Molecular Characterization of an Outbreak-involved Bacillus anthracis Strain Confirms the Spillover of Anthrax from West Africa

Shuchao Wang (wsc1026@126.com)
Chinese Academy of Agricultural Sciences Changchun Veterinary Research Institute
https://orcid.org/0000-0001-9782-9656

Roland Suluku
School of agriculture and food sciences, Njala University

Mohamed B. Jalloh
University of Sierra Leone College of Medicine and Allied Health Sciences

Ahmed F. Samba
Ministry of Agriculture, Sierra Leone

Baogui Jiang
Beijing institute of microbiology and epidemiology

Yubiao Xie
Chinese Academy of Agricultural Sciences Changchun Veterinary Research Institute

Doris Harding
Ministry of Health and Sanitation, Sierra Leone

Mengyao Zhang
Beijing institute of biotechnology

Foday Sahr
University of Sierra Leone College of Medicine and Allied Health Sciences

Mahmud E. Sesay
School of Agriculture and Food Sciences, Njala University

James S. Squire
Ministry of Health and Sanitation, Sierra Leone

Mohamed A. Vandi
Ministry of Health and Sanitation, Sierra Leone

Moinina N. Kallon
School of Agricultural sciences, Njala University

Shoufeng Zhang
Chinese Academy of Agricultural Sciences Changchun Veterinary Research Institute

Rongliang Hu
Chinese Academy of Agricultural Sciences Changchun Veterinary Research Institute
Abstract

Background

Anthrax, a zoonotic disease caused by the spore-forming bacterium *Bacillus anthracis*, remains a major global public health concern, especially in countries with limited resources. Sierra Leone, a West African country historically plagued by anthrax, has almost been out of report on this disease in recent decades, possibly due to limitations in diagnostic capabilities. In this study, we described a large-scale anthrax outbreak affecting both animals and humans and attempted to characterize the pathogen using molecular tools.

Methods

The causative agent of the animal outbreak was identified using the nanopore sequencing technique. Confirmation of suspected cases of cutaneous anthrax in humans, identified through active surveillance, was conducted using qPCR targeting both the chromosome and two virulent plasmids. Full-genome sequencing was accomplished by combining long-read and short-read sequencing methods. Subsequent phylogenetic analysis was performed based on the full-genome single nucleotide polymorphisms.

Results

The outbreak in Port Loko District, Sierra Leone, led to the death of 233 animals between March 26th and May 16th, 2022. We ruled out the initial suspicion of *Anaplasma* species and successfully identified *B. anthracis* as the causative agent of the outbreak. As a result of the government’s prompt response, out of the 49 suspected human cases identified during the one-year active surveillance, only six human cases, all within the first month after the official declaration of the outbreak, were confirmed. The phylogenetic analysis indicated that the BaSL2022 isolate responsible for the outbreak was positioned in the A.Br.153 clade within the TransEuroAsian group of *B. anthracis*.

Conclusions

The anthrax outbreak identified in this study was unprecedented in scale in Sierra Leone. With regard to the shared predicament, the utilization of nanopore sequencing and molecular techniques to identify unknown diseases in Sierra Leone has provided valuable insights into emergency preparedness in similar nations. The A.Br.153 clade, where BaSL2022 was phylogenetically placed, is distinct from the existing West African lineages but may have been prevalent in a large area of West Africa for a long period and undergone evolutionary changes to adapt to novel geographical environments, as evidenced by the spillover of this strain to the United Kingdom and United States.
1. Background

Anthrax is a zoonotic disease with a significant historical background. Despite being a long-standing issue, it remains a prominent global public health concern, particularly in resource-limited regions. The disease, affecting both humans and animals, is primarily caused by *Bacillus anthracis*, a gram-positive, spore-forming bacterium. Infection with *B. anthracis* typically results in cutaneous, gastrointestinal, inhalational, or injectional anthrax. Cutaneous anthrax is the prevailing form in humans, accounting for over 95% of reported human cases. This form of anthrax typically has a comparatively lower mortality rate, whereas gastrointestinal and inhalational forms of anthrax, although less prevalent, are characterized by a significantly higher lethality rate and are considered the primary cause of most anthrax-related fatalities in humans [1, 2]. Unlike humans, both wild herbivores and livestock predominantly encounter acute and lethal consequences as a result of gastrointestinal exposure to spores while grazing [3, 4]. Livestock anthrax frequently serves as a source of secondary human infections, mainly due to the handling and consumption of animal products that are inadvertently contaminated [5, 6]. According to available data, the global incidence of anthrax cases in humans is estimated to range between 20,000 and 100,000 per year, with approximately 1.83 billion individuals living in areas at risk of anthrax infection [2]. According to estimates, approximately 1.1 billion livestock are susceptible to the disease, mainly in rural rainfed systems, with a particular emphasis on sub-Saharan Africa [3].

