JOURNAL OF CLINICAL MICROBIOLOGY, Aug. 2010, p. 2821–2829 Vol. 48, No. 8
0095-1137/10/$12.00 doi:10.1128/JCM.00137-10

Genome-Wide Single Nucleotide Polymorphism Typing Method for Identification of *Bacillus anthracis* Species and Strains among *B. cereus* Group Species▽†

Makoto Kuroda,[1]* Masakuni Serizawa,[1] Akiko Okutani,[2] Tsuyoshi Sekizuka,[1] Satomi Banno,[1] and Satoshi Inoue[2]

Laboratory of Bacterial Genomics, Pathogen Genomics Center,[1] and Department of Veterinary Science,[2] National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

Received 22 January 2010/Returned for modification 24 May 2010/Accepted 4 June 2010

As an issue of biosecurity, species-specific genetic markers have been well characterized. However, *Bacillus anthracis* strain-specific information is currently not sufficient for traceability to identify the origin of the strain. By using genome-wide screening using short read mapping, we identified strain-specific single nucleotide polymorphisms (SNPs) among *B. anthracis* strains including Japanese isolates, and we further developed a simplified 80-tag SNP typing method for the primary investigation of traceability. These 80-tag SNPs were selected from 2,965 SNPs on the chromosome and the pXO1 and pXO2 plasmids from a total of 19 *B. anthracis* strains, including the available genome sequences of 17 strains in the GenBank database and 2 Japanese isolates that were sequenced in this study. Phylogenetic analysis based on 80-tag SNP typing showed a higher resolution power to discriminate 12 Japanese isolates rather than the 25 loci identified by multiple-locus variable-number tandem-repeat analysis (MLVA). In addition, the 80-tag PCR testing enabled the discrimination of *B. anthracis* from other *B. cereus* group species, helping to identify whether a suspected sample originates from the intentional release of a bioterrorism agent or environmental contamination with a virulent agent. In conclusion, 80-tag SNP typing can be a rapid and sufficient test for the primary investigation of strain origin. Subsequent whole-genome sequencing will reveal apparent strain-specific genetic markers for traceability of strains following an anthrax outbreak.

Many potential bioterrorism agents, including anthrax, present as pulmonary disease. Anthrax is caused by the spore-forming bacterium *Bacillus anthracis*, which is among the most severe zoonoses posing a serious threat to both public and animal health (7, 14). *B. anthracis* belongs to the *Bacillus cereus* group of bacteria, which is composed of closely related Gram-positive organisms with highly divergent virulent properties (14, 18). Infection with this bacterium can occur through the skin, gastrointestinal tract, or respiratory apparatus following contact, ingestion, or inhalation of spores, respectively (7, 14).

As an issue of biosecurity, a comprehensive molecular diagnosis system is considered for detecting potential infectious agents. For most potential bioterrorism agents, species-specific genetic markers have been well characterized (9), but strain-specific information is not sufficient for traceability to identify the origin of the strain.

A liquid suspension of *B. anthracis* was dispersed by the Aum Shinrikyo religious cult in Japan in 1993. The genotype of the *B. anthracis* isolate released was identical to that of the Sterne 34F2 strain, which is a member of the A3b diversity cluster (10). Fortunately, there were no victims of this attack because the strain was pXO2 plasmid defective and a low-virulent derivative used commercially in Japan to vaccinate animals against anthrax. The recent "postal anthrax attacks" in the United States aimed at the intentional release of *B. anthracis* spores underlies the growing importance of the identification of *B. anthracis* at the strain level in forensic and epidemiological investigations (2, 15, 17, 22). These cases indicate that rapid and adequate testing will be required for traceability.

Multiple-locus variable-number tandem-repeat analysis (MLVA) 25 (4, 13) or canonical SNPs (canSNPs) in combination with MLVA 15 (25, 26) facilitate the genotyping of *B. anthracis* strains. However, both typing systems require fragment analysis of multiple repeats, and the number of repeats is likely to be missassigned due to the use of different fragment analysis platforms in individual laboratories. In contrast, SNP alleles are correctly called by the DNA sequencing technique used, and they are more definitive than the ambiguous length of multiple repeats; moreover, whole-genome analysis enables the comprehensive identification of strain-specific genetic markers. In this study, we conducted genome-wide screening of whole SNPs among *B. anthracis* strains, including Japanese isolates, and constructed a simplified SNP-typing method using tag SNPs to facilitate the identification of strain lineage based on the whole-genome sequence of *B. anthracis.*

* Corresponding author. Mailing address: Laboratory of Bacterial Genomics, Pathogen Genomics Center, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan. Phone: 81-3-5285-1111. Fax: 81-3-5285-1166. E-mail: makokuro@nih.go.jp.

† Supplemental material for this article may be found at http://jcm.asm.org/.

