Erciyes Med J 2020; 42(4): 362–9 • DOI: 10.14744/etd.2020.64920
INVITED REVIEW – OPEN ACCESS

Bacillus anthracis Evolution: Taking Advantage of the Topology of the Phylogenetic Tree and Human History to Propose Dating Points

Gilles Vergnaud

ABSTRACT

This review expands on a talk I gave during the "Biology of Anthrax" meeting held in Bari, Italy (September 3rd–6th, 2019). The talk was a synthesis of recent investigations taking advantage of the topology of the *Bacillus anthracis* phylogenetic tree to propose tentative dating points and scenarios. Currently available whole-genome sequence (WGS) data allowed identifying single nucleotide polymorphisms (SNPs) among *B. anthracis* strains and drawing phylogenetic trees. The geographic origin of the strains and the topology of the tree were used to infer spreading events. Five star-like patterns in the tree (polytomies), each containing at least six branches, were detected. The analysis of the geographic distribution of the strains constituting one such polytomy suggests that it emerged not more than a few centuries ago. The key observation allowing this dating is the finding that the polytomy is anchored into Western Europe and that the main North-American lineage emerged from one of its branches, indicative of a post-Columbian export. From this point, I propose additional working hypotheses which may allow dating key nodes along the phylogeny of *B. anthracis* corresponding to four "Out-of-Africa" events. While trade of contaminated animal products seems to be the predominant driving force underlying modern long-distance spreading of *B. anthracis*, invasive military operations and more generally borders instabilities may have played an important role in earlier times. The testing of these hypotheses will require the sequencing of a significant number of additional strains from many countries.

Keywords: Anthrax, phylogeography, *Bacillus anthracis*, ecotype, communicable diseases, whole genome sequencing

Cite this article as:
Vergnaud G. *Bacillus anthracis* Evolution: Taking Advantage of the Topology of the Phylogenetic Tree and Human History to Propose Dating Points. Erciyes Med J 2020; 42(4): 362-9.

Paris-Saclay University, CEA, CNRS, Institute of Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France

Submitted
13.03.2020

Accepted
08.04.2020

Available Online Date
21.04.2020

Correspondence
Gilles Vergnaud,
Bâtiment 400, Rue Gregor Mendel Campus Orsay
Université Paris-Saclay
91405 Orsay Cedex, France
Phone: +33 (0)1 69 82 62 09
e-mail: gilles.vergnaud@universite-paris-saclay.fr

INTRODUCTION

Bacillus anthracis is a soil-borne spore-forming bacterial species highly pathogenic for a wide variety of herbivores. Anthrax, the associated disease, has had a high economic impact in the pre-vaccines and pre-antibiotics era (1). Despite major therapeutic and control successes during the last century, anthrax is still endemic in many countries due to the long-term survival of *B. anthracis* spores in the environment, difficult to control propagation in wild fauna, export of contaminated animal products and/or insufficient vaccination programs (2–8). The spread of anthrax is thought to be associated with the emergence of agriculture, and the biblical fifth plague of Egypt 3500 years before present (ybp) is usually considered to be anthrax (see (9) for a review and (10, 11) for a reminder on limitations of historical records interpretations regarding infectious diseases).

Twenty years ago, the availability of the first whole genome sequences allowed the development of simple genotyping assays based on tandem repeats polymorphisms (MLVA) or key (canonical) single nucleotide polymorphisms (canSNPs). These PCR-based assays were subsequently applied to thousands of strains collected worldwide (4, 12–15). Large-scale SNP typing demonstrated that the evolution of *B. anthracis* is strictly clonal, implying that robust phylogenies can be drawn from whole genome SNP analysis (wgSNP) (13, 16, 17).

Since 2010, whole genome sequencing (WGS) data from hundreds of *B. anthracis* strains have been made public and thousands of core-genome SNPs have been identified (7, 18–20). The genetic diversity of *B. anthracis* is now well known, a convenient nomenclature based on canSNPs and algorithms for naming lineages have been developed (13, 19, 21).

Despite these impressive achievements in the past twenty years, our understanding of the natural history of *B. anthracis* remains weak in two major respects. Firstly, the age of the *B. anthracis* species or even of the most recent common ancestor (MRCA) of currently identified lineages is essentially unknown. In contrast with other pathogens, such as *Mycobacterium tuberculosis* (22), data analysis illustrated the absence of a reliable molecular clock, which could have allowed such dating (13, 19). The lack of direct temporal signal in the sequence data is likely linked to the capacity of *B. anthracis* to sporulate and enter a quiescent state for years or decades (1), and to the frequency of infection opportunities per year, which may vary among different ecosystems (according to environmental conditions, host availability, and for the last 140 years the impact of vaccination programs). In addition, no *B. anthracis* sequences have been identified so far in ancient DNA investigations. This may reflect a lower epidemiological impact in humans, a lower relative amount of bacterial DNA in investigated remains, or a different spatiotempo-