Africa is one of the most threatened continents, facing the challenge of anthrax outbreaks. These outbreaks have been extensively documented in multiple countries, affecting wildlife, domestic animals, and humans alike [3, 5, 6, 7, 8]. However, it is believed that the actual number of cases is considerably underestimated due to inadequate diagnostic capabilities [8, 9]. Based on modeling techniques, previous studies have identified several regions in eastern, central, and southern Africa, as well as an east-west corridor spanning from Ethiopia to Sierra Leone in the Sahel region, as having favorable conditions for the spread of *B. anthracis* [8, 10]. However, thus far, only Botswana, Zambia, Uganda, Tanzania, Namibia, South Africa, Nigeria, Cameroon, Chad, Ghana, and Ethiopia have documented verified instances of anthrax through the utilization of molecular and/or bacteriological techniques [10, 11, 12]. In Sierra Leone, sporadic instances of suspected anthrax can be identified in the annual reports of the Ministry of Agriculture between 1947 and 1994. The latest recorded outbreak data date back to 1994, during which 8 individuals lost their lives as a result of consuming deceased goat meat sourced from Guinea (Personal interview). Recently, there have been four reported cases of suspected cutaneous anthrax in human individuals within the vicinity of Kamakwie, Sierra Leone [13]. Regrettably, none of these records have been supported by dependable molecular or validated bacteriological methods. The absence of molecular verification significantly impedes the utility of these outbreak records as a reliable genetic resource for disease control.
B. anthracis is characterized by a relatively monomorphic nature, exhibiting a highly conserved genome composition. This implies that the genetic relatedness between descendant strains can be readily ascertained through the utilization of genomic molecular techniques, such as multilocus variable-number-of-tandem-repeats analysis (MLVA) and canonical single nucleotide polymorphisms (SNPs) [14, 15]. According to the analyses conducted using MLVA and the 12 canonical SNPs, B. anthracis exhibited a phylogenetic division into three primary branches, namely, A, B, and C. These branches were further classified into several sublineages, each characterized by distinct geographical distributions [14, 16]. With the emergence of cost-effective whole-genome sequencing methods, the field of phylogenetic reconstruction has witnessed a significant increase in the utilization of genome-wide SNPs as well as core-genome-based multilocus sequence typing (cgMLST) [16, 17, 18, 19]. These methods offer improved strain clustering resolution, rendering them a potent tool for tracing outbreaks and modeling anthrax transmission [20, 21, 22]. Among the three primary phylogenetic branches of B. anthracis, isolates belonging to the B and C branches are exceedingly uncommon and geographically limited, often found only in specific locations or archival sources. In contrast, the A branch of B. anthracis exhibits a global distribution and accounts for more than 90% of outbreaks [14]. Within the A branch, the majority of the sublineages demonstrate a distinct geographical distribution. However, the TransEuroasian (TEA) sublineage has achieved significant global dissemination, possibly due to a rapid clonal radiation event within a short timeframe. This is evident from the remarkably short phylogenetic branches [21, 22]. In the West African region, there exists a limited number of publicly accessible complete genomes of B. anthracis, which were all isolated in Senegal and Gambia [23]. All these isolates belong to the TEA sublineage and exclusively form a specific subclade that has been recently designated A.Br.148, according to the nomenclature of the full genome SNP phylogeny system [19, 22]. Of note, despite the constraints of public genomic resources, West Africa bears a significant burden of anthrax caused by B. anthracis lineages that are primarily prevalent in this specific region [6, 9, 24]. A notable illustration is the exclusive presence of the specific lineage A.Br.148 in West Africa. This B. anthracis lineage, coupled with a distinct lineage found in Cameroon, Chad, Mali, and Nigeria that lacks spore-surface-associated arthrose [12, 25, 26], as well as the nontypical B. anthracis strains isolated from wild apes in Côte d'Ivoire and Cameroon [27, 28], provides a glimpse into the complex nature of anthrax epidemiology in West Africa.