▽ Published ahead of print on 16 June 2010.

TABLE 1. Information on strains used for SNP analysis[a]

Strain	Location isolated or type of strain	Yr isolated	pXO1	pXO2	MLVA 25[b]	Accession no. (chromosome; pXO1; pXO2)	Source or reference
Bacillus anthracis							
BA102	Miyagi, Japan	1983	+	+	A3b	NA	This study
BA103	Miyagi, Japan	1991	+	+	A3b	DRA000067 (short reads archive)	This study
BA104	Shizuoka, Japan	1982	+	+	A3a	DRA000068 (short reads archive)	This study
BA105	Shizuoka, Japan	1982	+	+	A3a	NA	This study
BA106	Okinawa, Japan	1956	+	+	A3a	NA	This study
BA107	Okinawa, Japan	1982	+	+	A3a	NA	This study
BA108	Shiga, Japan	1987	+	+	A3b	NA	This study
BA109	Mie, Japan	1970	+	+	A3a	NA	This study
BA110	Mie, Japan	1967	+	+	A3a	NA	This study
BA111	Okayama, Japan	1985	+	+	A3b	NA	This study
BA113	Okayama, Japan	NA	+	+	A3a	NA	This study
BA115	Shizuoka, Japan	NA	+	+	A3b	NA	This study
Ames	Laboratory strain	NA	−	−	A3b	NC_003997	18, 21
Ames 0581	Gold standard	NA	+	+	A3b	NC_007530; NC_007322; NC_007323	20
Sterne	Counterpart to the Pasteur strain	NA	+	−	A3b	NC_005945	18
A0174	Canada	NA	+	−	NA	NZ_ABLT00000000	Unpublished
A0193	South Dakota, USA	NA	+	+	A1a	NZ_ABKF00000000	Unpublished
A0389	Bekasi, Indonesia	NA	+	+	NA	NZ_ABLB00000000	Unpublished
A0442	Kruger National Park, South Africa	NA	+	+	NA	ABKG01000000	Unpublished
A0465	France	NA	+	+	B2	NZ_ABLH00000000	Unpublished
A0488	UK	1935	+	+	A4	NZ_ABJC00000000	Unpublished
A1055	Laboratory strain	NA	−	+	C	NZ_AAEO00000000	Unpublished
A2012	West Palm Beach, FL, USA	2001	+	+	NA	NZ_AAAC00000000	Unpublished
Australia 94	Australia	NA	+	+	A3a	NZ_AAES00000000	Unpublished
CNEVA-9066	France	NA	+	+	B2	NZ_AAEN00000000	Unpublished
Kruger B	Kruger National Park, South Africa	NA	+	+	B1	NZ_AAEQ00000000	Unpublished
Tsiankovskii-I	Former Soviet Union	1960	+	+	NA	NZ_ABDN00000000	Unpublished
Vollum	Laboratory strain	NA	+	+	A4	NZ_AAEP00000000	Unpublished
WesternNA USA6153	USA	NA	+	+	A1a	NZ_AAER00000000	Unpublished
Bacillus cereus							
AH187 (F4810/72)	London, UK	NA	NA	NA	NA	NC_011658	Unpublished
AH820	Akershus, Norway	1995	NA	NA	NA	NC_011773	Unpublished
ATCC 10987	Canada	NA	NA	NA	NA	AE017194, NC_003909	19
ATCC 14579	NA	NA	NA	NA	NA	AE016877, NC_004722	8
B4264	NA	1969	NA	NA	NA	NC_011725	Unpublished
E33L (ZK)	Namibia	1996	NA	NA	NA	CP000001, NC_006274	5
G9842	Nebraska, USA	1996	NA	NA	NA	NC_011772	Unpublished
NVH 391-98	NA	NA	NA	NA	NA	NC_009674	18
03BB108	NA	NA	NA	NA	NA	NZ_ABDM00000000	Unpublished
AH1134	Oklahoma, USA	NA	NA	NA	NA	NZ_ABDA00000000	Unpublished
G9241	NA	NA	NA	NA	NA	NZ_AAEK00000000	6
H3081.97	NA	NA	NA	NA	NA	NZ_ABDL00000000	Unpublished
NVH0597-99	NA	1999	NA	NA	NA	NZ_ABDK00000000	Unpublished
W	NA	NA	NA	NA	NA	NZ_ABCZ00000000	Unpublished
NBRC 3466	NA	NA	NA	NA	NA	NA	This study
NBRC 13494	NA	NA	NA	NA	NA	NA	This study
NBRC 15305	NA	NA	NA	NA	NA	NA	This study
GTC419	NA	NA	NA	NA	NA	NA	This study
GTC1777	Japan	NA	NA	NA	NA	NA	This study
GTC2886	Japan	NA	NA	NA	NA	NA	This study
GTC2903	Japan	NA	NA	NA	NA	NA	This study
GTC2926	NA	NA	NA	NA	NA	NA	This study
Bacillus thuringiensis							
97-27	NA	NA	NA	NA	NA	AE017355, NC_005957	5
Al Hakam	NA	NA	NA	NA	NA	NC_008600	3
ATCC 35646	Israel	NA	NA	NA	NA	NZ_AAJM00000000	Unpublished
NBRC 3951	NA	NA	NA	NA	NA	NA	This study
NBRC 13865	NA	NA	NA	NA	NA	NA	This study
NBRC 13866	NA	NA	NA	NA	NA	NA	This study
GTC2847	NA	NA	NA	NA	NA	NA	This study
Bacillus weihenstephanensis							
KBAB4	NA	NA	NA	NA	NA	NC_010184	18

