

## RESEARCH ARTICLE

# Epidemiological and comparative genomic analysis of *Bacillus anthracis* isolated from northern Vietnam

Thi Thu Ha Hoang<sup>1</sup>, Duc Anh Dang<sup>1</sup>, Thanh Hai Pham<sup>1</sup>, Minh Hoa Luong<sup>1</sup>, Nhu Duong Tran<sup>1</sup>, Tran Hien Nguyen<sup>1</sup>, Thuy Tram Nguyen<sup>1</sup>, Tran Tuan Nguyen<sup>2</sup>, Satoshi Inoue<sup>3</sup>, Shigeru Morikawa<sup>3</sup>, Akiko Okutani<sup>3\*</sup>

**1** National Institute of Hygiene and Epidemiology, Hanoi, Vietnam, **2** Ha Giang Center for Disease Control, Ha Giang, Vietnam, **3** Department of Veterinary Science, National Institute of Infectious Diseases, Tokyo, Japan

\* [okutani@nih.go.jp](mailto:okutani@nih.go.jp)



## OPEN ACCESS

**Citation:** Hoang TTH, Dang DA, Pham TH, Luong MH, Tran ND, Nguyen TH, et al. (2020) Epidemiological and comparative genomic analysis of *Bacillus anthracis* isolated from northern Vietnam. PLoS ONE 15(2): e0228116. <https://doi.org/10.1371/journal.pone.0228116>

**Editor:** Zhi Ruan, Zhejiang University, CHINA

**Received:** September 3, 2019

**Accepted:** January 7, 2020

**Published:** February 21, 2020

**Copyright:** © 2020 Hoang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The nucleotide sequence data of the Vietnamese *B. anthracis* isolates determined in the present study are available in the DDBJ Sequenced Read Archive (accession numbers DRR166482-DRR166487) and as assembled genomes (accession numbers BLE01000001-BLE01001786(TuanDB), BLEU01000001-BLEU01000153(DB), BLEV01000001-BLEV01000162(HG), BLEW01000001-BLEW01000130(LaLC), BLEX01000001-BLEX01000016(LamDB), BLEY01000001-BLEY01000118(QuyettLC)).

## Abstract

To understand the epidemiological and genetic background of anthrax cases occurring in Vietnam from 2011 to 2015, we surveilled and genetically analyzed *Bacillus anthracis* isolated in the north of the country. Epidemiological surveillance showed that most human cutaneous anthrax cases occurred in association with animal dissection. Whole-genome sequences were obtained from six *B. anthracis* strains from human patients with cutaneous anthrax in the endemic area. Comparative genomic analysis showed that the genetic homogeneity among Vietnamese *B. anthracis* strains was very high. All Vietnamese *B. anthracis* strains belonged to the canSNP lineage of A.Br.011/009, which mostly consists of strains of the trans-Eurasian (TEA) group, including the most closely related strain, Carbosap. To clarify the genetic diversity of Vietnamese strains and strains belonging to A.Br.011/009 and A.Br.008/011 canSNP lineages, we applied a reference genome-based single-nucleotide polymorphism (SNP) and gene-by-gene genomic analysis (whole-genome MLST) strategy. The phylogeny from core genome SNPs revealed that the Vietnamese strains were positioned close to each other; moreover, several SNPs specific to Vietnamese *B. anthracis* were identified. Whole-genome MLST analysis revealed the differences in the number of SNPs between Vietnamese strains, which could enable discrimination at the strain level.

## Introduction

Anthrax is a severe zoonotic disease caused by the spore-forming bacterium *Bacillus anthracis*. *B. anthracis* can survive in the environment for decades as a dormant and stable spore [1, 2]. Herbivores are susceptible to *B. anthracis* infection when grazing on contaminated grass. Naturally occurring human anthrax infections are caused by contact with infected animals or animal products (cutaneous anthrax), ingestion of undercooked infected meat (intestinal anthrax), or lung exposure to spores (inhalational anthrax) [3]. Recently, another form of anthrax infection among heroin users in Europe (injection anthrax) was reported [4].

In Vietnam, anthrax has a wide distribution and remains one of the most serious emerging infectious diseases. It is one of 28 infectious diseases that must be reported in the infectious

**Funding:** AO, SI and SM greatly acknowledge financial support by Japan Agency for Medical Research and Development (AMED; Grant No. 18fk0108049). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

