Whereas genotyping data from Western and Central European countries as well as from countries in the South and East of the Black Sea have been published [8, 11], there are no data available from the Bulgarian and Romanian area which forms a bridge between those regions.

In this study we analyzed 40 Bulgarian *B. anthracis* strains, mostly isolated from soil, by canSNP typing and MLVA15 and demonstrate that Bulgarian strains form a unique genetic cluster.

Material and Methods

Bacterial strains

A collection of 40 *B. anthracis* strains were obtained from the Historical Collection of the Medical Military Academy of Sofia, Bulgaria (table 1). Bulgaria is situated in the Eastern part of Europe on the Western shore of the Black Sea (figure 1). All strains have been isolated between 1960 and 1980 in different districts of Bulgaria from soil with the exception of two strains originating from cattle (isolate 3187 and 3186). The majority of the isolates (n=27) were collected in the South and North-East of Bulgaria (regions of Jambol, Razgrad and Silistra). All strains were cultivated in a biosafety level 3 (BSL) laboratory on Columbia blood agar at 37°C. Colony characteristics were in concordance with a *B. anthracis*-typical phenotype, i. e. flat, dull-grey, irregular bordered, sticky colonies with no hemolysis. DNA isolation was performed as previously described [12]. Strains were confirmed to be *B. anthracis* based on a positive amplification of the chromosomal target *dhp61* [12]. Presence or absence of plasmids pXO1 and pXO2 was determined by amplifying fragments of the target genes *pagA* and *capB* using the RealTime PCR Bacillus anthracis – Detection Kit (Roche Diagnostics, Penzberg, Germany) according to the protocol of the manufacturer.

Multi-Locus-Variable Number of Tandem Repeat-Analysis (MLVA)

MLVA was done using 15 markers according to the protocol published by van Ert *et al* [9]. Platinum Taq polymerase, $MgCl_2$ and water were purchased from Invitrogen (Karlsruhe, Germany) and Eppendorf (Hamburg, Germany), respectively. Primers were synthesized by Tib-MolBiol (Berlin, Germany). Multiplex-PCR reactions were performed using the Thermocycler GenAmp PCR System 9700 (Perkin Elmer, Waltham, USA). PCR products were purified using the PCR Purification Kit (Qiagen, Hilden, Germany) prior to loading on the ABI3100 sequencer (Applied Biosystems,

Darmstadt, Germany). Fragment lengths were determined by comparison to the internal length standard 1200LIZ (Applied Biosystems). Based on the size of the respective repeat unit and the fragment length, a repeat copy number was calculated for each marker according to the algorithm described by Vergnaud *et al* [13]. In order to compare data obtained in this study with typing data described by van Ert *et al.* [9, 14] we determined the repeat copy number of each marker of their strain collection after calibration with data of strains which had been analyzed in parallel in both laboratories (strains A9, A22, A30, A40, A41, A66, A67) and by including the entire genome sequencing data of strains Ames A2010 and Vollum (AAAC01000001.1 and NZ_AAEP00000000, resp.). In some cases sequencing of amplicons was performed using the respective MLVA primers and BigDye Terminator chemistry (Applied Biosystems) on a 3100 ABI sequencer (Applied Biosystems). The sequence of marker vrrC2 of strain 3189 was deposited in Genbank (accession number HM240285). Repeat units were identified by using the software Tandem Repeat (http://tandem.bu.edu/trf/trf.html). Data analysis was performed using the Software Bionumerics 5.1 (Applied Maths, Sint-Martens-Latem, Belgium). Cluster analysis was done using categorical coefficients and similarity trees based on the Unweighted Pair Group Method with Arithmetic mean (UPGMA) method.

Canonical Single Nucleotide Polymorphisms (canSNP) Analysis

CanSNP analysis was done according to van Ert *et al.* [9] with slightly modified probe sequences (see supplemental data, table 1), To further sub-classify strains belonging to the WNA cluster SNP analysis was performed according to Kenefic *et al.* [15]. All Primers and probes (supplemental data) were obtained from Tib-Molbiol (Berlin, Germany). PCR was performed with a final concentration of primers (0.3µM) and probes (0.1µM) using the ABI TaqMan-Mastermix (Applied Biosystems) and approx. 4 ng of template DNA using the MX3000 system (Stratagene, Agilent Technologies, Waldbronn, Germany). Initial denaturation at 96°C for 10 min was followed by 40 cycles, each consisting of denaturation (30 sec at 96°C), annealing (30 sec at 56°C) and elongation (30 sec at 72°C). Data evaluation was done using the algorithm published by van Ert *et al.* [9].