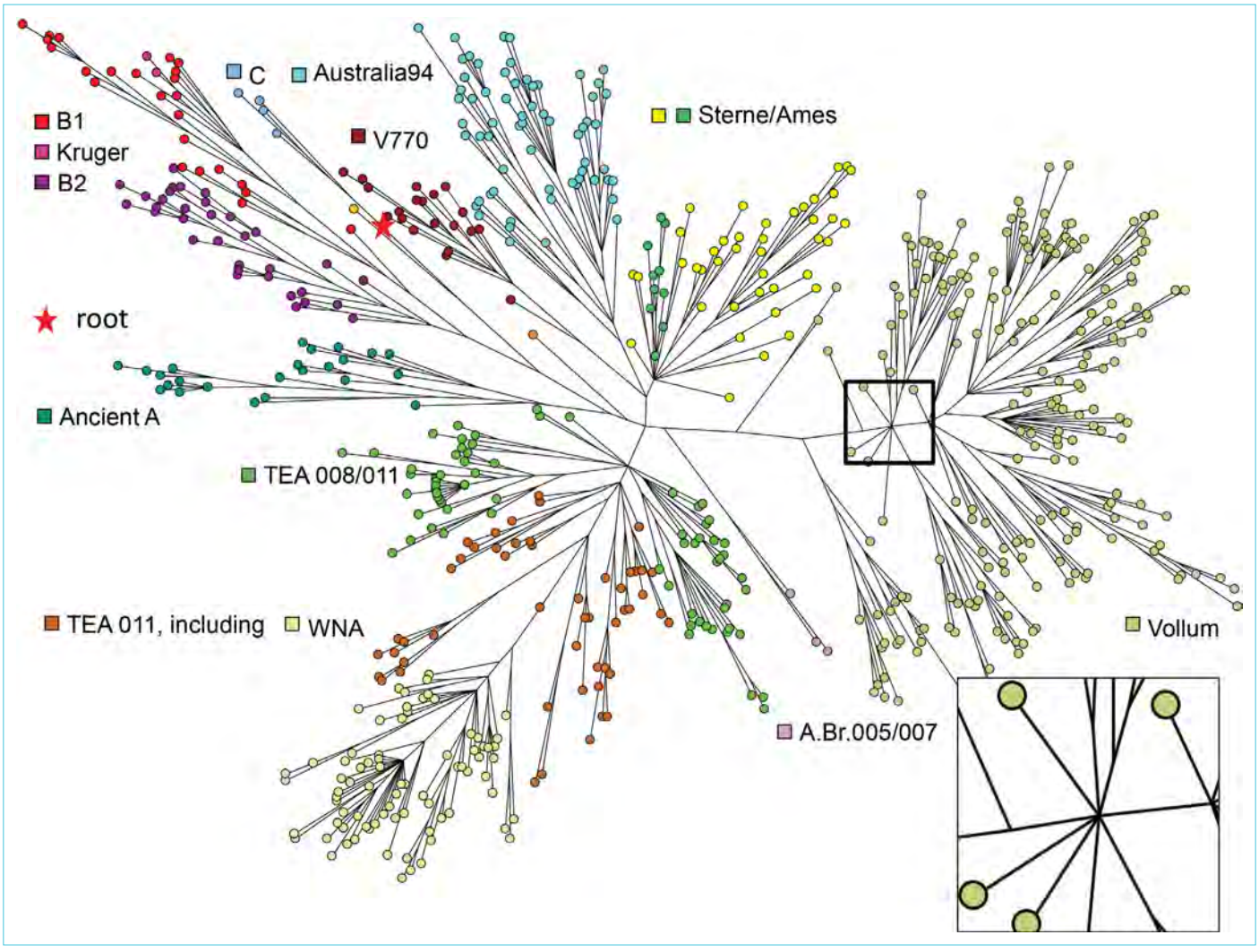

Figure 1. Whole genome SNP analysis of 636 strains. Color code reflects the canSNP assignment (13, 19). A logarithmic scale was used. All WGS data publicly released before December 21st, 2019 (genome assemblies and read archives) were analysed. Duplicates and redundancies (strains with an identical core-genome SNP genotype) are not included in the figure. The predominant A branch comprises "V770", "Australia94", "Sterne/Ames", "Vollum", "A.Br.005/007", "TEA 008/011", "TEA 011", "WNA"

ral distribution regarding *B. anthracis* infections as compared, for instance, to *Yersinia pestis* infections. Secondly, the place of birth of *B. anthracis* itself and most of its sub-lineages is unknown. This is due, at least in part, to the ability of *B. anthracis* to travel long distances in association with animal products (23–25), and to the biased sampling of strains. A significant part of our knowledge of *B. anthracis* genetic diversity results from sequencing efforts in the USA, where many investigated strains are the result of recent importation events and the actual source is often unknown.

In recent years, the analysis of the topology of the *B. anthracis* phylogeny has allowed proposing dating points (7, 20). This review will summarize these suggestions, update the underlying analyses, and propose additional dating hypotheses.

Polytomies, Origin and Fate

Figure 1 shows the global phylogeny of *B. anthracis* deduced from wgSNP identification and maximum parsimony analysis using all currently available non-redundant WGS data. The figure illustrates the existence in the phylogeny of *B. anthracis* of a small number of remarkable topological features, star-like patterns called polytomies. For instance, the "Vollum" clade contains a polytomy with more than six branches (inset in Fig. 1). Figure 2 recalls that polytomies naturally occur during outbreaks. At least three publications reported WGS data from multiple isolates associated with an outbreak: Girault et al. sequenced 53 strains associated with outbreaks in France (26, 27), Ågren et al. (28) sequenced ten strains from an outbreak in Sweden, Braun et al. (29) sequenced 18 strains recovered from two nearby burial pits in Pollino National Park, Southern Italy. In each case, a typical star-like pattern is seen, constituted by strains differing by usually a single SNP from the progenitor genotype at the center of the polytomy. Polytomies appear to be a systematic outcome of outbreaks, but they are rarely maintained in the long term. The apparent paradox is in agreement with the ecotype bacterial species evolution model (30) in which an ecosystem maintains only one lineage. The long-term fixation of additional lineages reflects the establishment in new

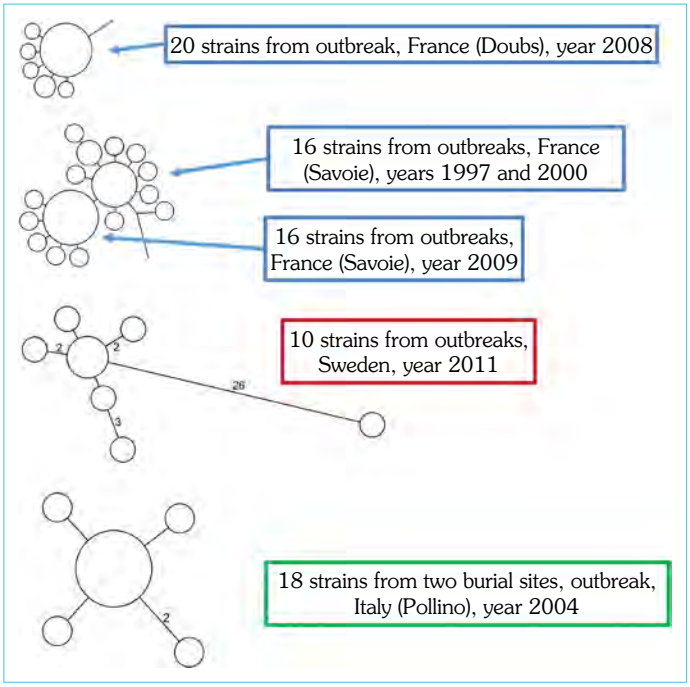

Figure 2. The birth of polytomies. Whole genome SNP analysis based on WGS data from *B. anthracis* outbreaks in Sweden, France and Italy (27–29). Circles' size is proportional to the number of strains sharing the same wgSNP genotype. The smallest circles correspond to one strain. Branch lengths longer than one SNP are indicated. In the Swedish outbreak, 1–2 SNPs per strain are observed. In the French outbreaks, the value is 0.4 SNP per strain (23 SNPs in 53 strains). In the Italian investigation, the ratio is 0.28 SNP per strain. Using the estimates recalled by (28) (a mutation rate of 5.2 10^{-10} mutation/bp/generation, a 5.2 10^{6} bp chromosome size and a 38–42 *B. anthracis* division cycles in an infected cow), an intra-animal heterogeneity of 0.1 SNP per genome would be expected. The relatively high value observed in the Swedish investigation might be due in part to the use of antibiotics. The lower value observed in the Italian investigation might reflect the fact that in this case, the sequenced strains were recovered from only two animals and corresponds to strict intra-animal genetic diversity (45)