disease surveillance system of Vietnam's Ministry of Health (MOH). From the surveillance conducted by Vietnam's MOH, 20–80 human anthrax cases are reported annually, most of which are from northern mountainous regions of the country. Recently, anthrax cases appear to have been concentrated in the mountainous border regions in the north of the country, including Ha Giang, Dien Bien, Lai Châu, Son La, Cao Bang, and Lao Cai provinces [5]. To understand the epidemiological and genetic background of anthrax cases occurring in Vietnam from 2011 to 2015, we surveilled and genetically analyzed *Bacillus anthracis* isolated in northern Vietnam. Epidemiological surveillance showed that most human cutaneous anthrax cases occurred in association with animal dissection. Whole-genome sequences were obtained from six *B. anthracis* strains from human patients with cutaneous anthrax in the endemic area. Comparative genomic analysis showed that the genetic homogeneity among Vietnamese *B. anthracis* strains was very high. All Vietnamese *B. anthracis* strains belonged to the canSNP lineage [6] of A.Br.011/009, which mostly consists of strains of the TEA group [7], including the most closely related strain, Carbosap. To clarify the genetic diversity of Vietnamese strains and strains belonging to A.Br.011/009 and A.Br.008/011 canSNP lineages, we applied a reference genome-based single-nucleotide polymorphism (SNP) and gene-by-gene genomic analysis (whole-genome MLST) strategy. The phylogeny from the core genome SNPs revealed that the Vietnamese strains were positioned close to each other; moreover, several SNPs specific to Vietnamese *B. anthracis* were identified. Whole-genome MLST analysis revealed the differences in the number of SNPs between Vietnamese strains, which could enable discrimination at the strain level. In the period from 1991 to 1995, three fatal cases out of 441 anthrax patients were reported from northern Vietnam, according to the MOH. The number of cases decreased in the period from 1996 to 2000 to 219, with no fatal cases. Among these latter cases, 211 were reported from northern parts of Vietnam and 8 from elsewhere in the country (southern Vietnam, 4 cases; central highlands, 3 cases; and central Vietnam, 1 case), according to a Health Statistical Profile by the MOH. From 2006 to 2011, three lethal cases out of 413 anthrax patients were reported in the northern provinces, with the most cases recorded in Dien Bien province. Human anthrax cases were reported year-round, but more occurred in the rainy season from June to September. Almost all patients were aged over 15 years, with 37.9% and 36.1% of patients aged 30–49 and 15–29 years old, respectively. The ratio of male to female patients was 7:3 [5]. More than 98% of human patients presented with skin lesions compatible with cutaneous anthrax. The numbers of intestinal anthrax and septicemia cases were very small, accounting for two and one of the fatal cases, respectively. Most cutaneous anthrax patients had a history of contact with sick or dead cattle, buffalo, horse, goat, and other animals through slaughtering or eating them. In one case, a cow had died suddenly. The owner had then opened the carcass and dressed the meat, which was then sold within the village for human consumption. Transportation of infected meat to neighboring villages also occurred, resulting in the spread of disease. To date, we have isolated *B. anthracis* only from lesions of human patients with cutaneous anthrax, but not from environmental sources. The prolonged *B. anthracis* spore phase slows the evolution of this species, resulting in high levels of genetic homogeneity [8]. Recently, single-nucleotide polymorphisms (SNPs) have been identified in *B. anthracis* and global and regional patterns of diversity have been described using canSNP typing [9] and core genome SNPs from *B. anthracis* whole-genome nucleotide sequences [10, 11]. The genetics of *B. anthracis* as determined using whole-genome analysis has been widely reported recently in the form of comparative genomic analysis with isolates from many countries. One of the most common methods used to date has been core genome SNP analysis. It is possible to use comparative whole-genome analysis as an unbiased approach in order to discover informative SNPs and trace the origin and previous spread of anthrax [12]. Whole-genome sequencing and phylogenetic analysis of isolates obtained globally can also be used to

generate hypotheses about the epidemiological connections among geographically disparate isolates and global patterns of anthrax dispersal [13]. It is also extremely useful to perform national sampling studies of *B. anthracis* collections to reveal the actual origin of a strain and identify it accurately [13, 14]. Comparative genome analysis and epidemiological surveillance have also been suggested to aid administrative actions, such as restrictions on the import of livestock products and animals, and thereby help to reduce anthrax cases [15]. SNP typing and subsequent phylogenetic analysis of *B. anthracis* strains isolated from worldwide anthrax cases have resulted in the identification of their genetic background [7, 13, 16, 17]. In the present study, whole-genome sequences of six Vietnamese *B. anthracis* strains isolated from human cutaneous anthrax patients were obtained for *in silico* canSNP typing and comparative genome analysis.

## Materials and methods

### Bacterial strains and DNA extraction

Vietnamese *B. anthracis* strains used for whole-genome sequence analysis are listed in the Table 1 with epidemiological data. The reference strains of *B. anthracis* from other countries used in the present study belonged to A.Br.011/009 and A.Br.008/011 canSNP lineages are listed in S1 Table.

*B. anthracis* strains were isolated from skin lesions of patients with typical symptoms of cutaneous anthrax. Biopsy samples from skin lesions of each patient were directly cultured on 5% sheep blood agar at 37°C overnight and a single colony was isolated as an independent strain. The colonies were identified using biochemical tests and PCR [18].

For DNA extraction, the strains were incubated overnight at 37°C on sheep blood agar/TSMF plates. One loop of the bacterial cells from the agar plates was inoculated into Luria-Bertani broth and grown overnight. Each overnight culture was lysed in lysis buffer (QIAGEN, Hilden, Germany) for 30 min at 37°C. Proteinase K was added at a final concentration of 1 mg/mL and incubated overnight at 56°C. RNase A was added at a final concentration of 20 mg/mL and incubated at 25°C for 2 min. Chromosomal DNA was obtained using the standard phenol-chloroform extraction method and was recovered in the aqueous phase after 20 min of centrifugation at 15,000 × g [19]. DNA was precipitated by the addition of two volumes of ice-cold 95% ethanol and collected by centrifugation for 20 min at 15,000 × g. The DNA was then dissolved in MilliQ water or TE buffer (10 mmol/L Tris and 1.0 mmol/L EDTA, pH 8.0) and filtered through a 0.2-μm filter (Millipore, Billerica, MA). The quantity and quality of DNA were analyzed by NanoDrop (Thermo Fisher Scientific, Waltham, MA).