Results

All Bulgarian strains could be identified as *B. anthracis* based on amplification of a *B. anthracis*-specific chromosomal sequence targeted by the *dhp*61-PCR assay. 9 out of 40 strains, including strain 3187 isolated from cattle, harbored both virulence plasmids (pXO1 and pXO2, respectively), whereas 4 strains contained plasmid pXO2 only (see table 1). The 27 others lack both plasmids.

MLVA-Typing

UPGMA cluster analysis of the 9 Bulgarian strains containing both plasmids revealed three different genotypes designated GT1, GT2 and GT3 (table 1 and figure 2). GT2 consisting of 2 strains and GT3 consisting of 6 strains differ in the number of repeat copies of marker vrrA (10 versus 11 repeat copies) and marker pXO1 (7 versus 10 repeat copies). GT1 is represented by a single strain (3189) with an unique allele size for marker vrrC2. Sequencing confirmed that the fragment length of marker vrrC2 is indeed much lower (333 bp) as compared to already described fragment lengths ranging from 458 bp to 635 bp. This allele (HM240285) is coded 6 in agreement with the currently used convention in which the 532bp allele present in the sequenced Ames genome when using the same primer set is coded 17 [9].

Analysis of four strains containing only plasmid pX02 (i.e. strains 3188, 3171, 3167 and 3180) allowed differentiation of two genotypes designated GT2a and GT3a, respectively. These strains have identical repeat codes with respect to the 11 chromosomal (the “chromosomal backbone”) and the 3 plasmid pX02-specific markers as strains with both plasmids belonging to GT2 and GT3 (table 1). Likewise, all plasmid-less strains (n=27) belong to 2 genotypes designated GT2b and GT3b because the 11 chromosomal markers are identical as compared to strains of the genotypes GT2 and GT3 (table 1).

Clustering resulted in a similarity tree shown in figure 2. Because of identical repeat codes for the 11 chromosomal markers all strains clustered closely with either genotype GT2 or GT3. This variation is shown in boxed profiles with either GT2, GT2a and GT2b or GT3, GT3a and GT3b, respectively (figure 2).

MLVA-data in a global context

To put typing results of Bulgarian strains in a more global context, a modUPGMA analysis of all MLVA datasets (n=40) was conducted with already published datasets. To this end we included MLVA data of 1,044 *B. anthracis* strains described by van Ert *et al*. and 148 strains of our in-house collection. The analysis demonstrated that the three genotypes GT1, GT2 and GT3 as well as the four "sub-genotypes" GT2a, GT2b, GT3a and GT3b are novel and cluster separately (figure 3). According to the MLVA15 system of van Ert *et al*., they all represent a new sub-cluster within the A1.a cluster. Strains collected in countries located in the South and East of Bulgaria, such as Georgia, Turkey and Iran, belong to the South Caucasian A3.a cluster [8]. In contrast, strains isolated in countries in the West of Bulgaria, such as Albania, Hungary [9] and Italy [16] form a different sub-cluster within the A1.a cluster (figure 3). The highest similarity of the Bulgarian cluster is seen with the WNA-cluster (figure 3) consisting of strains exclusively isolated in Canada and in the Northern States of the US [9, 17]. Both clusters differ in the repeat copy number of three markers of the "chromosomal backbone" (VNTR12, VNTR19 and VNTR32). This is shown with a representative strain (A1119USA) of the WNA-cluster in Figure 2.

canSNP-Typing

CanSNP analysis separated the Bulgarian isolates into two different phylogenetic groups. 9 out of 40 strains (see table 1) belong to the A.Br.WNA group, whereas 31 strains belong to the A.Br.008/009

group. It is interesting to note that genotypes GT2b and GT3b are found in both, the A.Br.WNA and the A.Br.008/009 group. To further sub-classify the Bulgarian WNA-strains we used a SNP-based strategy described by Kenefic *et al.* [15] and determined the nucleotides T, C, and A for the SNPs wna237471, wna1141774 and wna3368524, respectively. Based on this unique signature the 9 Bulgarian WNA strains could not be classified in one of those six WNA-sub-clades described by Kenefic *et al.* [15].