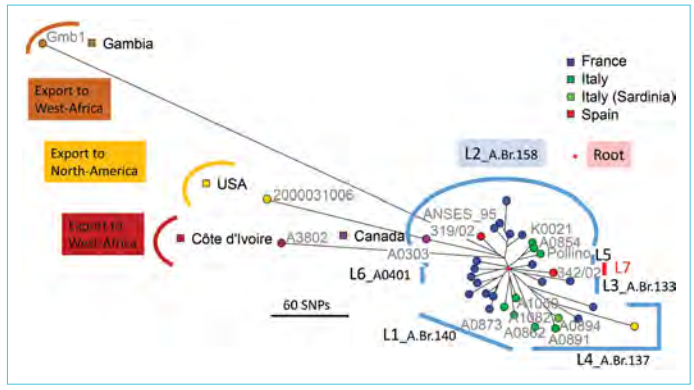

Figure 3. Maximum parsimony tree of the seven branches TEA 011 (A.Br.011/009) polytomy. Thirty-four representative strains were used (last updated December 21st, 2019, genome assemblies and read archives). The name of strains investigated by Sahl et al. (19) or constituting key new branches are indicated. Strain nodes are colored according to the country of isolation. The branch length is proportional to the number of SNPs (see scale). The long branches to North America or West Africa are represented by one or two selected strains. For clarity, short branches in the rest of the tree were pruned as follows: when two strains from the same country were separated by six SNPs or less, only one was kept. The tree is computed from 1551 SNPs. The tree size is 1556 (homoplasia 0.3%). The polytomy currently contains seven branches, numbered L1 to L7. Branch-specific SNPs defined by (19) are indicated alongside lineage number [except for L5 and L7, each represented by a single strain not included in (19)]. The red dot at the center (radiation point) of the polytomy is the root (branching point towards the Ames Ancestor reference strain used as outgroup)

ecosystems, such as potential host populations in new locations. The exceptional spreading events reflected by polytomies might result from historical events in particular those associated with invasive military conflicts, and might provide insights regarding *B. anthracis* phylogeography (7, 20).

Analysing the A.Br.011/009 Polytomy (TEA 011)

The first attempts to date nodes along the *B. anthracis* phylogeny based on molecular data were by van Ert et al. (13). The main parameter was the number of infection/death cycles per year. A "one death per year" average value resulted in a 13,000 ybp estimate for the dating of the MRCA of currently known *B. anthracis* lineages, whereas a "0.5 death per year" value predicted a 27,000 ybp old MRCA. Later on, Kenefic et al. (31) investigated the predominant lineage in Canada and the USA, called Western-North American (WNA). They interpreted (I) the high number of SNPs separating this lineage from its closest neighbors and (II) the evolution of the WNA lineage from north to south as indications of a pre-Columbian introduction via the Bering Strait more than 10,000 ybp. This was a reasonable assumption to interpret WGS data available in 2009, but currently available data made it clear that branch length expansion rates may show considerable variations among even very closely related lineages (19, 20, 27).

The TEA 011 (alias A.Br.011/009) polytomy was first described in 2014, based on the systematic whole genome sequencing of *B. anthracis* strains collected in France (19, 20, 27). Figure 3 shows the TEA 011 polytomy updated by considering currently available WGS data. TEA 011 represents approximately 10% of all sequenced strains. Seven branches numbered L1 to L7 radiate from the MRCA of the polytomy rooted using as an outgroup the "Ames ancestor" strain. The recently described Spanish strain 342/02 defines the new lineage L7 (32). Previously described sub-lineages L1, L2 and L4 [respectively containing the A.Br.140, A.Br.158 and A.Br.137 canSNP branches (19)] are the most frequently observed. In sub-lineage L1, a very early split separates French and Italian strains. One of the two Italian deep branching lineages present in sub-lineage L4 is represented by a strain from Sardinia. Sub-lineage L2 is particularly interesting as it contains three early-splitting sub-lineages, one of which contains secondary

splits to Spain, North-East France, West Africa and North America (Fig. 3). Strain 319/02 from Spain allows us to separate three departures outside of Europe, two towards West Africa (represented by strains from Senegal-Gambia and Côte d'Ivoire respectively) and one towards North America (the WNA lineage), all three characterized by very long branches. This observation, together with the otherwise restricted geographic distribution of strains belonging to TEA 011, provided a very strong argument in favor of post-Columbian contamination of North America from Western Europe instead of pre-Columbian contamination from Asia (20). The proposed interpretation for the emergence of the TEA 011 polytomy is that an outbreak occurred in one of the three European countries contributing founder lineages and that representatives were brought back in the other countries as a side effect of military operations. Southern Italy is the best candidate for being the place of the onset of the polytomy as lineages from the progenitor TEA 008/011 (7) are present in Sicilia (33). Figure 3 also suggests that the topology of TEA 011 reflects two layers of Spanish-French-Italian interactions. The first one is the initial outbreak. The second is illustrated by the more recent lineage L2 split with the Spanish and French branches, and by the split in lineage L4 involving the Sardinian and French lineages.

The three contamination events from West-European sources to West Africa and North America might have included intermediate points, i.e., one of the Atlantic islands colonized by West-European countries about one century before the Americas. However, the pattern of evolution of the WNA lineage from Canada (31) strongly suggests that the North America contamination occurred directly from Western Europe, with 16th or 17th century France as the current best candidate. No TEA 011 strain has been reported so far in other European countries relevant concerning exchanges with the northern part of North America, in particular UK or the Netherlands.