### Whole-genome sequencing of Vietnamese isolates

The preparation of a genomic library with a 250-bp or 300-bp read length was performed using the NEBNext DNA Library Prep Master Mix kit for Illumina (NEB, Ipswich, MA) with index primers from NEBNext Multiplex Oligos for Illumina (NEB), in accordance with the manufacturer's protocol. DNA library fragments were separated using Agencourt AMPure XP magnetic particles (Beckman Coulter, Brea, CA). Library quality and concentration were determined using BioAnalyzer (Agilent, Santa Clara, CA) with a High Sensitivity DNA Chip Set (Agilent). The libraries were applied to Illumina MiSeq (Illumina, San Diego, CA) using 250 or 300 paired-end sequencing with the MiSeq Reagent Kit v2 (500 cycles) or v3 (600 cycles) for whole-genome sequencing, respectively.

Quality assessment of the obtained reads was performed using the default parameters of CLC Genomic Workbench ver.9.5.3 (QIAGEN). Quality-filtered sequence data from the isolates were subjected to *in silico* canSNP typing determined based on 13 published canonical SNPs [8], as previously reported [15]. SRA files of *B. anthracis* belonging to A.Br.008/011 and

**Table 1. Profiles of six Vietnamese *B. anthracis* strains belonging to A.Br.011/009 canSNP lineages.**

Strain name	Isolation site	Isolation year	Isolation source	Accession number	canSNP lineage
TuanDB	Phinh Sang commune, Tuan Giao district, Dien Bien province, Vietnam	2011	cutaneous anthrax	DRR166487	A.Br.011/009
DB	Phinh Sang commune, Tuan Giao district, Dien Bien province, Vietnam	2011	cutaneous anthrax	DRR166482	A.Br.011/009
LaLC	Than Uyen district, Lai Chau province, Vietnam	2011	cutaneous anthrax	DRR166484	A.Br.011/009
QuyetcLC	Than Uyen district, Lai Chau province, Vietnam	2011	cutaneous anthrax	DRR166486	A.Br.011/009
LamDB	Phinh Sang commune, Tuan Giao district, Dien Bien province, Vietnam	2013	cutaneous anthrax	DRR166485	A.Br.011/009
HG	Niem Tong commune, Meo Vac District, Ha Giang province, Vietnam	2015	cutaneous anthrax	DRR166483	A.Br.011/009

<https://doi.org/10.1371/journal.pone.0228116.t001>

A.Br.011/009 canSNP lineages for comparison with Vietnamese isolates were obtained (Table 1) and converted to fasta or fastq files using SRA toolkit with the command fastq-dump. The nucleotide sequence data of the Vietnamese *B. anthracis* isolates determined in the present study are available in the DDBJ Sequenced Read Archive (accession numbers DRR166482–DRR166487) and as assembled genomes (accession numbers BLET01000001–BLET01001786 (TuanDB), BLEU01000001–BLEU01000153(DB), BLEV01000001–BLEV01000162(HG), BLEW01000001–BLEW01000130(LaLC), BLEX01000001–BLEX01000016(LamDB), BLEY01000001–BLEY01000118(QuyetcLC)).

### Comparative genomic analysis

We used the Parsnp analytical tool from the Harvest Suite software for fast multiple alignment of genomic sequences [20]. Assembled contigs with the default parameters using CLC Genomics Workbench version 11.0.1 (QIAGEN) were used as input for Parsnp analysis (parameters -c -x). The detected SNPs were extracted to a VCF file, using HarvestTools (version 1.0) from the same software package. To improve the overall quality of the data, any SNPs less than 10 bp away from other SNPs, as well as positions that carry an unspecified nucleotide (“N”), were removed. The edited file was used as an input file in HarvestTools to compile the fasta file.

Assembled contigs of Vietnamese strains and five strains from the TEA group [7] were applied to BacWGSTdb [21] for determining the phylogenetic relatedness among the user-uploaded multiple genome sequences, following both SNP and whole-genome multi-locus sequence type (wgMLST) approaches.