In this study, we provide a detailed account of an anthrax outbreak in Sierra Leone that exhibited an unprecedented magnitude. This outbreak led to the unfortunate demise of more than 200 animals and caused infections in six individuals. The analysis of the full-genome sequence indicated that the strain responsible for the outbreak was distinct from all existing West African lineages but may be associated with the anthrax outbreaks in London, Scotland, and New York between 2006 and 2008.

2. Methods

2.1 Rapid identification of the causative agent responsible for the outbreak
All experimental procedures were conducted within a designated A2-type biological safety cabinet situated in a biosafety level 3 (BSL-3) facility. The liver sample underwent two rounds of rinsing with 70% ethanol and was subsequently divided into smaller sections at the Tropical Infectious Diseases Prevention and Control Center of Sierra Leone (IDPC). The materials underwent a washing process using 1 mL of distilled water, and 200 µl of the supernatant was utilized for the extraction of DNA/RNA using the TGuid S32 magnetic DNA/RNA isolation kit (Tiangen, Beijing) following the manufacturer's instructions. The nucleic acid was sterilized at 95 °C for 20 minutes. The quality and concentration of the nucleic acid were determined using a Qubit 4 fluorometer (Invitrogen). A nanopore sequencing library was prepared by utilizing the eligible DNA sample and the Rapid Barcoding and Sequencing kits provided by Oxford Nanopore Technologies. The library was subsequently subjected to sequencing on a MinION Mk1B device using an R9.4.1 flow cell (Oxford Nanopore Technologies). During the one-hour sequencing process, the acquisition of raw data, base calling, and demultiplexing were carried out in real time using MinKNOW software associated with the device, utilizing a "superaccuracy" model. A swift annotation was then performed using the software Centrifuge (release 1.0.3) with the "Bacteria, Archaea, Viruses, Human (compressed)" database [29].

2.2 Confirmation of *B. anthracis* through quantitative PCR (qPCR) and cultivation

Tissue samples were processed according to the aforementioned procedure, with a minor adjustment. A homogenization process was performed on each tissue sample, where 200 mg of the sample was mixed with 1 mL of normal saline solution. A volume of 200 µl of the supernatant obtained from the tissue homogenate or blood sample was used for DNA extraction. The analysis of *B. anthracis* DNA was conducted through qPCR to examine the presence of chromosomal marker *dhp61* (BA_5345) as well as plasmid markers *pagA* (pXO1) and *capC* (pXO2) using the primers as previously described [30, 31].

For cultivation, the liver tissue sample was subjected to two rounds of rinsing with 70% ethanol and distilled water and subsequently partitioned into two distinct sections. The cutting plane was applied onto the nutrient agar, which was incubated at a temperature of 37 °C for 24 hours. Colonies were subsequently purified using the streaking method. One colony was selected to prepare a slide for Gram staining.

2.3 Full-genome sequencing

Cells from the culture were harvested and digested with lysosome at 20 mg/ml and RNase A at 2 mg/ml for 1 hour, followed by genomic DNA extraction using the TGuide Bacteria DNA Kit (Tiangen, Beijing). Sterilization of the resulting DNA was performed using 0.22-µm filters, which was further confirmed by culturing on nutrient agar. Eligible DNA with no positive colonies on the plate within 72 hours was subsequently quantified using the Qubit 4 fluorometer and transferred out of the BSL-3 lab. Full-genome sequencing was conducted by a combination of short-read sequencing and long-read sequencing techniques [32, 33]. For long-read sequencing, a nanopore sequencing library was prepared employing the Ligation Sequencing Kit and the Native Barcoding Kit (Oxford Nanopore Technologies). The library was
sequenced on a MinION Mk1B device for 24 hours using an R9.4.1 flow cell (Oxford Nanopore Technologies) following the manufacturer’s manuals. Base calling, demultiplexing, and de novo assembly of the long-read sequencing data were performed using the corresponding pipelines and tools in EPI2ME application v5.1.1 (Oxford Nanopore Technologies).

Short-read sequencing was performed on an Illumina MiSeq platform using the Nextera XT DNA Library Preparation Kit and the MiSeq Reagent Kit v3 (2×300 bp) chemistry (Illumina). High-quality paired-end reads were subjected to de novo assembly into contigs in SPAades 3.11.1 [34] prior to further refining by SNP and indel correction using SAMtools 1.7 and Pilon 1.22 [35, 36]. The assembly and polishing of the combined long reads and short reads data were ultimately conducted using MicroPIPE v0.9 [37]. The primary data generated in this study were deposited in the NCBI Sequence Read Archive (SRA) repository under the BioProject number PRJNA875505.