[a] NA, not available.
[b] Referred by Lista et al. (13) and Okutani et al. (16).

MATERIALS AND METHODS

***B. anthracis* strains.** Japanese isolates of *B. anthracis* BA103 or BA104 were used for whole-genome sequencing as representative strains in the A3a or A3b cluster, respectively, and analyzed using the MLVA 25 method (16). Genomic information of the other strains is shown in Table 1.

Short read DNA sequencing using the Illumina genome analyzer II (GA II). Library preparation was performed using a genomic DNA sample preparation kit (Illumina, San Diego, CA), and DNA clusters were generated on a slide using the cluster generation kit (v.2) on an Illumina cluster station (Illumina) according to the manufacturer's instructions. To obtain ~10 million clusters for one lane, the general procedure described in the standard protocol (Illumina) was performed as follows: template hybridization, isothermal amplification, linearization, blocking, denaturation, and hybridization of the sequencing primer (Illumina). All sequencing runs were performed with the GA II using the Illumina sequencing kit (v.3). Fluorescent images were analyzed with the Illumina base-calling pipeline v.1.3.2 to obtain FASTQ-formatted sequence data of 50-mer short reads.

Whole SNP extraction. A schematic flowchart of the data processing procedure is shown in Fig. 1. To identify whole SNPs compared with the reference sequence of *B. anthracis* Ames 0581, Maq software (v.0.7.1) (12), a mapping assembler for short reads generated by the next-generation sequencer, was used with the "easyrun" command as the default parameter. Strain-specific SNPs were extracted from the "cns.final.snp" files (12) by comparison between the genome sequence of the tested strain and that of Ames 0581. Read alignment for the validation of SNPs was performed using the MapView graphical alignment viewer (1). To extract whole SNPs from the available genomic sequences of other *B. anthracis* strains, Maq software (v.0.7.1) (12) was used with the "maq simulate" command with a modification of the following default parameters: number of pairs of reads, "-N 10000000"; mutation rate, "-r 0"; and fraction of 1-bp indels, "-R 0." These parameters indicate that 20 million 36-mer hypothetical reads were generated with neither mutations nor indels from the genomic sequence for SNP identification. SNPs located in repetitive sequence regions (e.g., variable-number tandem repeats [VNTRs], rRNA, and insertion sequence) were excluded from the analysis. Furthermore, a BLASTN search was performed for the validation of the SNP findings.

Tag SNP selection. Tag SNPs, representative SNPs in a region of the genome with high linkage disequilibrium, were selected from whole SNPs. Each SNP allele was assigned as major or minor, followed by conversion to 1 or 0 as major or minor, respectively (Fig. 1). These allele patterns were sorted and classified into each tag SNP group. A single representative SNP was selected from each tag SNP group.

PCR amplification. PCR amplification was performed using 50 ng of genomic DNA and Ex*Taq* DNA polymerase (Takara, Shiga, Japan) with a PE Applied Biosystems PCR 9600 machine (Applied Biosystems, Foster City, CA) with the following program: initial denaturation, 95°C for 5 min; and 3 steps of amplification (×30 cycles), 95°C for 30 s, 55°C for 30 s, and 72°C for 1 min. PCR primer sequences for 80-tag SNPs are shown in Table S2 in the supplemental material. PCR products were verified by 1% agarose gel electrophoresis, followed by Sanger sequencing using the BigDye Terminator v.3.1 cycle sequencing kit (Applied Biosystems) with PCR primers.

Phylogenetic analysis. Multiple sequence alignment was performed using ClustalW (11), and the phylogenetic tree was constructed by using neighbor-joining (NJ) methods with 1,000 times bootstrapping or the unweighted pair group method with arithmetic mean (UPGMA). FigTree v.1.2.3 software was used to display the tree.

Short read archive accession numbers. Short read archives have been deposited in the DNA Data Bank of Japan (DDBJ; accession numbers DRA000067 and DRA000068 for BA103 and BA104, respectively).