Figure 4 illustrates the rate of expansion of the WNA lineage. From the MRCA indicated by the red star to the tips of the phylogenetic tree, the L2 lineage expanded by approximately 20 SNPs in France, whereas WNA lineages expanded by up to 230 SNPs. This divergence occurred in the same time-span and the most parsimonious explanation for such a behavior is a ten-fold average increase in the number of infection/death cycles per year after arrival in North America. The more basal branches are observed in Canada, in agreement with previous investigations (31). Strain A0303 isolated from cattle is the closest to the ancestral European genotype. The strain is positioned directly on the branch leading to the rest of the WNA lineages. It contributes no specific SNP as if this strain was representing a still extant ancestral genotype.

One interpretation for the pattern shown in Figure 4 is that *B. anthracis* imported from Western Europe, possibly France, contaminated wildlife and spread from Eastern Canada towards Central Canada, where it reached plains bison herds. It was then carried south via bison migrations, as previously proposed (31). The A0303 strain would be a representative of the *B. anthracis* genotype as it was when it first reached plains bisons. Places of past outbreaks would still be contaminated, allowing the reemergence of such ancestral genotypes (2). The success of the subsequent infections would indicate that the bison population was naive and highly susceptible. This recent introduction of anthrax in North America may have been one cause of the massive extinction of bisons (34).

Figure 4 contains a major cluster. The geographic origin of many of the associated strains is not known, but eight among the nine strains with a known location are from Haiti. This cluster defines a polytomy with eight branches. The shortest is defined by three SNPs, the longest by 77. As reviewed in (10, 11), major anthrax outbreaks were observed in Haiti in 1770–1776. It was suggested at that time that the outbreaks resulted from the import of horses from North America. Consequently, Figure 4 further indicates that bisons would have been reached well before the contamination of Haiti, possibly during the 17th century or very early 18th century.

Carrying *B. anthracis* to Canada: the "Régiment de Carignan-Salières" Hypothesis

It is tempting to search for a potential scenario regarding the introduction of the WNA lineage ancestor to Canada and the contamination of bisons. Interestingly, there has been one major military operation between France and Canada during the 16th and 17th centuries, the sending of the Infantry Regiment "Carignan-Salières" to Quebec in the year 1665 (35). The "Régiment de Carignan-Salières" was based in Marsal, North-East France (Lorraine), where all extant French sublineage L2 strains have been collected (27). The "Régiment de Carignan-Salières" resulted from the merging in 1659 of two regiments, (36) previously sent to fight in Northern Italy in 1655–1656 against Spanish troops, many of which were coming from the Napoli area [then under Spanish rule and where a major anthrax outbreak was documented in 1617 (9)]. The infantry regiment crossed France from Marsal to La Rochelle in January-February 1665 and left La Rochelle in April and May 1665. As part of the same global operation aimed at strengthening the French colony in Quebec, twelve horses were transported from La Rochelle in May 1665 (37, 38), suggesting that they were selected among the regiment horses. The horses were the first among approximately 80 sent from France in years 1665–1671 (38).

Although it was not yet a general rule at that time, this regiment was wearing a uniform. Consequently, similar contamination might have been carried by many soldiers if some pieces of equipment (leather, wool, horsehair) were contaminated by *B. anthracis* spores. On the first winter after its arrival in Quebec, the regiment launched a military expedition against Mohawks. Estimates are that the expedition comprised approximately 500 men, tens of them died of exhaustion and hunger, with soldiers from the regiment representing half of the participants (the casualties estimates vary according to accounts, and do not distinguish soldiers and volunteers). The expedition lasted two months from January to March 1666 and reached Schenectady, New York state, 500 kilometers from Quebec (35). Corpses were presumably abandoned along the way back, and some might have subsequently constituted contaminated spots. Consequently, it is tempting to speculate that *B. anthracis* contaminated uniforms have been the seed of wildlife contamination of Eastern Canada in 1666. Another possibility is that horses themselves were carriers of the infection, or were contaminated after their arrival via contact with soldiers. Although not proved impossible, subclinical infection in horses is clearly not established (1), and the possibility of such a long term healthy carriage has never been documented. Contamination after arrival is similarly unlikely as the fate of each horse imported from France was individually recorded, and the introduction of the horses was considered as very successful (38). In addition, the amount of other