2.4 Full-genome SNP calling and phylogenetic construction

To determine the phylogenetic position of the *B. anthracis* strain isolated in this outbreak, we retrieved 236 representative full genome sequences that were used in the initial global phylogenetic reconstruction [22] and the most recent molecular typing practice based on full-genome SNPs of *B. anthracis* [32]. The Parsnp tool from the Harvest Suite [38] was then used for the core chromosome multiple alignment with the ‘Ames Ancestor’ (NC_007530.2) stain serving as the reference genome. HarvestTools 1.2 from the same suite was employed to identify the SNP positions and export the concatenated SNP sequences, as described previously [32, 39]. Phylogenetic trees were constructed using RAxML v8.2.12 with the maximum likelihood model and annotated using the online iTol platform [40].

3. Results

3.1 Identification and confirmation of the anthrax outbreak in Sierra Leone

The salient aspects of the outbreak are illustrated in Fig. 1. On March 26th, 2022, herbivorous livestock in Thinka Barreh village (8°48′32″N, 12°55′40″W) of the Bakeloko Chiefdom, located in the Port Loko District of Sierra Leone, were reported to be afflicted by an unidentified disease. The observed symptoms encompassed manifestations such as weakness, depression, shivering, mortality, and postmortem hemorrhage from the oral and nasal cavities (Fig. 2A). Local veterinarians hypothesized that the outbreak could potentially be attributed to a severe infection caused by *Anaplasma* species, given the observation of leukemia cases in cattle on the same farm. On April 11th, samples of tissue from the spleen and liver of a deceased bovine, as well as blood samples from eight live bovines, were submitted to the IDPC for verification of the initial suspicion. In the BSL-3 lab, PCR-based screening was conducted on the samples to identify the presence of the *Anaplasma* species [41]. The screening revealed that all samples yielded negative results, thereby refuting the initial hypothesis.
On May 3rd, the Department of Animal Sciences of Njala University led a team to the outbreak area in attempt to identify the etiology of the ongoing outbreak, as the disease's rapid dissemination resulted in substantial livestock mortality, leading to significant economic losses for farmers. On May 4th, a comprehensive collection of clinical samples was conducted, comprising 40 blood samples from live cattle and 6 blood samples from live sheep, as well as tissue samples from the heart, liver, spleen, lung, and kidney of a newly deceased sheep. These samples were collected from two neighboring farms, Chernor Bah and Amadu Wurie Bah, where the grazing areas overlapped with the Affina Jalloh farm. After collection, the samples were promptly dispatched to the IDPC in a cold chain. We conducted a prompt identification of the potent causative agent utilizing a random long-read sequencing technique, which has been demonstrated to be efficacious in emergency preparedness [17]. The results revealed a significant presence of anthrax infection in the deceased sheep, as the majority of the sequencing data from this sample were identified as originating from *B. anthracis* or closely related *Bacillus cereus* and *Bacillus thuringiensis*, with the exception of host sequences (Table S1). Given the inherent error-prone nature of long-read sequencing, we conducted additional validation of the results by employing the qPCR method to assess the presence of the chromosomal marker *dhp61* (BA_5345), as well as the plasmid markers *pagA* (pXO1) and *capC* (pXO2) of *B. anthracis* in the samples. It was not surprising that the DNA samples collected from the spleen, liver, kidney, and lung all yielded strong positive results, as evidenced by cycle threshold (Ct) values ranging from 14 to 22. In contrast, all samples obtained from live animals tested negative for *B. anthracis* genes. These observations clearly indicated that when considering the Ct value of the positive control (which was 31), there was a very high bacterial load in the tissue sample of the sheep that died in the outbreak. In addition, a qPCR analysis was conducted on the bovine samples used for *Anaplasma* screening, yielding comparable outcomes. The analysis of the tissue samples exhibited a positive curve with low Ct values (ranging from 15–20), indicating the significant presence of *B. anthracis*. Conversely, the blood samples obtained from live animals tested negative for *B. anthracis* genes. Additional isolation and purification of *B. anthracis* from the liver and spleen tissues of the deceased sheep were performed through culturing techniques. Gram staining of the colony revealed a unique morphology characterized by elongated and rod-shaped structures, displaying a gram-positive appearance (Fig. 2B). The aforementioned findings provided compelling evidence supporting the assertion that the outbreak was indeed a result of a *B. anthracis* infection.