RESULTS

Summary of sequencing reads and coverage for Japanese isolates BA103 and BA104. The GA II sequencer produced 6.8 million to 7.6 million 50-base-long reads per strain after applying the quality filter of the Illumina base-calling pipeline v.1.3.2 (Table 2). The filter-passed reads were aligned to the reference sequence of the Ames 0581 strain using Maq software, resulting in more than a 56-fold coverage depth on average and a 98.3 to 99.3% coverage of the sequence,

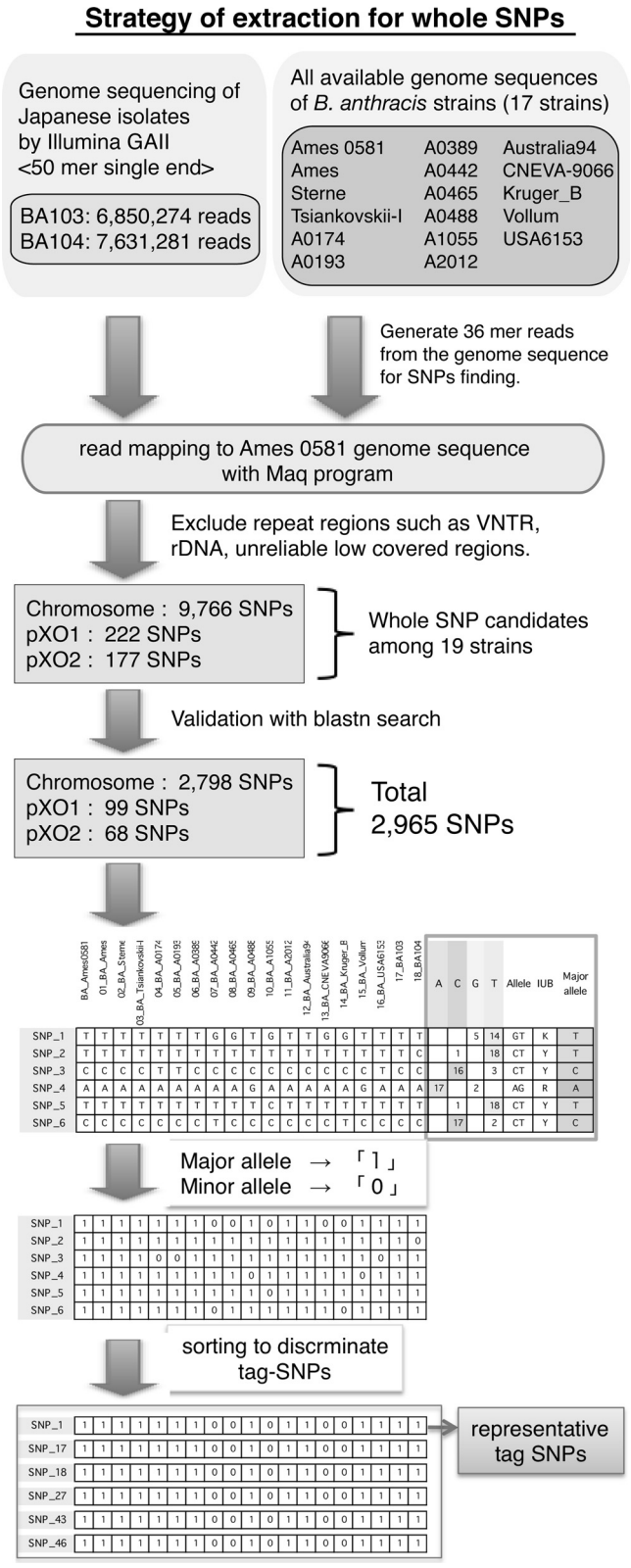

FIG. 1. Schematic representation of the extraction of whole SNPs from the genomic sequences of *B. anthracis* strains. In total, 19 strains including 2 Japanese isolates in this study and 17 available genomic sequences were used for SNP extraction.

TABLE 2. Experimental parameters obtained in whole-genome sequencing of Japanese *B. anthracis* strains

Parameter	BA103	BA104
Total no. of reads passing quality filter[a]	6,850,274	7,631,281
Total no. of bases passing quality filter	342,513,700	381,564,050
No. of reads aligned to each reference sequence by Maq[b]		
Chromosome	5,862,074	6,657,825
pXO1	486,900	442,538
pXO2	178,142	166,943
No. (%) of reads unaligned by Maq	323158 (4.72)	363975 (4.77)
Average coverage depth (fold)		
Chromosome	56.07	63.68
pXO1	133.94	121.73
pXO2	93.94	88.04
Total length of covered regions by reads (%)		
Chromosome	5,150,227 (98.5)	5,149,114 (98.5)
pXO1	179,906 (99.0)	180,573 (99.3)
pXO2	92,880 (98.0)	93,155 (98.3)

[a] Total read number that passed quality check procedure of Illumina base-calling pipeline 1.3.2.
[b] Obtained total reads were mapped onto a reference genome sequence of *B. anthracis* Ames 0581.

excluding ambiguous repetitive regions (Table 2). Since genomic analysis of laboratory strains of *B. subtilis* using the Illumina GA I was performed at a maximum coverage of 51.7-fold (24), the coverage in the present study would be sufficient for the identification of SNPs.