### Ethical approval

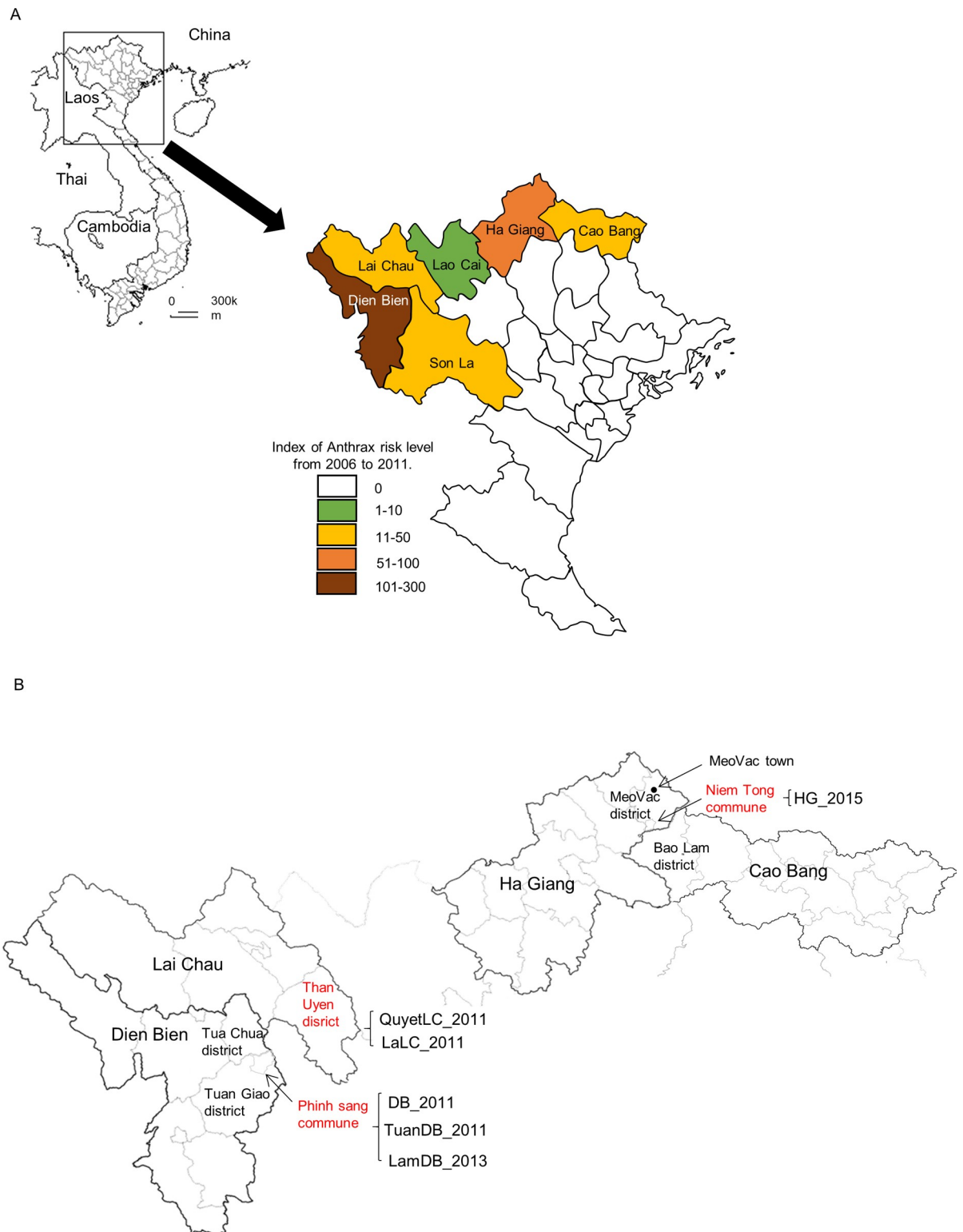
The study protocol was reviewed by the local ethics committee at the National Institute of Hygiene and Environment (NIHE) with approval numbers 01 IRB and 14 IRB and at the National Institute of Infectious Diseases (NIID) with approval number 955.

## Results

### Epidemiological characteristics of Vietnamese anthrax

Recently, we have observed that annual anthrax outbreaks occur in northern Vietnam during the period of heavy rainfall from June to September. This region includes six provinces that are close to the Chinese and Laos borders (Fig 1A). These anthrax outbreaks spread to neighboring villages, where both humans and animals can become infected. The predominant form of human anthrax is cutaneous and includes skin lesions with black eschars on the patients’ necks, fingers (Fig 2), arms, and feet. Diagnosis, by a medical doctor at the Center for Protective Medicine in each province, has usually been based on this typical symptom.

In the study period from 2011 to 2015, 155 suspected anthrax cases were reported based on clinical symptoms. Hotspots of anthrax infection in the communities in the northern



**Fig 1. Epidemiological and geographical information about Vietnamese anthrax.** (A) Map of Vietnam showing the distribution of human anthrax in northern provinces from 2006 to 2011. The color indicates the human anthrax risk level depending on the number of human anthrax cases occurring from 2006 to 2011 [5]. (B) The anthrax hotspot commune and district are shown. The isolation sites of strains used in the present study are indicated in red, from Dien Bien, Lai Chau, and Ha Giang provinces in northern Vietnam.

<https://doi.org/10.1371/journal.pone.0228116.g001>

provinces included Niem Tong commune, Meo Vac town in Meo Vac district in Ha Giang province, Tuan Giao and Tua Chua districts in Dien Bien province, Than Uyen district in Lai Chau province, and Bao Lam district in Cao Bang province (Fig 1B). Human patients were usually treated with traditional herbal medicines at home and upon admission to hospital. In cases with more severe symptoms, patients were treated with the antibiotic ciprofloxacin. Most patients showed clinical features of cutaneous anthrax. Seven patients had acquired infection through contact with animals, while for the others no clear exposure information was available. A total of 50 eschar and vesicular fluid samples were collected from the patients, to identify and confirm the infectious agent by laboratory testing. Six *B. anthracis* strains were isolated from these clinical samples, including three isolates from including three isolates from Dien Bien province, two from Lai Chau province, and one from Ha Giang province. (Fig 1B). All of them were isolated from the cutaneous anthrax patients (Table 1 and Fig 1B).