As of May 9th, the Affina Jalloh farm reported that 66.7% (14 out of 21) of the cattle, 100% (5 out of 5) of the sheep, and 100% (4 out of 4) of the goats had been affected by the disease. Similarly, the Chernor Bah farm reported that 31 out of 85 cattle (36.5%), 18 out of 20 sheep (90%), and 16 out of 16 goats (100%) had also perished due to the disease. These numbers signified a substantial outbreak. In contrast, in an effort to mitigate the economic losses, the farmers sold the meat of the animals that succumbed to the outbreak in the market. This action led to the extensive spread of the threat, which originated from the Port Loko District and had implications for the entire nation, affecting both animal and human populations. Consequently, the findings were promptly communicated to the Ministry of Agriculture.

3.2 The official response and active surveillance
On May 16th, 2022, the Ministry of Agriculture and the Ministry of Health and Sanitation collaboratively announced the official declaration of an anthrax outbreak to effectively mitigate and reduce the risk of transmission associated with the emergency. Subsequently, a comprehensive set of control measures was implemented, encompassing stringent regulations that govern the production, processing, and marketing of livestock and livestock products. As of the declaration of the outbreak, a cumulative number of 233 livestock (91 cattle, 53 goats, and 79 sheep) had been reported as deceased in the impacted region, with a significant portion of the meat having entered the market. The Ministry of Health and Sanitation expeditiously implemented an enhanced surveillance program for anthrax within the local communities. On May 16th, a total of four samples (consisting of three swabs and one blood sample) were collected from seven suspected cases in a village located in the Karene District (Fig. 2C) and subsequently submitted to the IDPC, where all three swabs tested positive for *B. anthracis* by qPCR. All three confirmed cases exhibited a distinctive symptom of cutaneous anthrax (Fig. 2D, 2E) and had been in close proximity to deceased livestock within two weeks. Between May 19th and June 17th, an additional three cases of cutaneous anthrax were confirmed in the IDPC. These cases were reported in the Districts of Karene, Freetown, and Kailahun, as shown in Fig. 2C. All three cases had a history of contacting animal meat within a week prior to the onset of symptoms. Since June 2022, the Ministry of Health and Sanitation has implemented an ongoing active surveillance program to monitor the occurrence of anthrax and other diseases characterized by skin lesions. Over the course of a twelve-month duration that included both a rainy season and a dry season, a total of 43 samples obtained from individuals with suspected cases were subjected to testing for anthrax using both qPCR and cultivation methods. The results of these tests revealed that none of the 43 samples exhibited any traces of anthrax.

### 3.3 The phylogenetic placement of the *B. anthracis* strain responsible for the outbreak

From the tissue samples obtained from deceased cattle and sheep, we successfully isolated and purified two strains of *B. anthracis*. The integration of long-read sequencing data and short-read sequencing data derived from the genomic DNA sample resulted in the identification of three expected contigs, encompassing the complete genome as well as two virulent plasmids, pX01 and pX02. The average sequencing depth achieved in this study was 124-fold. The application of short-read sequencing data for polishing purposes effectively mitigated the indel-introducing effect commonly associated with the long-read sequencing technique [37], thereby exhibiting a high level of reliability [17, 32]. The full-genome SNP sequences were identical between the two isolates. Consequently, considering the shared origin of these samples from a single outbreak, we categorized them as a unified strain, denoted as BaSL2022 (Bacillus Sierra Leone 2022). The construction of the global dendrogram of the *B. anthracis* phylogeny involving the analysis of 236 representative genomes and BaSL2022 was conducted using the maximum likelihood method, which positioned BaSL2022 within the A.Br.008/009 (TEA) sublineage (Fig. 3A) of the A branch. The precise phylogenetic placement of BaSL2022 within the TEA group is illustrated in Fig. 3B, which was generated using 72 representative genomes from the TEA lineage. BaSL2022 was classified within the canonical SNP group A.Br.011/009, more specifically falling under the A.Br.153 subgroup, which was strongly supported by bootstrapping analysis.
4. Discussion

Despite the extensive documentation of anthrax in West Africa, there exists a notable dearth of comprehensive data, particularly pertaining to the diagnosis and genetic sequences of *B. anthracis*, within this region. In the present investigation, we successfully identified an anthrax outbreak impacting both animals and humans in Sierra Leone through the utilization of sequencing and molecular techniques. This advancement signifies a noteworthy achievement in the realm of anthrax prevention and control in Sierra Leone.