Extraction of whole strain-specific SNPs among *B. anthracis* strains. To extract whole candidates for SNP alleles, 17 other available genomic sequences of *B. anthracis* strains were also compared to that of the Ames 0581 strain for SNP identification using *in silico* analysis (Table 1; Fig. 1). A total of 2,965 reliable SNPs in the chromosome, and the pXO1 and pXO2 plasmids, were identified among the 19 strains examined, including 2 Japanese isolates (see Table S1 in the supplemental material). A phylogenetic tree was constructed based on the concatenated sequences of whole SNP alleles (Fig. 2). SNP variation showed the corresponding phylogenetic relationship as well as the results obtained by MLVA 25 (16), indicating that these genetic alterations appear to be inherited as strain-specific markers even though VNTRs and SNPs are distinct types of genetic variation. Regarding the stability of these genetic alterations, SNPs are assumed to be more stable and definitive markers than VNTRs; therefore, we further developed a more simplified SNP typing method to enable rapid testing.

PCR analysis of 80-tag SNP groups for primary investigation of *B. anthracis* strains. We classified 80-tag SNP groups from all 2,965 SNPs (Fig. 1). To select a single SNP locus from the multiple loci of each tag group, an SNP locus located on the coding sequence and specific to *B. anthracis* was preferentially selected; otherwise, some SNPs were located in a noncoding sequence (see Table S2 in the supplemental material). Through the selection of the representative SNP locus in the 80-tag SNP group, each locus was chosen in advance so that there would be no amplification from other *B. cereus* group species such as *B. cereus*, *B. thuringiensis*, and *B. weihenstephanensis*, if possible. Figure 3 shows the *in silico* simulation of PCR amplification using 80-tag SNPs among the *B. cereus* group species. PCR simulation indicated that most SNPs could be amplified in natural isolates of *B. anthracis* (pXO1- and pXO2-positive), while the animal vaccine source Sterne strain was easily discriminated as a pXO2-defective strain. Based on the simulation, we determined that the 80-tag SNP PCR enabled the discrimination of *B. anthracis* from other *B. cereus* group species (Fig. 3).

However, the virulent strains *B. cereus* E33L_ZK and *B. thuringiensis* 97-27, which are closely related to *B. anthracis* (5), showed more positive amplifications of the SNP loci than the other strains, suggesting that their genomic sequences share similar variations to some extent. Furthermore, one striking report revealed that the *B. cereus* G9241 strain possesses a pXO1-like plasmid carrying the edema factor and lethal factor (6); this present study indicates that the 80-tag SNP PCR testing method also has the potential to identify a strain carrying anthrax toxins among the *B. cereus* group.

To validate whether the PCR typing method works, 12 Japanese strains of *B. anthracis*, 8 strains of *B. cereus*, and 4 strains of *B. thuringiensis* were investigated (Fig. 4). All tested *B. anthracis* strains were definitely discriminated from other members of the *B. cereus* group. In contrast to the *in silico* simulation shown in Fig. 3, some of the chromosomal tag SNPs are likely to be amplified in *B. cereus* strains. For instance, *B. cereus* type strain NBRC 15305, corresponding to ATCC 14579, showed different profiles between the *in silico* simulation (Fig. 3) and the actual PCR trial (Fig. 4). The simulation represents a virtual result with perfect

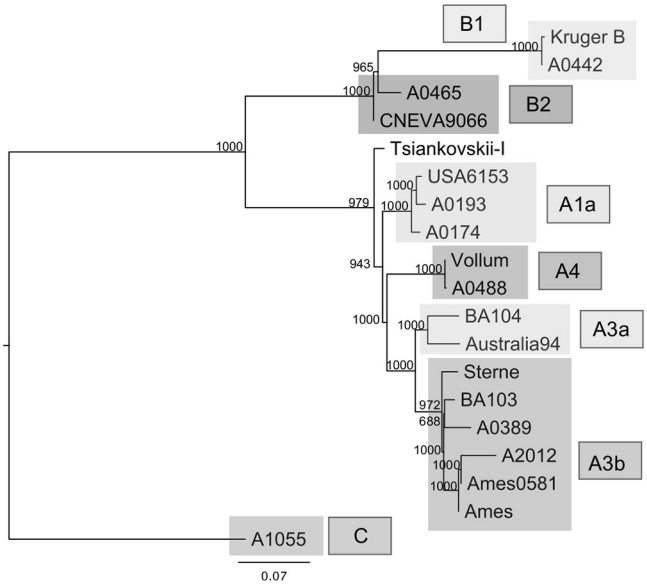

FIG. 2. Obtained whole 2,965 SNP alleles were concatenated into a single nucleotide sequence for each strain and examined by phylogenetic analysis. The indicated cluster was previously defined into a category such as A1a to C by MLVA 25 (4, 13). The scale indicates the nucleotide substitution rate per site.