Discussion

The actual method used for forensic investigations of *B. anthracis* strains and isolates is based on a typing strategy called "progressive hierarchical resolving assay using nucleic acids" (PHRANA) introduced by Keim *et al*. [18]. This strategy proposes in a first step to determine the canSNP group and in a second step to conduct MLVA.

canSNP-Typing

An unexpected result of the canSNP typing described here was that Bulgarian *B. anthracis* strains could be divided into two distinct canSNP groups, i.e. A.Br.008/009 and A.Br.WNA. According to our findings 9 out of 40 Bulgarian strains (all derived from soil) belong to canSNP group A.Br.WNA which is the first description of A.Br.WNA strains on the Eurasian continent. Based on typing data described in the literature for South-Caucasian strains from Georgia and Turkey [8] and for strains from Kazakhstan [11] we had assumed that Bulgarian strains would belong to the canSNP group A.Br.008/009. To further sub-classify WNA strains Kenefic *et al*. [15] had investigated DNA of 352 WNA-strains (according to van Ert *et al*.) and showed that 351 WNA-strains could be sub-classified into 6 distinct WNA-sub-clades. One strain (#2000032989), however, derived from a historical collection hold at CDC could not be grouped [15]. SNP analysis of the Bulgarian WNA-strains revealed identical SNPs as this strain for the investigated three nucleotide positions. The explanation that trading of contaminated animal products from North America has established a focus of A.Br.WNA

strains in Bulgaria is highly unlikely. Instead, the most parsimonious explanation is homoplasy of the WNA-specific point mutation: The same A.Br.WNA-specific mutation has occurred in the Bulgarian area as it had led to the formation of canSNP group A.Br.WNA in North America from its ancestor group A.Br.008/009. Differences in three chromosomal MLVA-markers observed between "American" representatives of the A.Br.WNA-SNP-group and Bulgarian-WNA-isolates support this theory and genome sequencing of a Bulgarian strain of the WNA cluster might improve our knowledge with respect to classification.

Homoplasy, which is very rarely seen in *B. anthracis* (1 in >25.0000 data points), has already been postulated to have occurred for strain A0303: this strain has been allocated to the WNA-clade although canSNP 2589947 shows the ancestral A type instead of the G expected in the WNA clade [15, 19].

MLVA-Typing

Historically, MLVA typing of *B. anthracis* strains started in the year 2000 and was based on eight markers (6 chromosomal and one plasmid marker each) which allowed the discrimination of 89 different genotypes out of 400 isolates [20]. Based on this 1st generation typing system, all Bulgarian strains have the identical MLVA profile with regard to the 6 chromosomal marker as strain A1119USA of the WNA-cluster [9], recently published Kazakh strains [11] and five genotypes of the MLVA8 system [20] (GT1, 2, 4, 5, 13 and 14). A few years later two different extensions of the MLVA8 system were developed by Le Fléche *et al*., Lista *et al.* and van Ert *et al*. [9, 21, 22] and are still in use for typing. The MLVA25 system described by Lista *et al.* showed a high discriminatory power especially for strains of their European collection. The MLVA15 system introduced by van Ert *et al*. was applied to 1,044 strains and contributed significantly to the understanding of the geographical diversity of *B. anthracis* in the world. Analysis of the Bulgarian strains with the MLVA15 system yielded a higher discriminatory power as compared to the MLVA8 system: Differences could be observed in three

chromosomal markers (vrrA, vrrC2, BaVNTR32) between the Bulgarian and the WNA-cluster, whereas the patterns only show only up to two differences dependent on the strain investigated (figure 2).

In a given geographic location strains of *B. anthracis* may have different origins. Either the strains are indeed genuine for this region or they had been imported into the region from other areas by animal or human factors, such as trading. Investigation of spores collected in a Belgian wool factory [23] and retrospective investigations of outbreaks in Switzerland in 1979 [24] show, that the outbreaks could be linked to spore-contaminated wool either from Kazakhstan or the Middle East. These examples highlight that only accurate and detailed information about an outbreak can identify its real origin. Typing data derived from a strain with no further information do not necessarily contribute to the understanding of the global diversity of anthrax.

An interesting option to identify genuine strains of a given geographical region is to analyze soil-associated *B. anthracis* strains. It is known that strains may lack one or even both plasmids [25] after long-term-storage based on repeated passage, instead of lyophilization. We assume that with Bulgarian collection, a loss of plasmids had occurred over the years resulting in 27 plasmid-less strains out of 40 strains.

However, using the existing algorithms based on a categorical coefficient and an “unweighted pair group mean arithmetic-analysis”, plasmid-less strains or strains lacking one plasmid would be grouped into different clusters distinct to strains with the identical chromosomal backbone and plasmids (data not shown).