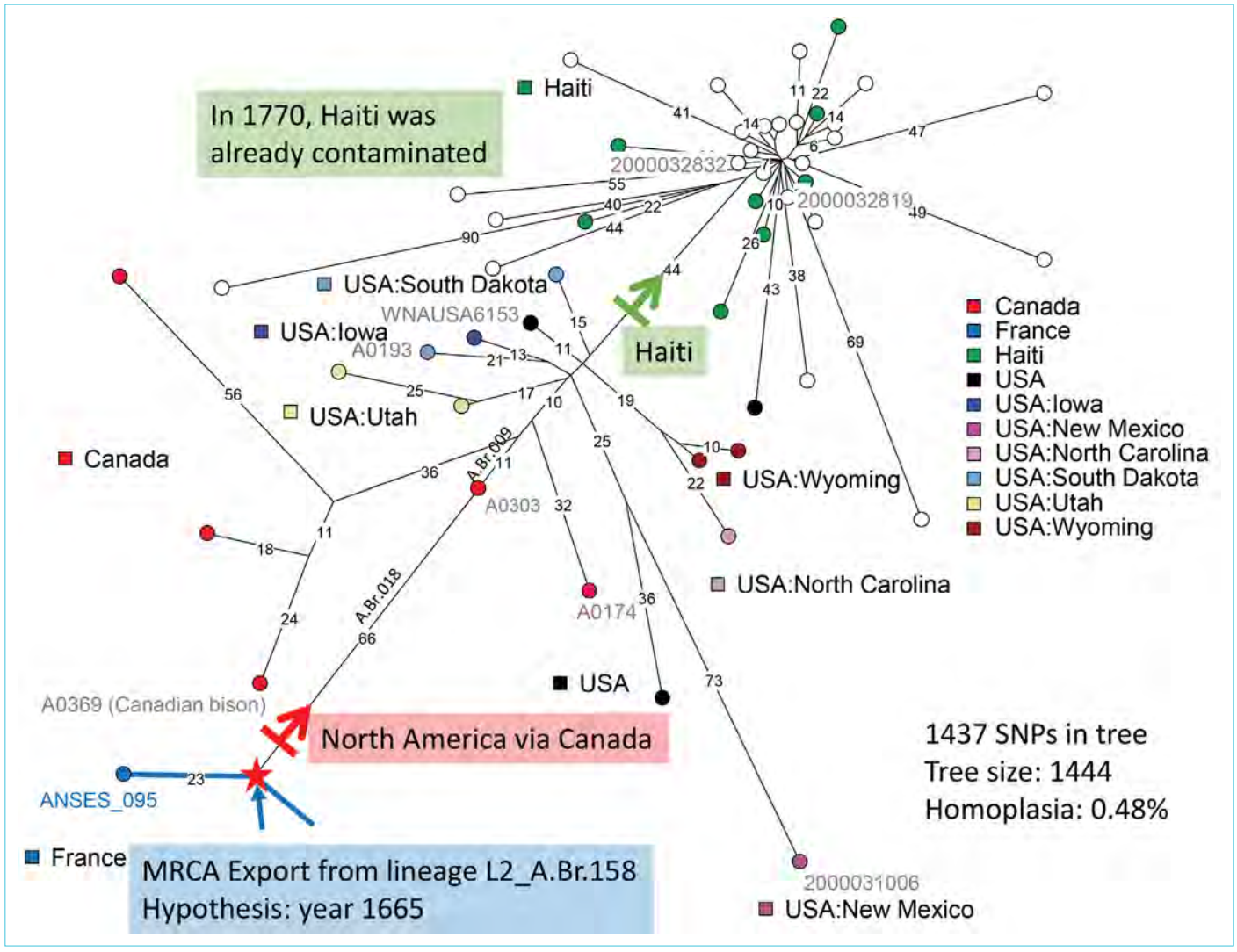

Figure 4. Maximum parsimony analysis of the WNA lineage. Fifty strains were analysed, including one from the A.Br.011/009 polytomy sublineage 2 from which WNA was exported. The color code reflects the geographic origin. Branch length is proportional to the number of SNPs, values of six SNPs or more are indicated. Strain IDs investigated by Sahl et al. (19) are indicated except for some of the strains in the densely populated "Haiti" cluster containing the eight branches polytomy. The red star shows the position of the MRCA of the European progenitor lineage L2_A.Br.158 and the WNA lineage. The green and red arrows illustrate the geographic transitions

livestock present in Quebec at that time would certainly not have been sufficient to fuel the extremely fast expansion of the *B. anthracis* lineage after arrival in North America, illustrated in Figure 4.

Tentatively Dating the Birth of the TEA 011 (A.Br.011/009) Polytomy

Interestingly, when excluding strains not isolated in Europe, or strains which might have been extensively cultivated (strains from public repositories or strains derived from the PasteurII-Carbosap vaccine), branch length heterogeneity within TEA 011 appears to be limited. The shortest lineage in the TEA 011 polytomy is lineage L5 with 18 SNPs and the longest is L4 with up to 65 SNPs (Fig. 3). This suggests that there might be a temporal signal within this subset, reflecting a homogeneous ecosystem in France, Italy and Spain. Figure 5 shows the result of a Bayesian Evolutionary Analysis Sampling Trees (BEAST) software (39) analysis. The reconstituted ancestral genotype (MRCA) as exported towards North America and West Africa was included in the analysis and dated to 1650 under the «Carignan-Salières» hypothesis. The analysis dates the emergence of the TEA 011 polytomy to the year 1450, with a 95% credibility interval of 1220–1620 CE (Fig. 5).

Dating Internal Nodes Along *B. anthracis* Phylogeny

The proposed dating of the TEA 011 polytomy, i.e., only a few centuries back, indicates that *B. anthracis*, like other major pathogens, such as *M. tuberculosis* (22, 40), is quite young, and that its emergence may indeed be associated with the development of agriculture and pastoralism. An African origin for *B. anthracis* might be proposed based on the finding of an African *Bacillus cereus* lineage causing anthrax-like disease in wildlife in African tropical forests (41). *B. cereus* biovar anthracis (BcvA) possesses plasmids (pBCXO1 and pBCXO2) showing sequence similarity higher than