### Phylogenetic analysis of *B. anthracis* using core genome SNPs and wgMLST from whole-genome sequences

In the present study, we determined the whole-genome sequences of six Vietnamese *B. anthracis* isolates. These six isolates were shown by *in silico* canSNP analysis to belong to the A.Br.011/009 canSNP lineage. Assembled contig files constructed with the CLC Genomic Workbench with the default parameters were used for rapid core genome multiple alignment using Parsnp with 76 strains from other countries belonging to the A.Br.011/009 and A.Br.008/011 canSNP lineages and the Ames Ancestor genome as a reference (Fig 3A). Total coverage among all sequences was 56.4% by Parsnp analysis. The phylogeny using the maximum likelihood method with 4,976 core genome SNPs revealed that an Italian strain, Carbosap, was closely related to the Vietnamese strains. Among the Vietnamese strains, DB2011, LaLc\_2011, LamDB\_2013, and HG\_2015 belonged to the same genetic branch, while TuanDB\_2011 and QuyetLC\_2011 belonged to independent branches (Figs 3B and 4A). TuanDB\_2011 and QuyetLC\_2011 clearly possessed additional SNP patterns compared with the other four Vietnamese strains (Fig 3B). We identified seven SNPs specific for Vietnamese strains from different provinces (S2 Table).

The phylogeny from the core genome SNPs by BacWGSTdb revealed that the Vietnamese strains were positioned close to each other within the TEA polytomy for A.Br.011/009 and A.Br.008/011 canSNP lineages, and that Carbosap was closely related to them, as revealed by Parsnp analysis (Fig 3).

Among Vietnamese strains, strains isolated in the same year in 2011 (DB2011, LaLc\_2011, TuanDB\_2011, and QuyetLC\_2011) showed genetic diversity as assessed by the difference in the number of SNPs, as shown in Fig 4B. TuanDB\_2011 and QuyetLC\_2011 possessed 685 and 688 SNP differences from LaLc\_2011, respectively. DB\_2011 and HG\_2015 had one SNP difference from LaLc\_2011 and LamDB\_2013 possessed 200 SNP differences from LaLc\_2011.

## Discussion

To understand the epidemiological and genetic background of anthrax cases occurring in Vietnam from 2011 to 2015, we surveilled and genetically analyzed *Bacillus anthracis* isolated in northern provinces of Vietnam. Northern mountainous areas such as Lai Châu, Dien Bien, and Ha Giang provinces are considered as high-risk areas for this disease (Fig 1A). The predominant form of human anthrax was cutaneous, but we had not investigated the genetics of *B. anthracis* strains isolated from patient skin. To understand the genetic background of *B. anthracis* isolated from human cutaneous anthrax, whole-genome analysis of six strains isolated from Lai Châu, Dien Bien, and Ha Giang provinces (Table 1) was performed by a