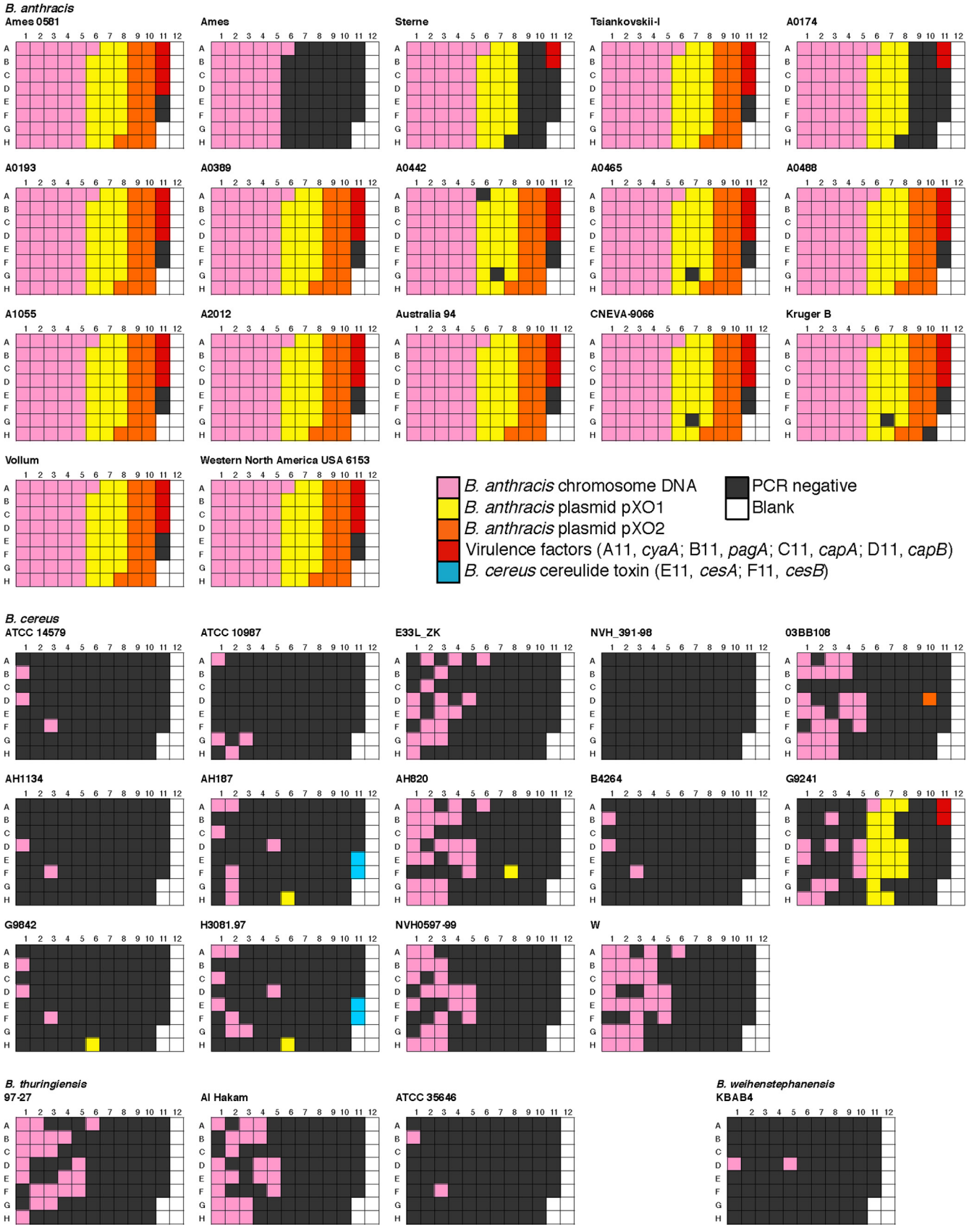

FIG. 3. *In silico* 80-tag SNP PCR amplification. The genomic sequences of all tested strains shown in Table 1 are available from the GenBank database. In addition to the 80-tag SNPs, 6 loci were included as positive controls for anthrax toxins or the *B. cereus* cereulide toxin. The predicted PCRs are shown in a 96-well plate format for each strain. The differential colors represent a positive PCR result at the SNP site located in either the chromosome or the plasmids. All negative PCR results are shown in black.