This results in false clustering of plasmid-less strains in the UPGMA dendrogram. Up to now this mathematical problem has not been discussed yet, because only a few genotypes of plasmid-less strains are part of analyses described , e.g. 7 GTs out of 221 GT [9]. In the context of the “global” *B. anthracis* collection of van Ert *et a.l* [9] consequences of false clustering were not obvious for the authors. In our geographically rather limited collection consequences of false clustering stand out. By using the feature of the categorical coefficient clustering of Bionumerics 5.1 missing markers are

ignored and the data analysis is based on chromosomal markers only. As a consequence, all strains with the same chromosomal background cluster together (boxed genotypes in figure 2) and are assigned as subgroups of the respective genotypes. A typing system relying on chromosomal markers only seems to be useful for epidemiology and interpretation of diversity, in forensic investigations of virulent *B. anthracis* strains all typing data have to be taken into account.

Molecular evolution in Bulgaria

Data interpretation of epidemiologically relevant MLVA profiles in combination with phylogenetically relevant results of the canSNP typing revealed an interesting finding. *B. anthracis* strains with genotypes GT2b and GT3b, lacking plasmid markers, are present in both canSNP groups A.Br.008/009 and A.Br.WNA. A similar finding has been reported by Simonson *et al.* [17]. They identified the identical MLVA15 genotype for Chinese strains A0576CHI and A0731CHI, which belong to two different canSNP groups (A.Br.Ames and A.Br.001/002, respectively). With regard to phylogeny this points to two different populations existing in the same area. Against the background of different mutation frequencies of MLVA markers and canSNPs, it has to be assumed, that first the mutation of the SNP occurred resulting in two different populations in Bulgaria. Lateron, MLVA markers evolved in parallel. The Bulgarian and the Chinese data demonstrate that MLVA and canSNP typing do not always led to the same result. Therefore, interpretation of only MLVA-based typing results with regard to phylogeny has to be performed with caution.

Conclusion:

In this study we demonstrate that Bulgarian *B. anthracis* strains form a geographically separated cluster with 3 unique MLVA genotypes different from genotypes found in neighboring countries. In addition, canSNP analysis differentiated Bulgarian *B. anthracis* strains into two groups, i.e. the A.Br.008 and the A.Br.WNA group, represented by a new WNA sub-clade. This the first description of A.Br.WNA strains on the Eurasian continent. The genetic typing data provided here, including data of strains lacking either one or both plasmids, increase our knowledge of the diversity of *B. anthracis* in

the Black Sea-region and may help to extend the understanding of the global distribution and genetic relationship of *B. anthracis* in the world.

Acknowledgment

We gratefully acknowledge Claudia Kahlhofer and Linda Dobrzykowski for excellent technical assistance. We also thank the unknown reviewers for helpful comments.

References

[1] WHO (2008) Anthrax in humans and animals. http://www.who.int/csr/resources/publications/AnthraxGuidelines2008/en/index.html, WHO Library Cataloguing, Accessed 11 December 2010

[2] Greene CM, Reefhuis J, Tan C, Fiore AE, Goldstein S, Beach MJ, Redd SC, Valiante D, Burr G, Buehler J, Pinner RW, Bresnitz E, Bell BP (2002) Epidemiologic investigations of bioterrorism-related anthrax, New Jersey, 2001. Emerg Infect Dis 8:1048-1055

[3] Holtz TH, Ackelsberg J, Kool JL, Rosselli R, Marfin A, Matte T, Beatrice ST, Heller MB, Hewett D, Moskin LC, Bunning ML, Layton M (2003) Isolated case of bioterrorism-related inhalational anthrax, New York City, 2001. Emerg Infect Dis 9:689-696

[4] Jernigan JA, Stephens DS, Ashford DA, Omenaca C, Topiel MS, Galbraith M, Tapper M, Fisk TL, Zaki S, Popovic T, Meyer RF, Quinn CP, Harper SA, Fridkin SK, Sejvar JJ, Shepard CW, McConnell M, Guarner J, Shieh WJ, Malecki JM, Gerberding JL, Hughes JM, Perkins BA (2001) Bioterrorism-related inhalational anthrax: the first 10 cases reported in the United States. Emerg Infect Dis 7:933-944

[5] Keim P, Smith KL, Keys C, Takahashi H, Kurata T, Kaufmann A (2001) Molecular investigation of the Aum Shinrikyo anthrax release in Kameido, Japan. J Clin Microbiol 39:4566-4567