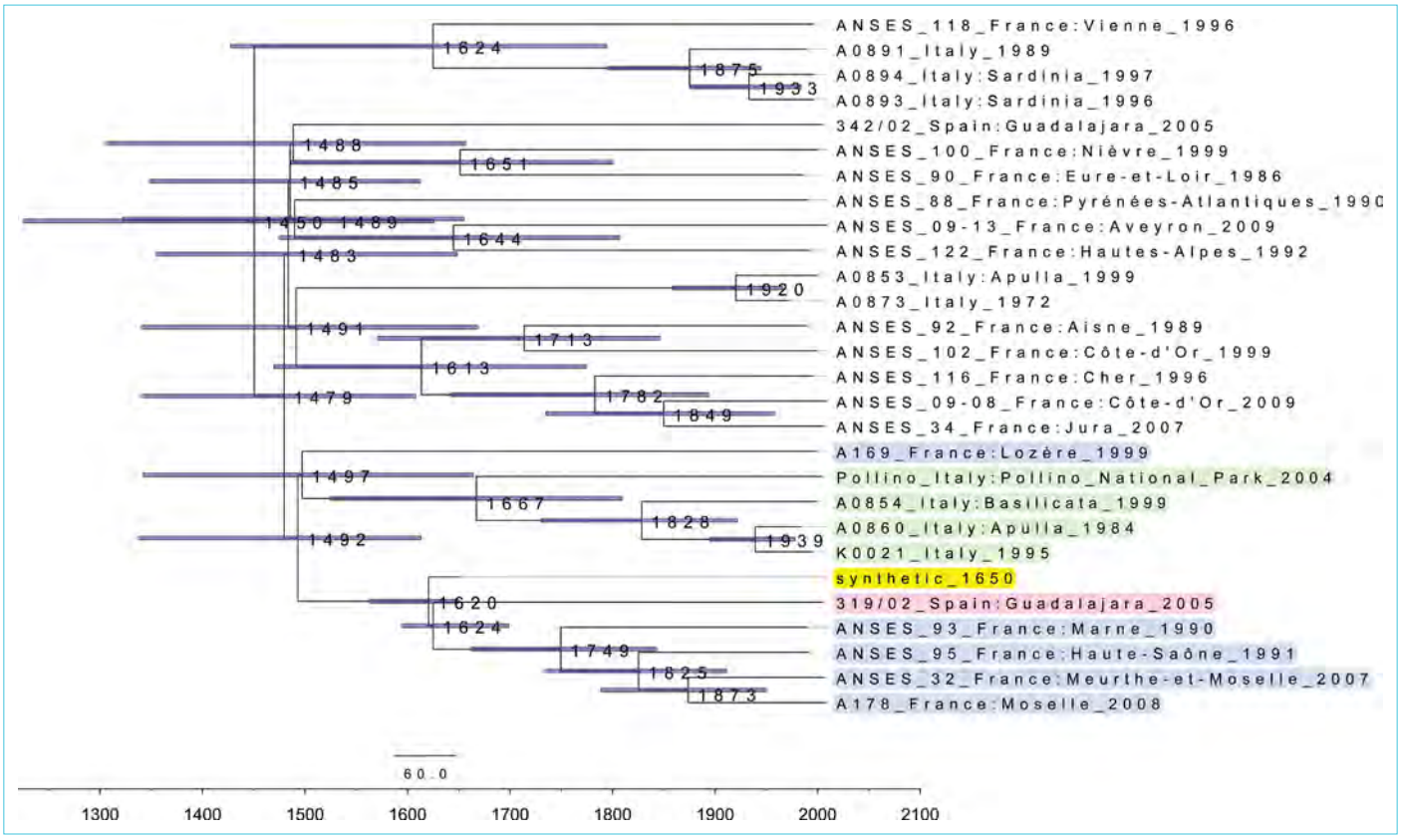

Figure 5. Analysis of wgSNP data from Italian, French and Spanish TEA 011 strains using BEAST. BEAST version 1.10 (39) was run under a general time-reversible model of nucleotide substitution with a gamma distribution between sites, a relaxed molecular clock, Bayesian skyline plot (BSP) demographic model, lognormal distribution for population sizes. Years are indicated in the current era (CE). Bars on nodes indicate the 95% credibility dating interval. Strain ID, place of isolation, and year of isolation are indicated. Strains from the L2 sub-lineage are colored according to the country of origin. The artificial synthetic hypothetical node representing the ancestral genotype at the time (circa 1650) of the export towards West Africa and North America is colored in yellow

99.9% with the *B. anthracis* virulence plasmids pXO1 and pXO2. The BcvA versus *B. anthracis* chromosomal sequence similarity is much lower (95%), implying that the virulence plasmids can occasionally be exchanged by horizontal transfer among different *B. cereus* lineages (41). A different *B. cereus* lineage carrying a pXO1 plasmid has been identified in North America (42). As a result, North America is a place where all three of the main *B. anthracis* lineages, A, B, and C, are represented, as well as *B. cereus* strains causing anthrax-like disease. It could be argued that North America, rather than Africa, is the birthplace of *B. anthracis* virulence plasmids. However, this would imply a very unlikely export from North America to the West-African tropical forest. Instead, a more parsimonious hypothesis might be that these lineages were imported into the United States. Although more information on the *B. cereus* lineages causing anthrax-like disease is required to reach a conclusion, I will hypothesize an African origin for the virulence plasmids in the rest of this review. Phylogenetic analysis showed that the virulence plasmids present in these *B. cereus* lineages are outgroups concerning the plasmids present among extant *B. anthracis* lineages (41). This is as if one among possibly many *B. cereus* strains carrying the virulence plasmids and circulating within the tropical forest had, due to human activities, escaped this ecosystem to give birth to the *B. anthracis* ecotype, reminiscent of the scenario proposed to explain the emergence of *M. tuberculosis* from *Mycobacterium canettii* (40). Since this emergence, the plasmids carried by the founder strain have undergone a clonal evolution strictly congruent with the chromosome. One *B. anthracis* lineage, A.Br.005/006, also called "Ancient A" (19), is a candidate for representing the original ecotype that emerged from the African tropical forest "cradle". This lineage is remarkable by its strong geographic association with Africa (Botswana, Cameroon, Chad, Nigeria, South-Africa, Tanzania, Uganda, Zambia). Among sequenced strains, the shortest branches are constituted by strains from Tanzania, Zambia and Uganda. No other *B. anthracis* lineage was reported yet from these countries. They are adjacent to the Democratic Republic of Congo and Central African Republic from which *B. cereus* biovar anthracis strains were recovered and in which the genetic diversity of *B. anthracis* has not yet been reported.

Figure 6 is a tentative representation of the evolution of *B. anthracis* using the ecotype speciation model (30). In this representation, a local ecotype ("Ancient A" for Ancient anthracis) is descending from a *B. cereus* ancestor carrying pXO1 and pXO2. At some points along its evolution, the ecotype would have caused anthrax outbreaks, which spread out of Africa, giving birth to new geotypes. In the current topology of *B. anthracis* popula-

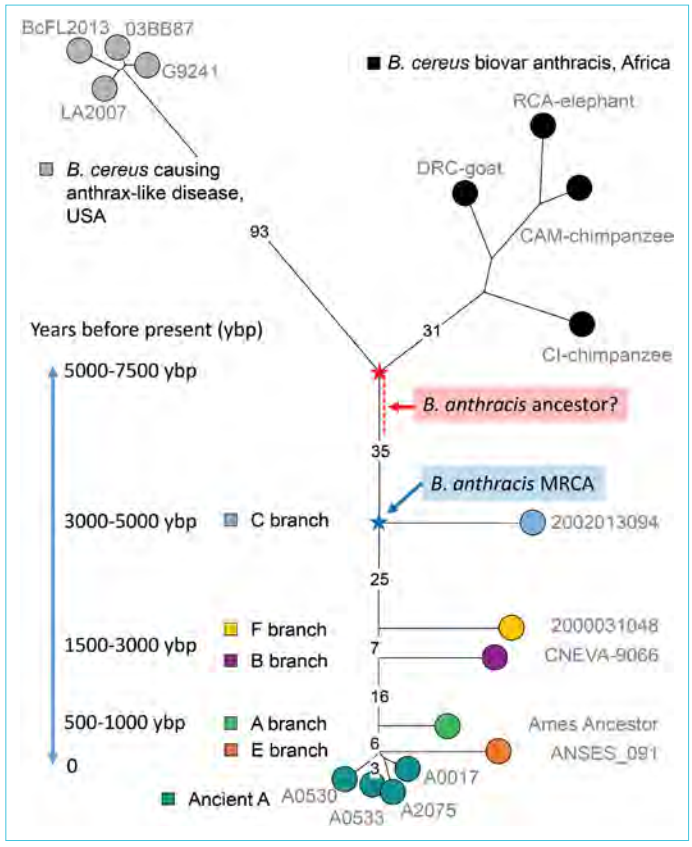

Figure 6. Out-of-Africa time-points along with *B. anthracis* evolution. Maximum parsimony analysis based on SNPs identified in the pXO1 plasmid present in *B. cereus* causing anthrax-like disease and in representative strains from the main *B. anthracis* clades. The new F branch is defined by a unique strain. While the position of the *B. anthracis* MRCA is clearly defined, the position of the *B. anthracis* species ancestor is uncertain, partly because of our currently very limited knowledge of pXO1 plasmids hosted by *B. cereus* lineages. The estimate for the emergence of each branch reflects the uncertainty of the position of the *B. anthracis* ancestor