In this anthrax outbreak, the mortality of 233 animals in Sierra Leone represents the highest number ever reported in Sierra Leone. Although laboratory confirmation was not conducted for the majority of the animal cases, the confirmation of anthrax infection in the sequentially deceased bovine and sheep, which were determined to be caused by the same strain, provides strong evidence to suggest that the suspected animal cases succumbed to anthrax exposure. From 1947 to 1994, Sierra Leone witnessed a succession of outbreaks that impacted both livestock and humans, as documented in the annual reports of the Ministry of Agriculture, with a particular focus on the Kamakwie area situated in the Northern Region. However, all these instances were diagnosed based solely on symptoms and, at most, with the assistance of microscopy [42]. This situation arose as a result of inadequate laboratory capacities and facilities, which were further compromised during the onset of the civil war in 1991. Such a predicament could potentially provide an explanation for the absence of any officially recorded instances of anthrax cases in the country since 1994. The scarcity of publicly accessible data on the disease further posed a potential contributing factor to the underdiagnosis and misdiagnosis of the disease [7, 42, 43]. The misdiagnosis also occurred at the beginning of the investigation of this anthrax outbreak, which notably hampers the timeliness of the implementation of intervention measures. In view of this, the identification of a widespread outbreak by employing nanopore sequencing to examine clinical samples of an unidentified ailment [17], as demonstrated in this study where we rapidly determined the causative agent without any preconceived notions, provides significant insights for tackling outbreaks of anthrax and other neglected diseases in Sierra Leone and similar countries.

It is worth mentioning that in 2018, a documented report was published regarding four suspected cases of cutaneous anthrax originating from the Kamakwie region of Sierra Leone [13]. In the report, the diagnosis of all cases was conducted using clinical interviews and enzyme-linked immunosorbent assays [13]. However, the application of this assay for diagnosing anthrax has not been widely endorsed due to its limited accuracy in this particular context [42]. This aspect holds particular significance when considering that the cases under investigation occurred in Kamakwie, an area where anthrax was endemic. Consequently, the exclusion of historical anthrax infection in these cases poses a significant challenge, as it may result in misleading outcomes in antibody-based diagnostic procedures. While the diagnosis of anthrax and similar diseases in Sierra Leone and neighboring countries has been impacted by various challenges [11], this paper presents the first report analyzing *B. anthracis* isolates and clinical specimens from suspected cases of cutaneous anthrax using the qPCR method, which has recently
become more accessible and feasible in resource-limited countries due to capacity-building initiatives [11].

The prompt and proactive response from the Ministry of Agriculture and the Ministry of Health and Sanitation to the outbreak played a crucial role in effectively mitigating the risk of spreading. As evidenced by previous investigations [44] and the findings in this study, only six cutaneous anthrax cases were confirmed, all within a month after the implementation of intervention measures. The confirmation of human cases was achieved through qPCR, although attempts to culture *B. anthracis* were unsuccessful. This failure can be attributed to the presence of other bacterial species in the samples, as *B. anthracis* can be readily outcompeted [42]. Consequently, the utilization of molecular techniques for the purpose of tracking human cases presented notable difficulties, thereby raising the question of whether the occurrence of these cases was a consequence of secondary infection stemming from the animal outbreak [6] or if they constituted separate outbreaks. The results obtained from the subsequent long-term active surveillance indicated that the confirmed human cases were predominantly clustered within a limited time period after the outbreak, as opposed to being spread out over the entire duration of the surveillance. This finding provides support for the hypothesis that the human cases resulted from the animal outbreak rather than occurring independently. Notably, despite the absence of confirmed cases during post-outbreak surveillance, it is imperative to avoid underestimating the risk of future anthrax outbreaks, given the prolonged viability of *B. anthracis* spores in the environment, which can persist for several decades [1, 9, 22]. When taking into account the circumstances in which the processing and consumption of contaminated meat occurred locally before the outbreak was officially declared, it is crucial to express considerable concern regarding the repeated occurrence of this disease in Sierra Leone.