[6] Takahashi H, Keim P, Kaufmann AF, Keys C, Smith KL, Taniguchi K, Inouye S, Kurata T (2004) *Bacillus anthracis* incident, Kameido, Tokyo, 1993. Emerg Infect Dis 10:117-120

[7] Van Ert MN, Easterday WR, Simonson TS, U'Ren JM, Pearson T, Kenefic LJ, Busch JD, Huynh LY, Dukerich M, Trim CB, Beaudry J, Welty-Bernard A, Read T, Fraser CM, Ravel J, Keim P (2007) Strain-specific single-nucleotide polymorphism assays for the *Bacillus anthracis* Ames strain. J Clin Microbiol 45:47-53

[8] Merabishvili M, Natidze M, Rigvava S, Brusetti L, Raddadi N, Borin S, Chanishvili N, Tediashvili M, Sharp R, Barbeschi M, Visca P, Daffonchio D (2006) Diversity of *Bacillus anthracis* strains in Georgia and of vaccine strains from the former Soviet Union. Appl Environ Microbiol 72:5631-5636

[9] Van Ert MN, Easterday WR, Huynh LY, Okinaka RT, Hugh-Jones ME, Ravel J, Zanecki SR, Pearson T, Simonson TS, U'Ren JM, Kachur SM, Leadem-Dougherty RR, Rhoton SD, Zinser G, Farlow J, Coker PR, Smith KL, Wang B, Kenefic LJ, Fraser-Liggett CM, Wagner DM, Keim P (2007) Global genetic population structure of *Bacillus anthracis*. PloS one 2:e461

[10] Kuroda M, Serizawa M, Okutani A, Sekizuka T, Banno S, Inoue S (2010) Genome-wide single nucleotide polymorphism typing method for identification of *Bacillus anthracis* species and strains among *B. cereus* group species. J Clin Microbiol 48:2821-2829

[11] Aikembayev AM, Lukhnova L, Temiraliyeva G, Meka-Mechenko T, Pazylov Y, Zakaryan S, Denissov G, Easterday WR, Van Ert MN, Keim P, Francesconi SC, Blackburn JK, Hugh-Jones M, Hadfield T (2010) Historical distribution and molecular diversity of *Bacillus anthracis*, Kazakhstan. Emerg Infect Dis 16:789-796

[12] Antwerpen M, Zimmermann P, Frangoulidis D, Bewley K, Meyer H (2008) Real-time PCR system targeting a chromosomal marker specific for *Bacillus anthracis*. Mol Cell Probes 22:313-315

[13] Vergnaud G, Pourcel C (2006) Multiple Locus VNTR (Variable Number of Tandem Repeat) Analysis (MLVA). In: E S (ed) Molecular Identification, Systematics and Population Structure of Prokaryotes. Springer Publishing, pp 83-104

[14] Keim P, Kalif A, Schupp J, Hill K, Travis SE, Richmond K, Adair DM, Hugh-Jones M, Kuske CR, Jackson P (1997) Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers. J Bacteriol 179:818-824

[15] Kenefic LJ, Pearson T, Okinaka RT, Schupp JM, Wagner DM, Hoffmaster AR, Trim CB, Chung WK, Beaudry JA, Jiang L, Gajer P, Foster JT, Mead JI, Ravel J, Keim P (2009) Pre-Columbian origins for North American anthrax. PloS one 4:e4813

[16] Fasanella A, Van Ert M, Altamura SA, Garofolo G, Buonavoglia C, Leori G, Huynh L, Zanecki S, Keim P (2005) Molecular diversity of *Bacillus anthracis* in Italy. J Clin Microbiol 43:3398-3401

[17] Simonson TS, Okinaka RT, Wang B, Easterday WR, Huynh L, U'Ren JM, Dukerich M, Zanecki SR, Kenefic LJ, Beaudry J, Schupp JM, Pearson T, Wagner DM, Hoffmaster A, Ravel J, Keim P (2009) *Bacillus anthracis* in China and its relationship to worldwide lineages. BMC Microbiol 9:71

[18] Keim P, Van Ert MN, Pearson T, Vogler AJ, Huynh LY, Wagner DM (2004) Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. Infect Genet Evol 4:205-213

[19] Pearson T, Busch JD, Ravel J, Read TD, Rhoton SD, U'Ren JM, Simonson TS, Kachur SM, Leadem RR, Cardon ML, Van Ert MN, Huynh LY, Fraser CM, Keim P (2004) Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. Proc Natl Acad Sci U S A 101:13536-13541

[20] Keim P, Price L, Klevytska A, Smith K, Schupp J, Okinaka R, Jackson P, Hugh-Jones M (2000) Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J Bacteriol 182:2928-2936