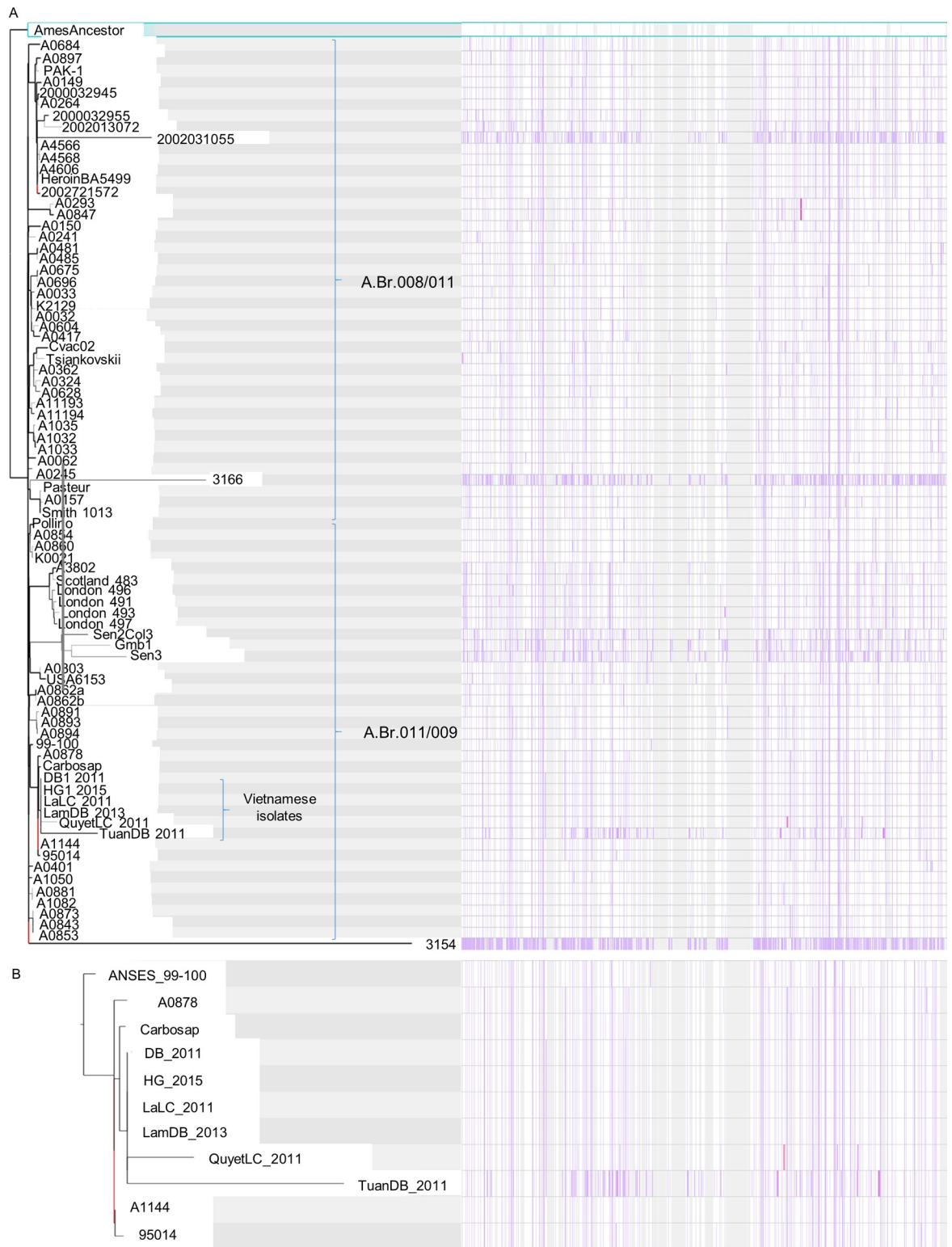


**Fig 2. The affected finger of a cutaneous anthrax patient.** A typical black eschar on the finger of a cutaneous anthrax patient who had a history of contact with *B. anthracis*-infected animals and carcasses.

<https://doi.org/10.1371/journal.pone.0228116.g002>

reference genome-based single-nucleotide polymorphism (core genome SNP) strategy, as well as a gene-by-gene genomic analysis (wgMLST) encompassing not only core genes/loci but also accessory genes/loci using Parsnp and BacWGSTdb. This was the first whole-genome analysis of Vietnamese *B. anthracis* strains and revealed that they were all classified into the A.Br.011/009 canSNP lineage; no strains from Asia have previously been reported to belong to this lineage. The Carbosap clade that Sahl et al. reported [9] was the nearest genetic clade, which consisted of strains isolated in Italy, France, and Argentina. Seven SNPs specific to Vietnamese *B. anthracis* were found using the Parsnp program (S2 Table); however, Vietnamese *B. anthracis* strains isolated in the same year, namely, 2001, were rather heterogeneous compared with those isolated in 2013 or 2015 (Figs 3 and 4A). In addition, the strains from the same province showed heterogeneity, contrary to expectations. For example, strains in Dien Bien province, DB\_2011, TuanDB\_2011, and LamDB\_2013, were classified into different branches by core genome SNP analysis (Figs 3B and 4A) and the numbers of SNP differences by wgMLST analysis (Fig 4B). Moreover, strains in Lai Châu province, LaLC\_2011 and QuyetLC\_2011, belonged to different branches with more than 600 SNP differences between them.

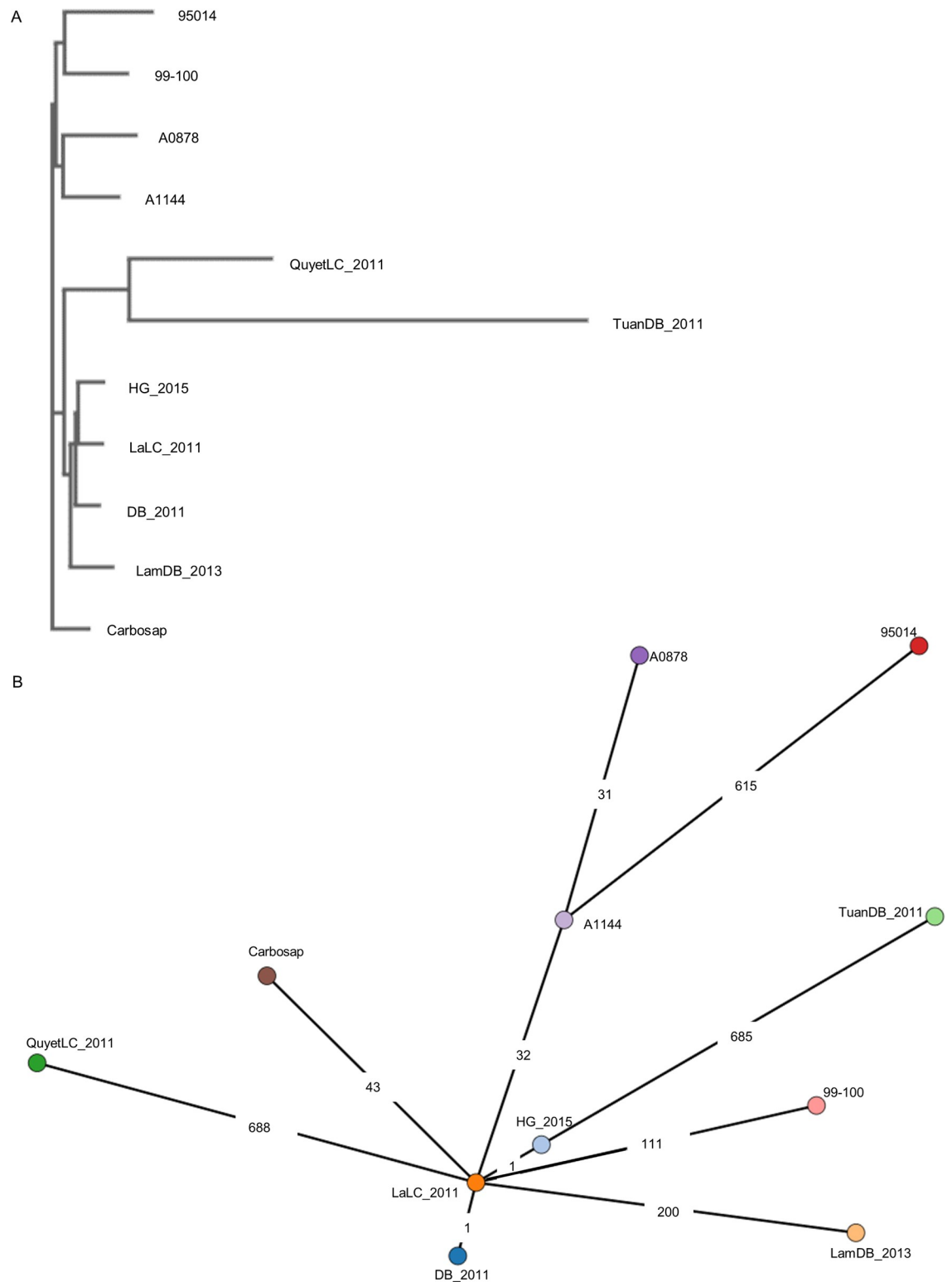
The numbers of SNP differences between strains (Fig 4B) suggested that genetically diverse strains may have expanded in several northern provinces in Vietnam around 2011, however, genetically clonal strains like LamDB\_2013 and HG\_2015 have spread beyond these provinces since then.