The BaSL2022 strain represents the initial identification of the *B. anthracis* isolate discovered in Sierra Leone. The analysis of full-genome SNPs revealed that BaSL2022 was classified within the A.Br.153 phylogenetic group and exhibited a close relationship with isolates (Fig. 3B) obtained from three outbreaks related to drum-making and drumming activities [45] in Scotland (Scotland_484, 2006) [20, 45], London (London_493, 2008) [46], and New York (A3802, 2006) [20, 47]. The animal skins and hides used for the drums associated with the Scotland and New York cases were believed to have been imported from Guinea [20] and Cote d'Ivoire [47], respectively. The London cases were investigated in connection with animal hide sourced from various origins, including Gambia [46]. The observation that the genomes of the strains isolated from the outbreaks were largely similar to each other and constituted a distinct phylogenetic clade (A.Br.153) separate from other existing clades in the TEA lineage (Fig. 3B) further supported the inference that these strains, obtained from cases associated with animal-hide drums, likely originated from the same geographical region, most likely West Africa [20]. However, the verification of this hypothesis was pending, as only three sequences, which were obtained at a later time from samples collected in 2010 in West African Senegal and, specifically, Gambia [23], were found to be different from the sequences obtained from the cases in Scotland, London, and New York. In contrast, a distinct clade known as A.Br.148 [20] was formed, as depicted in Fig. 3B. In light of the aforementioned evidence, our findings present strong support for the conclusive element of the hypothesis, suggesting that the anthrax
incidents in the United Kingdom and United States during the 2000s were caused by the spillover of \textit{B. anthracis} from West Africa.

Considering the geographical origins of the hides associated with the outbreaks in Scotland (Guinea), New York (Cote d'Ivoire), and London (multiple origins including Gambia), the identification of BaSL2022 in Sierra Leone has revealed the potential for long-term endemicity of A.Br.153 in at least the west coast of West Africa. Furthermore, it is important to note that all the strains identified in West Africa were found to be genetically distinct from other phylogenetic groups worldwide [12, 23]. Therefore, the association of BaSL2022 with other A.Br.153 isolates signifies the first instance of \textit{B. anthracis} lineages being transmitted from this specific geographical area to other continents. This statement implies that this particular lineage might have undergone evolutionary changes that resulted in successful adaptation to novel environments [14].

5. Conclusion

Sierra Leone, a West African country, is grappling with the endemicity of anthrax. However, the lack of publicly accessible data on this disease, stemming from historical deficiencies in capacity building, poses a significant obstacle to emergency preparedness efforts for anthrax and other neglected diseases within the nation. In the current investigation, the causative agent responsible for a widespread outbreak was successfully identified as \textit{B. anthracis} through the utilization of the nanopore sequencing technique to analyze clinical specimens. This identification was achieved without any prior assumptions or biases. The initial application of qPCR for the diagnosis of anthrax in this nation resulted in the confirmation of six instances of cutaneous anthrax within a month following the outbreak. Additionally, it effectively ruled out the possibility of anthrax in 43 suspected cases during the year-long surveillance period. These findings demonstrate the successful implementation of a prompt and proactive government response in containing the spread of the disease. We successfully isolated and purified the first Sierra Leonean \textit{B. anthracis} strain BaSL2022 from the outbreak. Phylogenetic analysis revealed that this strain is distinct from the currently available West African lineages and belongs to the A.Br.153 clade. This clade exclusively includes isolates obtained from animal-hide-associated cases in the United Kingdom and United States between 2006 and 2008. This finding provides strong evidence to support the conclusive aspect of the hypothesis that the cases in the United Kingdom and United States were a result of the spillover of \textit{B. anthracis} from West Africa.

Abbreviations

\textit{B. anthracis}: \textit{Bacillus anthracis}

SNP: single nucleotide polymorphism

TEA: TransEuroAsian

MLVA: multilocus variable-number-of-tandem-repeat analysis
cgMLST: core genome multilocus sequence typing

Ct: cycle threshold

qPCR: quantitative polymerase chain reaction

**Declarations**

**Availability of data and materials**

The sequencing data were submitted to the NCBI Sequence Read Archive under the BioProject number PRJNA875505. All the other data yielded in this study are shown in the paper as well as the supplementary materials.