[21] Le Fléche P, Hauck Y, Oteniente L, Prieur A, Denoeud F, Ramisse V, Sylvestre P, Benson G, Ramisse F, Vergnaud G (2001) A tandem repeat database for bacterial genomes: application to the genotyping of *Yersinia pestis* and *Bacillus anthracis*. BMC Microbiol 1:2

[22] Lista F, Faggioni G, Valjevac S, Ciammaruconi A, Vaissaire J, le Doujet C, Gorge O, De Santis R, Carattoli A, Fasanella A, Orsini F, D'Amelio R, Pourcel C, Cassone A, Vergnaud G (2006) Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. BMC Microbiol 6:33

[23] Wattiau P, Klee SR, Fretin D, Van Hessche M, Ménart M, Franz T, Chasseur C, Butaye P, Imberechts H (2008) Occurence and genetic diversity of *Bacillus anthracis* strains isolated in an active wool-cleaning fcatory. Appl Environ Microbiol 74:4005-4011

[24] Pilo P, Perreten V, Frey J (2008) Molecular epidemiology of *Bacillus anthracis*: determining the correct origin. Appl Environ Microbiol 74:2928-2931

[25] Marston CK, Hoffmaster AR, Wilson KE, Bragg SL, Plikaytis B, Brachman P, Johnson S, Kaufmann AF, Popovic T (2005) Effects of long-term storage on plasmid stability in *Bacillus anthracis*. Appl Environ Microbiol 71:7778-7780

Table 1: Variable Number of Tandem Repeat sizes for 40 Bulgarian *Bacillus anthracis* strains.
#Analysis was performed according to van Ert *et al.* [9] and the repeat codes were calculated according to Vergnaud *et al.* [13].
*fragment length which has not been described so far for allele vrrC2.
§ isolated from cattle

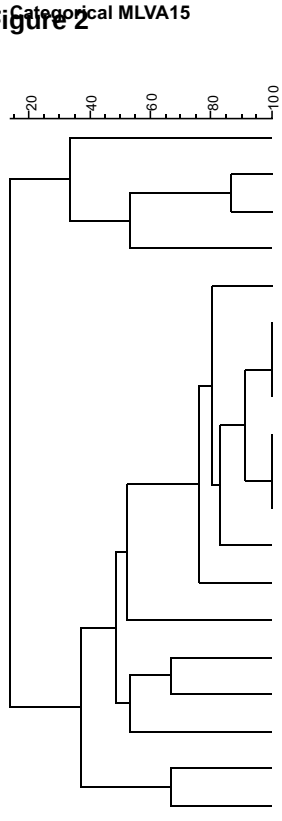
Figure 1: Map of Europe with Bulgaria located on the Western shore of the Black Sea.

Figure 2: MLVA based dendrogram of Bulgarian *Bacillus anthracis* strains. One representative strain each belonging to one of the 12 canSNP group was included for comparison. Bulgarian strains cluster separately from the A.Br.WNA canSNP group. Boxed MLVA patterns highlight the three Bulgarian genotypes with respective sub-branches.

Figure 3: Dendrogram based on MLVA typing data of 1.191 *Bacillus anthracis* strains and isolates using the modUPGMA-clustering method. MLVA clusters were assigned normal letters, geographical associated clusters are in italic letters.