**Fig 3. Phylogeny and multiple alignments of *B. anthracis* strains belonging to A.Br.008/011 and A.Br.011/009 canSNP lineages.** A total of 4,976 core genome chromosomal SNPs (S3 Table) from 82 strains were extracted using Parsnp. (A) The phylogenetic tree and SNP heatmap showing SNP density across the genome of each strain of the A.Br.008/011 and A.Br.011/009 canSNP lineages used in the present study. The Ames Ancestor strain was used as a reference and an outgroup. (B) The phylogenetic tree and SNP heatmap of TEA group strains and Vietnamese strains. Screenshots of the phylogenetic tree and SNP heatmap were obtained as output by Gingr.

<https://doi.org/10.1371/journal.pone.0228116.g003>





**Fig 4. Multiple genome analysis was used to determine the phylogenetic relatedness among the genome sequences of TEA group strains following both SNP and whole-genome MLST approaches using tools in BacWGSTdb. (A)** The phylogenetic tree based on the SNP strategy from the assembled contigs of six Vietnamese strains and five genetically closest TEA group strains (Carbosap, A0878, 95014, A1144, and ANSES\_99-100) was visualized using the neighbor joining method. **(B)** Minimum spanning tree based on the whole-genome MLST strategy encompassing not only core genome genes/loci but also accessory genes/loci was visualized with static redraw and log scale modes. The numbers on the branches indicate the number of SNPs between strains.

<https://doi.org/10.1371/journal.pone.0228116.g004>

In the present study, whole-genome analysis with a rather small number of strains identified distinct relationships between genetic features and the geographical distribution of strains. Although there was a limited number of strains for genetic analysis, it was interesting to find that strains isolated in northern Vietnam belonged to only one canSNP lineage, with high homogeneity. SNPs specific for strains from different provinces as determined in the present study could be useful for further analysis of isolates. We believe that genomic analysis with detailed epidemiological information and continuous surveillance and sampling of humans, animals, and livestock products could shed light on the infectious source or route into the endemic area in northern Vietnam.

## Supporting information

**S1 Table. The reference strains of *Bacillus anthracis* used in the present study belonged to A.Br.011/009 and A.Br.008/011 canSNP lineages.**

(XLSX)

**S2 Table. Seven SNPs specific for Vietnamese strains were identified by Parsnp.**

(XLSX)

**S3 Table. The list of 4,976 core genome SNPs extracted by Parsnp.**

(XLSX)

## Acknowledgments

We thank Dr. Ken Maeda from the National Institute of Infectious Diseases, Tokyo, Japan, for his critical review of the manuscript. We also thank Edanz ([www.edanzediting.co.jp](http://www.edanzediting.co.jp)) for editing the English text of a draft of this manuscript.

## Author Contributions

**Conceptualization:** Thi Thu Ha Hoang, Akiko Okutani.

**Data curation:** Duc Anh Dang, Satoshi Inoue, Shigeru Morikawa.

**Funding acquisition:** Satoshi Inoue, Shigeru Morikawa, Akiko Okutani.

**Investigation:** Thi Thu Ha Hoang, Thanh Hai Pham, Minh Hoa Luong, Nhu Duong Tran, Tran Hien Nguyen, Thuy Tram Nguyen, Tran Tuan Nguyen, Akiko Okutani.

**Methodology:** Nhu Duong Tran.

**Resources:** Tran Tuan Nguyen.

**Supervision:** Duc Anh Dang, Tran Hien Nguyen, Shigeru Morikawa.

**Visualization:** Thi Thu Ha Hoang, Akiko Okutani.

**Writing – original draft:** Thi Thu Ha Hoang, Akiko Okutani.

**Writing – review & editing:** Shigeru Morikawa.

## References

1. Dragon DC, Rennie RP. The ecology of anthrax spores: tough but not invincible. *Can Vet J.* 1995; 36(5):295–301. PMID: [7773917](https://pubmed.ncbi.nlm.nih.gov/7773917/); PubMed Central PMCID: PMC1686874.
2. Hugh-Jones M, Blackburn J. The ecology of *Bacillus anthracis*. *Mol Aspects Med.* 2009; 30(6):356–67. <https://doi.org/10.1016/j.mam.2009.08.003> PMID: [19720074](https://pubmed.ncbi.nlm.nih.gov/19720074/).