tion, remnants of four such departures are visible along the line going from the *B. anthracis* MRCA to Ancient A. They are represented by (starting from the earliest departure) lineage C, lineage B, and the most recent lineage A. The fourth departure, represented by a unique strain, is detected immediately preceding the lineage B departure. I propose to call this lineage "F" in keeping with the previous lineages D and E designations (43). Lineage E, also called Aβ (44), was observed in Cameroon and Chad and one Cameroon strain was sequenced. Both D and E belong to A.Br.005/006 (13). Figure 6 is also proposing a timescale, under the approximate assumptions of emergence of *B. anthracis* 5,000 ybp, corresponding to the current estimates for the emergence of agriculture in Central Africa, and of a regular rate of expansion of the local ecotype Ancient A. This estimate would make the C lineage a candidate for being a remnant of the fifth plague of Egypt approximately 3500 ybp (9). Egypt would be central in these "Out-of-Africa" events, as the Nile itself and its floods could have directly carried contaminations that arised from outbreaks in Central Africa.

CONCLUSION

This review is proposing a global hypothetical scheme for the phylogeography of *B. anthracis* and its place of origin. This scheme will need to be challenged by the collection of new sequence data from many countries worldwide, including in particular African and Mediterranean countries.

Acknowledgements: I wish to thank Pr Les Baillie and Dr Antonio Fasanella for organizing the "Biology of Anthrax" meeting in Bari, Italy 2019 and inviting me to present ongoing work on the dating of *B. anthracis*. I particularly thank Pr Mehmet Doganay for offering to write a review on this topic for the Erciyes Medical Journal. This writing provided the opportunity to explore further the scenario presented during the conference. The present review contains significant modifications and major additions in terms of dating. I thank Pr Martin Hugh-Jones and Dr Robert Gainer, met during the meeting, for subsequent numerous exchanges regarding healthy carriage and latent infections issues. I thank Christine Pourcel for her critical reading of a near to final version. Due to space limitations, I apologize for having to omit a number of valuable references. These analyses are made possible owing to the public availability of whole genome sequence data produced by many laboratories worldwide.

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author have no conflict of interest to declare.

Financial Disclosure: The author declared that this study has received no financial support.

REFERENCES

1. WHO. Anthrax in humans and animals. 4 ed. Geneva (CH): World Health Organization Press; 2008.
2. Carlson CJ, Getz WM, Kausrud KL, Cizauskas CA, Blackburn JK, Bustos Carrillo FA, et al. Spores and soil from six sides: interdisciplinarity and the environmental biology of anthrax (*Bacillus anthracis*). Biol Rev Camb Philos Soc 2018; 93(4): 1813–31. [CrossRef]
3. Carlson CJ, Kracalik IT, Ross N, Alexander KA, Hugh-Jones ME, 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. [CrossRef]
4. Beyer W, Bellan S, Eberle G, Ganz HH, Getz WM, Haumacher R, et al. Distribution and molecular evolution of *Bacillus anthracis* genotypes in Namibia. PLoS Negl Trop Dis 2012; 6(3): e1534. [CrossRef]
5. Turner WC, Kausrud KL, Beyer W, Easterday WR, Barandongo ZR, Blaschke E, et al. Lethal exposure: An integrated approach to pathogen transmission via environmental reservoirs. Sci Rep 2016; 6: 27311.
6. Hugh-Jones M, Blackburn J. The ecology of *Bacillus anthracis*. Mol Aspects Med 2009; 30(6): 356–67. [CrossRef]
7. Timofeev V, Bahtejeva I, Mironova R, Titareva G, Lev I, Christiany D, et al. Insights from *Bacillus anthracis* strains isolated from permafrost in the tundra zone of Russia. PLoS One 2019; 14(5): e0209140. [CrossRef]
8. Pilo P, Frey J. Pathogenicity, population genetics and dissemination of *Bacillus anthracis*. Infect Genet Evol 2018; 64: 115–25. [CrossRef]
9. Turnbull PCB, Shadomy SV. Anthrax from 5000 BC to AD 2010. In: Bergman NH, editor. *Bacillus anthracis* and Anthrax. John Wiley & Sons Inc; 2010.p.1–15. [CrossRef]
10. Morens DM. Characterizing a "new" disease: epizootic and epidemic anthrax, 1769-1780. Am J Public Health 2003; 93(6): 886–93.
11. Morens DM. Epidemic anthrax in the eighteenth century, the Americas. Emerg Infect Dis 2002; 8(10): 1160–2. [CrossRef]
12. Keim P, Price LB, Klevytska AM, Smith KL, Schupp JM, Okinaka R,