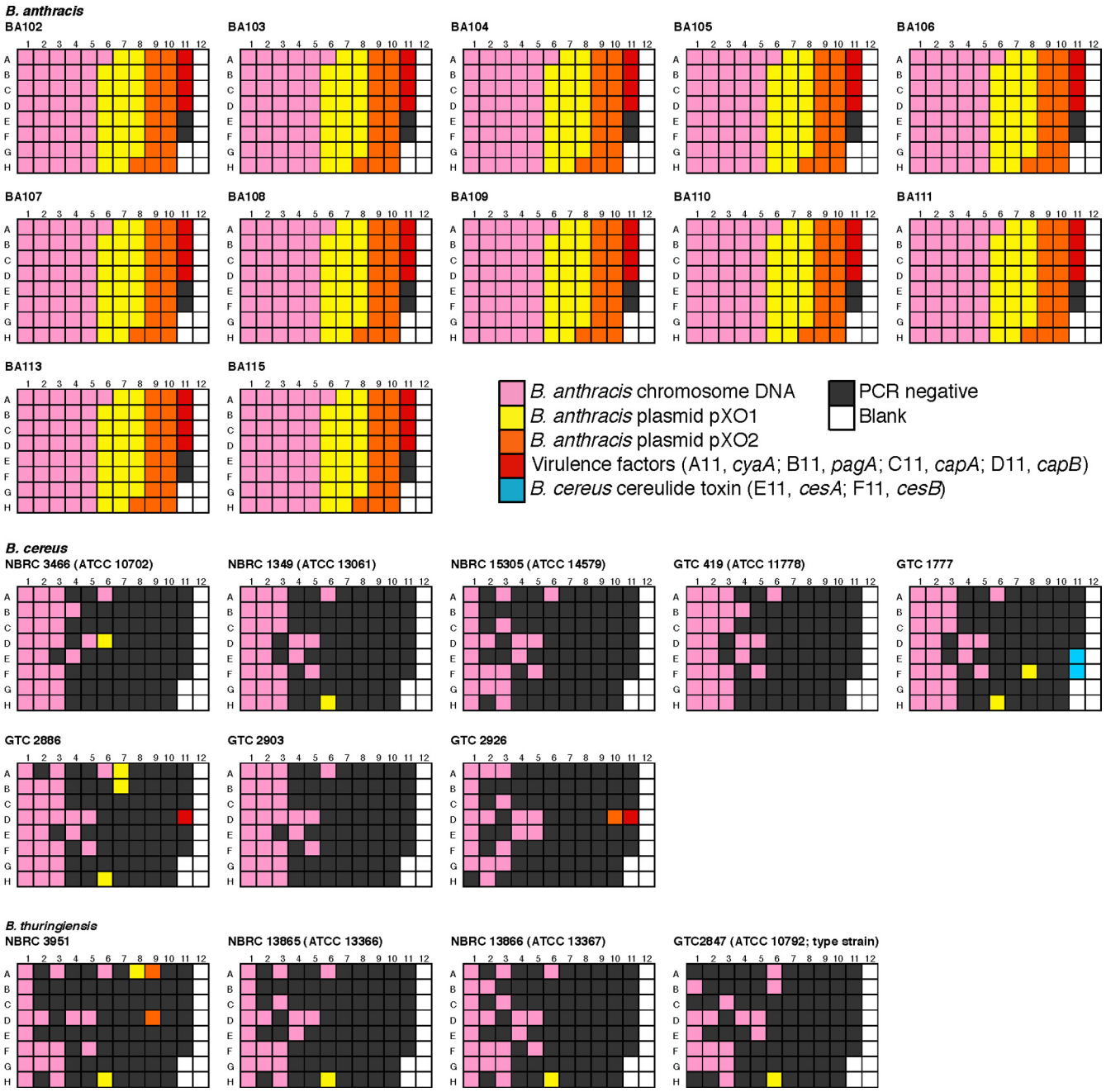

FIG. 4. PCR amplification of 80-tag SNPs in 12 Japanese isolates of *B. anthracis*, 8 *B. cereus* strains, and 4 *B. thuringiensis* strains.

matches of the primer pair; thus, these positive amplifications could be detected by the mispriming of these primers due to the high similarity of genomic sequences in the *B. cereus* group.