**Acknowledgments**

We would like to extend our utmost gratitude to the personnel of the Department of Animal Sciences at Njala University for their unwavering commitment in the field, as they diligently worked to identify and verify this significant outbreak. The dedication of individuals involved played a crucial role in mitigating the impact of the outbreak. We would like to express our gratitude to Dr. Yajun Song from the Beijing Institute of Microbiology and Epidemiology, Dr. Lingwei Zhu from the Changchun Veterinary Research Institute, and Dr. Lifeng Zhao from Jilin Agricultural Science and Technology University for their invaluable contributions in providing diagnosis consultation and technical support. We would like to extend our gratitude to the 2022 Group of China CDC in Sierra Leone for their invaluable technical support in the field of short-read sequencing.

**Author information**

**Authors and affiliations**

Changchun Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Changchun, China

Shuchao Wang, Yubiao Xie, Shoufeng Zhang, and Rongliang Hu

Department of Animal Sciences, School of Agriculture and Food Sciences, Njala University, Njala, Sierra Leone

Roland Suluku, Mahmud E. Sesay, and Moinina N. Kallo

Department of Microbiology, College of Medicine and Allied Health Sciences, University of Sierra Leone, Freetown, Sierra Leone

Mohamed B. Jalloh and Foday Sahr

Ministry of Agriculture and Food Sciences, Freetown, Sierra Leone
Ahmed F. Samba  
Beijing Institute of Microbiology and Epidemiology, Beijing, China  

Baogui Jian, Yuee Zhao, and Zhiqiang Mi  
Ministry of Health and Sanitation, Freetown, Sierra Leone  

Doris Harding, James S. Squire, and Mohamed A. Vandi  
Beijing Institute of Biotechnology, Beijing, China  

Mengyao Zhan  

Contributions  

SW and ZM performed the identification, isolation, and sequencing of the pathogen. RS, AFS, MNK, and MES collected and processed the samples. DH, JSS, and MAV performed the active surveillance and human sample management. SW, BJ, YX, YZ, MZ, and ZM performed the laboratory detection of active surveillance. SW analyzed the data and drafted the manuscript. SW, MBJ, FS, SZ, RH, and ZM revised and edited the manuscript. All authors have read and approved the submitted version of the manuscript.  

Ethics declarations  

This investigation was performed in response to an emergency and thus not considered purely as scientific research. The animal samples were collected upon request for diagnosis, which was approved by Njala University. The sampling of suspected human cases was approved by the Ministry of Health and Sanitation under an emergency mechanism. All patients or their guardians provided verbal consent to the collection of their samples. Case information other than the location and results was not included in the text.  

Funding  

This work was supported by the Tropical Infectious Diseases Prevention and Control Center of Sierra Leone.  

Competing interests  

The authors declare that there are no conflicts of interest.  

Consent for publication  

Not applicable.  

References


Figures

**Figure 1**

The identification and official response to the anthrax outbreak in Port Loko District. The upper half of the diagram shows points related to animal outbreaks in lilac, while the lower half depicts points associated with human cases in cyan.
Figure 2

The characteristics of animal and human outbreaks. A. Symptoms observed in cattle during the outbreak. B. Gram staining of the culture (400x). C. The geographical locations of the outbreak and cases. The red cross marks the location of the initial animal outbreaks. The circles represent the locations of the Freetown case (left), the Karene cases (middle), and the Kailahun case (right). Specific signs indicating cutaneous cases from Karene are shown in D and E (red arrow).
Figure 3

Phylogenetic analysis of the BaSL2022 isolate based on full-genome SNPs. A. Global distribution of BaSL2022 shown in a maximum likelihood tree. The phylogenetic tree was constructed using full-genome SNPs of the representative genomes described previously [22, 32, 39]. Leaves are categorized and labeled into traditional groups. The TEA group, including BaSL2022, is highlighted in red. B. The specific position of BaSL2022 within the TEA group. The phylogenetic tree was constructed using the maximum likelihood model. It included 72 representative genomes from the TEA group, as well as BaSL2022, with the reference strain being Ames and serving as the outgroup. The tree was drawn to scale, accurately representing the length of the branches. Nodes with a bootstrapping value above 90 are indicated by a black circle, while those between 70 and 90 are indicated by a gray circle. The leaves are displayed with the isolate’s name, country of isolation, and year of isolation. The color blocks on the leaves represent the genetic subgroups of the TEA group, and the nicknames of these subgroups are also displayed in colored blocks outside the leaves. The outer layer displays the canonical SNP groups of isolates within the TEA group using circular blocks. The BaSL2022 is marked with star and shown in blue.

Supplementary Files

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

- AppendixTableS1.docx
- GraphicalAbstract.png