3. World Health Organization. Guidelines for the surveillance and control of anthrax in humans and animals. 1998. [http://www.who.int/csr/resources/publications/anthrax/WHO\\_EM\\_C\\_ZDI\\_98\\_6/en/](http://www.who.int/csr/resources/publications/anthrax/WHO_EM_C_ZDI_98_6/en/)
4. Hanczaruk M, Reischl U, Holzmann T, Frangoulidis D, Wagner DM, Keim PS, et al. Injectional anthrax in heroin users, Europe, 2000–2012. *Emerg Infect Dis*. 2014; 20(2):322–3. <https://doi.org/10.3201/eid2002.120921> PMID: 24447525
5. Duong TN, Cam HPT. Epidemiological characteristics of human anthrax in the northern region of Vietnam, 2006–2011. [In Vietnamese]. *Viet J Prevent Med*. 2012; 22(2): 14–23.
6. Pilo P, Frey J. *Bacillus anthracis*: molecular taxonomy, population genetics, phylogeny and patho-evolution. *Infect Genet Evol*. 2011; 11(6):1218–24. <https://doi.org/10.1016/j.meegid.2011.05.013> PMID: 21640849.
7. Sahl JW, Pearson T, Okinaka R, Schupp JM, Gillette JD, Heaton H, et al. A *Bacillus anthracis* Genome Sequence from the Sverdlovsk 1979 Autopsy Specimens. *mBio*. 2016; 7(5). <https://doi.org/10.1128/mBio.01501-16> PMID: 27677796; PubMed Central PMCID: PMC5050339.
8. 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. <https://doi.org/10.1371/journal.pone.0000461> PMID: 17520020.
9. Van Ert MN, Easterday WR, Simonson TS, U'Ren JM, Pearson T, Kenefic LJ, et al. Strain-specific single-nucleotide polymorphism assays for the *Bacillus anthracis* Ames strain. *J Clin Microbiol*. 2007; 45(1):47–53. Epub 2006/11/10. doi: JCM.01233-06 [pii] <https://doi.org/10.1128/JCM.01233-06> PMID: 17093023; PubMed Central PMCID: PMC1828967.
10. 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. Epub 2004/09/07. <https://doi.org/10.1073/pnas.0403844101> PMID: 15347815; PubMed Central PMCID: PMC518758.
11. 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. <https://doi.org/10.1016/j.meegid.2009.05.014> PMID: 19477301.
12. Girault G, Thierry S, Cherchame E, Derzelle S. Application of High-Throughput Sequencing: Discovery of Informative SNPs to Subtype *Bacillus anthracis*. *Scientific Research*. 2014; 5(7):669–77. <https://doi.org/10.4236/abb.2014.57079>
13. Derzelle S, Aguilar-Bultet L, Frey J. Whole genome SNP analysis of bovine *Bacillus anthracis* strains from Switzerland reflects strict regional separation of Simmental and Swiss Brown breeds in the past. *Vet Microbiol*. 2016; 196:1–8. <https://doi.org/10.1016/j.vetmic.2016.10.014> PMID: 27939145.
14. Derzelle S, Girault G, Roest HI, Koene M. Molecular diversity of *Bacillus anthracis* in the Netherlands: investigating the relationship to the worldwide population using whole-genome SNP discovery. *Infect Genet Evol*. 2015; 32:370–6. <https://doi.org/10.1016/j.meegid.2015.03.030> PMID: 25843650.
15. Okutani A, Inoue S, Morikawa S. Comparative genomics and phylogenetic analysis of *Bacillus anthracis* strains isolated from domestic animals in Japan. *Infect Genet Evol*. 2019; 71:128–39. <https://doi.org/10.1016/j.meegid.2019.03.022> PMID: 30928604
16. Derzelle S, Aguilar-Bultet L, Frey J. Comparative genomics of *Bacillus anthracis* from the wool industry highlights polymorphisms of lineage A.Br.Vollum. *Infect Genet Evol*. 2016; 46:50–8. <https://doi.org/10.1016/j.meegid.2016.10.019> PMID: 27793731.
17. 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. <https://doi.org/10.1371/journal.pone.0146216> PMID: 26901621; PubMed Central PMCID: PMC4763433.
18. Fasanella A, Taranto PD, Garofolo G, Colao V, Marino L, Buonavoglia D, et al. Ground Anthrax *Bacillus anthracis* Refined Isolation (GABRI) method for analyzing environmental samples with low levels of *Bacillus anthracis* contamination. *BMC Microbiol*. 2013; 13(1):167. <https://doi.org/10.1186/1471-2180-13-167> PMID: 23865983
19. Sambrook J, Russell DW. Purification of nucleic acids by extraction with phenol:chloroform. *Cold Spring Harbor Protocols*. 2006; 2006(1). <https://doi.org/10.1101/pdb.prot4455> PMID: 22485786.
20. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol*. 2014; 15(11):524. <https://doi.org/10.1186/s13059-014-0524-x> PMID: 25410596; PubMed Central PMCID: PMC4262987.
21. Ruan Z, Feng Y. BacWGSTdb, a database for genotyping and source tracking bacterial pathogens. *Nucleic Acids Res*. 2016; 44(D1):D682–7. Epub 2015/10/04. <https://doi.org/10.1093/nar/gkv1004> PMID: 26433226; PubMed Central PMCID: PMC4702769.