et al. Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J Bacteriol 2000; 182(10): 2928–36. [CrossRef]
13. Van Ert MN, Easterday WR, Huynh LY, Okinaka RT, Hugh-Jones ME, Ravel J, et al. Global genetic population structure of *Bacillus anthracis*. PLoS One 2007; 2(5): e461. [CrossRef]
14. Thierry S, Tourterel C, Le Flèche P, Derzelle S, Dekhil N, Mendy C, et al. Genotyping of French *Bacillus anthracis* strains based on 31-loci multi locus VNTR analysis: epidemiology, marker evaluation, and update of the internet genotype database. PLoS One 2014; 9(6): e95131. [CrossRef]
15. Le Flèche P, Jacques I, Grayon M, Al Dahouk S, Bouchon P, Denoeud F, et al. Evaluation and selection of tandem repeat loci for a *Brucella* MLVA typing assay. BMC Microbiol 2006; 6: 9. [CrossRef]
16. Pearson T, Busch JD, Ravel J, Read TD, Rhoton SD, U'Ren JM, et al. Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. Proc Natl Acad Sci U S A 2004; 101(37): 13536–41. [CrossRef]
17. Simonson TS, Okinaka RT, Wang B, Easterday WR, Huynh L, U'Ren JM, et al. *Bacillus anthracis* in China and its relationship to worldwide lineages. BMC Microbiol 2009; 9: 71. [CrossRef]
18. Pearson T, Okinaka RT, Foster JT, Keim P. Phylogenetic understanding of clonal populations in an era of whole genome sequencing. Infect Genet Evol 2009; 9(5): 1010–9. [CrossRef]
19. Sahl JW, Pearson T, Okinaka R, Schupp JM, Gillece JD, Heaton H, et al. A *Bacillus anthracis* Genome Sequence from the Sverdlovsk 1979 Autopsy Specimens. mBio 2016; 7(5): e01501–16. [CrossRef]
20. Vergnaud G, Girault G, Thierry S, Pourcel C, Madani N, Blouin Y. Comparison of French and Worldwide *Bacillus anthracis* Strains Favors a Recent, Post-Columbian Origin of the Predominant North-American Clade. PLoS One 2016; 11(2): e0146216. [CrossRef]
21. Lärkeryd A, Myrtennäs K, Karlsson E, Dwibedi CK, Forsman M, Larsson P, et al. CanSNPer: a hierarchical genotype classifier of clonal pathogens. Bioinformatics 2014; 30(12): 1762–4. [CrossRef]
22. O'Neill MB, Shockey A, Zarley A, Aylward W, Eldholm V, Kitchen A, et al. Lineage specific histories of *Mycobacterium tuberculosis* dispersal in Africa and Eurasia. Version 2. Mol Ecol 2019; 28(13): 3241–56.
23. Keim P, Grunow R, Vipond R, Grass G, Hoffmaster A, Birdsell DN, et al. Whole Genome Analysis of Injectional Anthrax Identifies Two Disease Clusters Spanning More Than 13 Years. EBioMedicine 2015; 2(11): 1613–8. [CrossRef]
24. Pullan ST, Pearson TR, Latham J, Mason J, Atkinson B, Silman NJ, et al. Whole-genome sequencing investigation of animal-skin-drum-associated UK anthrax cases reveals evidence of mixed populations and relatedness to a US case. Microb Genom 2015; 1(5): e000039.
25. Szablewski CM, Hendricks K, Bower WA, Shadomy SV, Hupert N. Anthrax Cases Associated with Animal-Hair Shaving Brushes. Emerg Infect Dis 2017; 23(5): 806–8. [CrossRef]
26. Madani N, Mendy C, Moutou F, Garin-Bastuji B. La fièvre charbonneuse en France. Episodes de l'été 2009 et foyers enregistrés sur la dernière décennie (1999-2009). [Article in French]. Bulletin épidémiologique 2010; 38(Spécial zoonoses): 17–9.
27. Girault G, Blouin Y, Vergnaud G, Derzelle S. High-throughput sequencing of *Bacillus anthracis* in France: investigating genome diversity and population structure using whole-genome SNP discovery. BMC Genomics 2014; 15: 288. [CrossRef]
28. Ågren J, Finn M, Bengtsson B, Segerman B. Microevolution during an Anthrax outbreak leading to clonal heterogeneity and penicillin resistance. PLoS One 2014; 9(2): e89112. [CrossRef]
29. Braun P, Grass G, Aceti A, Serrecchia L, Affuso A, Marino L, et al. Microevolution of Anthrax from a Young Ancestor (M.A.Y.A.) Suggests a Soil-Borne Life Cycle of *Bacillus anthracis*. PLoS One 2015; 10(8): e0135346. [CrossRef]
30. Cohan FM. Transmission in the Origins of Bacterial Diversity, From Ecotypes to Phyla. Microbiol Spectr 2017; 5(5): MTBP-0014-2016.
31. Kenefic LJ, Pearson T, Okinaka RT, Schupp JM, Wagner DM, Hoffmaster AR, et al. Pre-Columbian origins for North American anthrax. PLoS One 2009; 4(3): e4813. [CrossRef]
32. Antwerpen M, Beyer W, Bassy O, Ortega-García MV, Cabria-Ramos JC, Grass G, et al. Phylogenetic Placement of Isolates Within the Trans-Eurasian Clade A.Br.008/009 of *Bacillus anthracis*. Microorganisms 2019; 7(12): 689. [CrossRef]
33. Rondinone V, Serrecchia L, Parisi A, Fasanella A, Manzulli V, Cipolletta D, et al. Genetic characterization of *Bacillus anthracis* strains circulating in Italy from 1972 to 2018. PLoS One 2020; 15(1): e0227875.
34. Holt SDS. Reinterpreting the 1882 Bison population collapse. Rangelands 2018; 40(4): 106–14. [CrossRef]
35. Verney J. The Good Regiment: The Carignan-Salieres Regiment in Canada, 1665-1668. McGill-Queen's University Press, 1991.p.240.
36. Quillivic B, Nicol J. Le régiment de Carignan. Available from: URL: http://www.migrations.fr/. Accessed September 20, 2019.
37. Campeau CV. Navires venus en Nouvelle-France, gens de mer et passagers, des origines à 1699. Available from: URL: https://www.naviresnouvellefrance.net/. Accessed September 25, 2019.
38. Seguin RL. Le cheval et ses implications historiques dans l'Amérique française. [Article in French]. Revue d'histoire de l'Amérique française 1951; 5(2): 227–51. [CrossRef]
39. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 2007; 7: 214. [CrossRef]
40. Blouin Y. A new scenario for the early evolution of *Mycobacterium tuberculosis*. Université Paris Sud - Paris XI, 2014.
41. Antonation KS, Grützmacher K, Dupke S, Mabon P, Zimmermann F, Lankester F, et al. *Bacillus cereus* Biovar Anthracis Causing Anthrax in Sub-Saharan Africa-Chromosomal Monophyly and Broad Geographic Distribution. PLoS Negl Trop Dis 2016; 10(9): e0004923. [CrossRef]
42. Marston CK, Ibrahim H, Lee P, Churchwell G, Gumke M, Stanek D, et al. Anthrax Toxin-Expressing *Bacillus cereus* Isolated from an Anthrax-Like Eschar. PLoS One 2016; 11(6): e0156987. [CrossRef]
43. Lista F, Faggioni G, Valjevac S, Ciammaruconi A, Vaissaire J, le Doujet C, 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. [CrossRef]
44. Maho A, Rossano A, Hächler H, Holzer A, Schelling E, Zinsstag J, et al. Antibiotic susceptibility and molecular diversity of *Bacillus anthracis* strains in Chad: detection of a new phylogenetic subgroup. J Clin Microbiol 2006; 44(9): 3422–5. [CrossRef]
45. Easterday WR, Ponciano JM, Gomez JP, Van Ert MN, Hadfield T, Bagamian K, et al. Coalescence modeling of intrainfection *Bacillus anthracis* populations allows estimation of infection parameters in wild populations. Proc Natl Acad Sci U S A 2020; 117(8): 4273–80. [CrossRef]