Figure

Figure 2



modUPGMA

vrrA	vrrB1	vrrB2	vrrC1	vrrC2	CG3	pX01	pX02	Bavntr12	Bavntr16	Bavntr17	Bavntr19	Bavntr23	Bavntr32	Bavntr35		
11	23	8	28	21	2		4	5	24	6	5	2	22	5	A1055USA	C.Br.A1005
9	19	8	53	17	2	7	10	5	24	3	5	3	11	3	A0382SOA	B.Br.001/002
9	19	8	53	17	2	6	9	5	24	3	5	3	11	3	A0442SOA	B.Br.Kruger
9	15	5	53	17	2	12	6	5	22	3	5	3	14	5	A0333GER	B.Br.CNEVA
10	16	7	57	21	1	8	7	6	21	4	5	4	9	4	A1119USA	A.Br.WNA
11	16	7	57	21	1	10	8	6	21	4	5	4	13	4	3177 + 5 isolates	GT3
11	16	7	57	21	1		8	6	21	4	5	4	13	4	3180 + 2 isolates	GT3a
11	16	7	57	21	1			6			5	4	13	4	3151 + 17 isolates	GT3b
10	16	7	57	21	1	7	8	6	21	4	5	4	13	4	3187 + 1 isolate	GT2
10	16	7	57	21	1		8	6	21	4	5	4	13	4	3188	GT2a
10	16	7	57	21	1			6			5	4	13	4	3150 + 8 isolates	GT2b
10	16	7	57	6	1	10	11	6	21	4	5	4	13	4	3189	GT1
10	16	7	57	21	1	7	7	6	20	4	4	4	13	4	A0280ITA	A.Br.008/009
8	16	6	48	21	2	10	8	6	20	4	5	2	13	4	A0416ZIM	A.Br.Vollum
8	16	7	57	17	2	9	9	6	20	5	4	4	12	4	A03881ARG	A.Br.003/004
10	16	7	57	17	2	8	11	6	20	3	4	4	16	4	A0406IND	A.Br.Aust94
10	16	7	57	17	2	7	8	5	4	5	6	4	13	4	A0421IDO	A.Br.005/006
10	16	7	53	17	2	6	7	6	4	4	4	3	16	5	A1045USA	A.Br.001/002
10	16	6	53	17	2	7	9	6	20	4	4	4	16	5	A2012USA	A.Br.Ames

Figure 3
Click here to download high resolution image

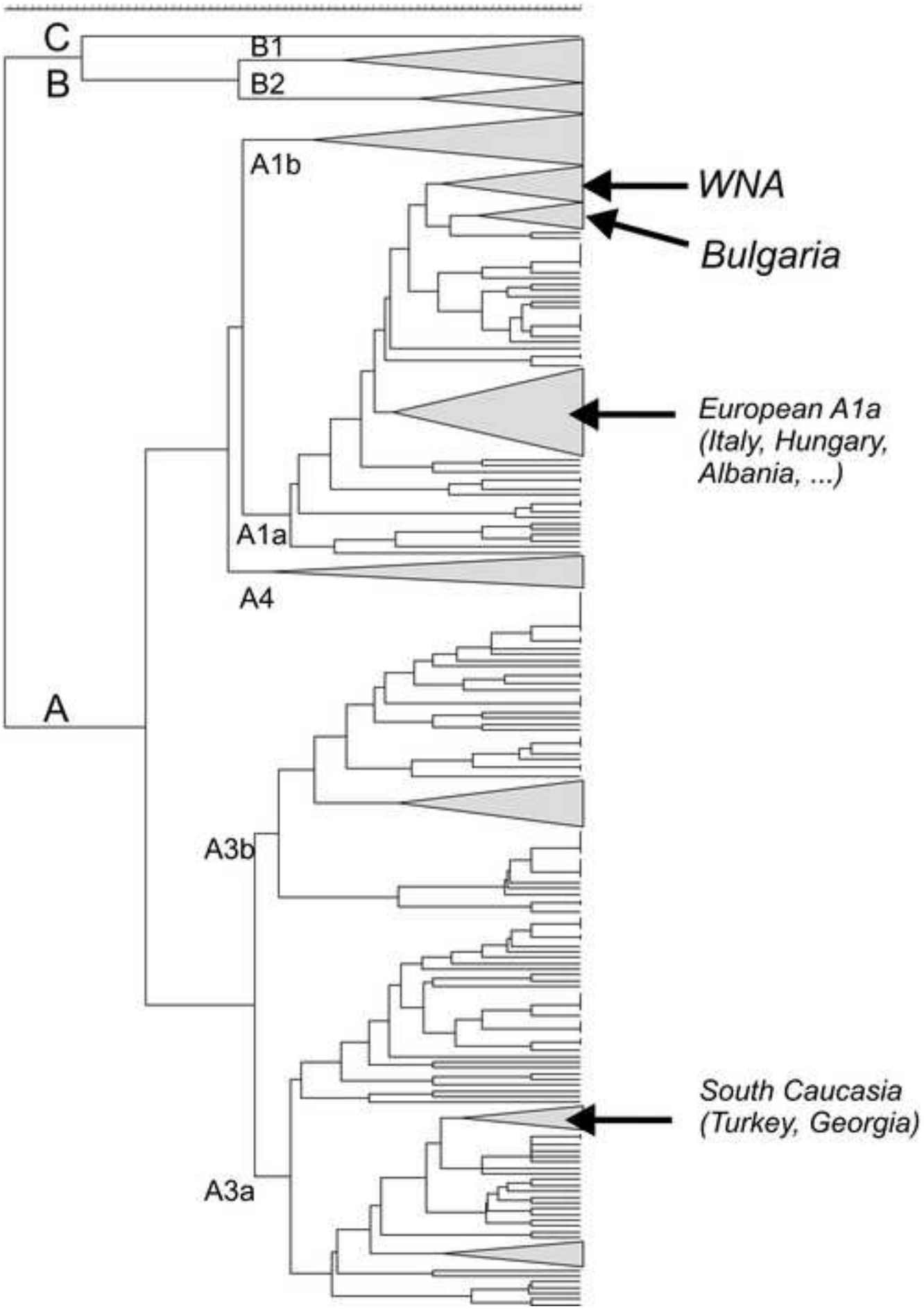


Table 1

designation	district	canSNP	MLVA-Genotype	vrrA	vrrB1	vrrB2	vrrC1	vrrC2	CG3	pXO1	pXO2	BaVNTR12	BaVNTR16	BaVNTR17	BaVNTR19	BaVNTR23	BaVNTR32	BaVNTR35
3189	Sofia	A.Br.008/009	GT1	10	16	7	57	6*	1	10	11	6	21	4	5	4	13	4
§3187	Kjustendil		GT2	10	16	7	57	21	1	7	8	6	21	4	5	4	13	4
3185	Sofia			10	16	7	57	21	1	7	8	6	21	4	5	4	13	4
3188	Sofia		GT2a	10	16	7	57	21	1		8	6	21	4	5	4	13	4
3170	Silistra		GT2b	10	16	7	57	21	1			6			5	4	13	4
3179	Silistra		GT3	11	16	7	57	21	1	10	8	6	21	4	5	4	13	4
3177	Silistra			11	16	7	57	21	1	10	8	6	21	4	5	4	13	4
3168	Tarnovo			11	16	7	57	21	1	10	8	6	21	4	5	4	13	4
3166	Tarnovo			11	16	7	57	21	1	10	8	6	21	4	5	4	13	4
3165	Jambol			11	16	7	57	21	1	10	8	6	21	4	5	4	13	4
3161	Jambol			11	16	7	57	21	1	10	8	6	21	4	5	4	13	4
3171	Silistra		GT3a	11	16	7	57	21	1		8	6	21	4	5	4	13	4
3180	Silistra			11	16	7	57	21	1		8	6	21	4	5	4	13	4
3167	Tarnovo			11	16	7	57	21	1		8	6	21	4	5	4	13	4
3184	Sofia		GT3b	11	16	7	57	21	1			6			5	4	13	4
3183	Sofia			11	16	7	57	21	1			6			5	4	13	4
3182	Sofia			11	16	7	57	21	1			6			5	4	13	4
3176	Silistra			11	16	7	57	21	1			6			5	4	13	4
3175	Silistra			11	16	7	57	21	1			6			5	4	13	4
3178	Silistra			11	16	7	57	21	1			6			5	4	13	4
3164	Jambol			11	16	7	57	21	1			6			5	4	13	4
3163	Jambol			11	16	7	57	21	1			6			5	4	13	4
3162	Jambol			11	16	7	57	21	1			6			5	4	13	4
3181	Sofia			11	16	7	57	21	1			6			5	4	13	4
3160	Razgrad			11	16	7	57	21	1			6			5	4	13	4
3159	Razgrad			11	16	7	57	21	1			6			5	4	13	4
3158	Razgrad			11	16	7	57	21	1			6			5	4	13	4
3156	Jambol			11	16	7	57	21	1			6			5	4	13	4
3154	Razgrad			11	16	7	57	21	1			6			5	4	13	4
3152	Jambol			11	16	7	57	21	1			6			5	4	13	4
3151	Jambol			11	16	7	57	21	1			6			5	4	13	4
3174	Silistra	A.Br.WNA	GT2b	10	16	7	57	21	1			6			5	4	13	4
3173	Silistra			10	16	7	57	21	1			6			5	4	13	4
3172	Silistra			10	16	7	57	21	1			6			5	4	13	4
3150	Jambol			10	16	7	57	21	1			6			5	4	13	4
3157	Jambol			10	16	7	57	21	1			6			5	4	13	4
3155	Jambol			10	16	7	57	21	1			6			5	4	13	4
3153	Razgrad			10	16	7	57	21	1			6			5	4	13	4
§3186	Sofia			10	16	7	57	21	1			6			5	4	13	4
3169	Tarnovo		GT3b	11	16	7	57	21	1			6			5	4	13	4

Electronic Supplementary Material

Click here to download Electronic Supplementary Material: supplementary data .xls