Phylogenetic analysis based on 80-tag SNPs among the *B. anthracis* strains. After checking the PCR amplification described above, the PCR products were subjected to DNA sequencing and 80-tag SNP alleles were concatenated into one nucleotide sequence for phylogenetic analysis. Alignment of the concatenated nucleotide sequences was performed using non-gap insertion between nucleotides to order each SNP site (Fig. 5). A phylogenetic tree based on the alignment indicated that the Japanese isolates were classified into two groups: A3a including Australia 94 and A3b including Ames 0581. Basically, the result of the SNP typing corresponded to that of MLVA 25 (16). In addition, a previous study with MLVA 25 could not discriminate between BA106 and BA107, or between BA109 and BA110, because MLVA 25 showed an identical fragment length at 25 VNTR loci (16), while the 80-tag SNP typing revealed that these were distinct strains (Fig. 6). Although whole-genome sequence information is definitely required for the complete

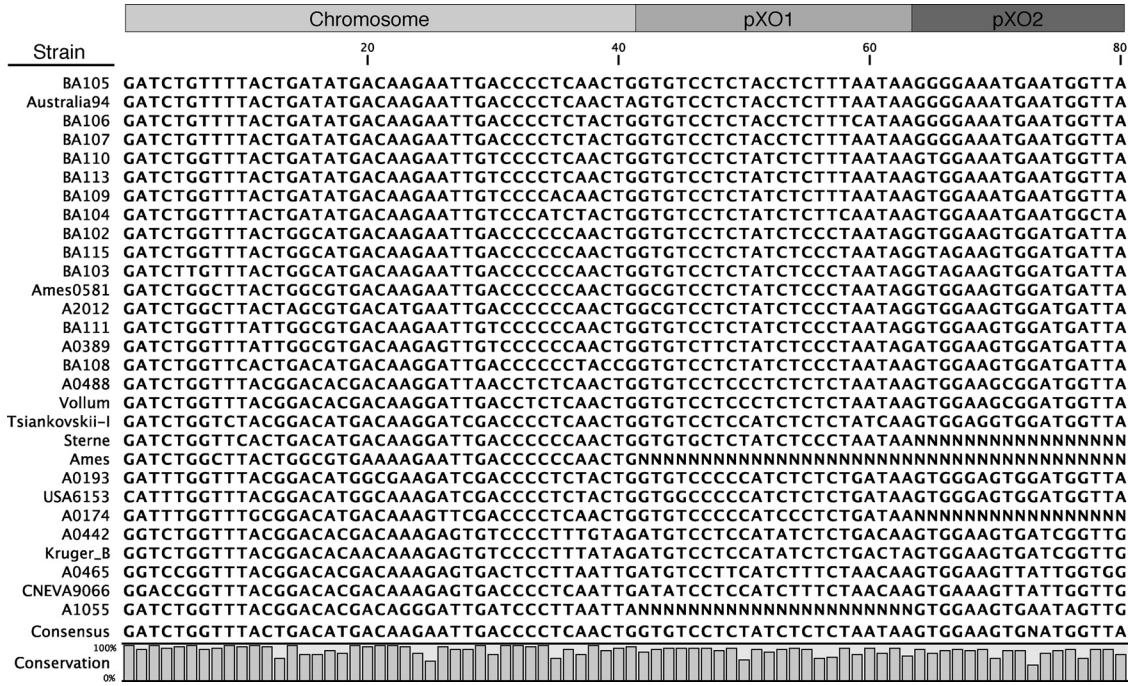

FIG. 5. Alignment of the concatenated sequences from 80-tag SNP alleles of each *B. anthracis* strain. These 80 SNPs consist of 41, 22, and 17 alleles located on the chromosome, pXO1, and pXO2, respectively. The "N" nucleotide indicates no SNP alleles due to the lack of the pXO1 or pXO2 plasmid.

discrimination of the isolates, 80-tag SNP typing was sufficient and effective for strain typing.

DISCUSSION

Following a bioterrorism anthrax attack, strain-specific genetic markers represent crucial information for traceability and have practical implications in reducing the risk of a pandemic. Testing methods using MLVA 25 (13) or canSNPs in combination with MLVA 15 (23, 25) have been reported as rapid strain genotyping systems for *B. anthracis*; however, the MLVA method is complicated given the need to estimate the correct number of repeat units. Indeed, we have experienced that the observed length of fragments obtained using MLVA 25 must be normalized into their actual length by DNA sequencing every time the assay is performed (16). The results by Lista et al. also indicated that the detected length at all 25 loci should be normalized (13); thus, such normalization may cause incorrect processing when assigning strain-specific information for comparison with other strains.

Conversely, SNPs are more definitive genetic markers than the number of repeat units (15, 17); furthermore, every sequencing technique enables the correct call of SNP alleles as a specific genetic marker in every laboratory. However, the investigation of whole SNP alleles on genomic sequences is more laborious than the methods reported above and this was behind our proposal of a simplified 80-tag SNP typing method for the rapid strain typing of *B. anthracis*.

Phylogenetic analysis using whole SNPs generated a corresponding relationship with previous reports using other genotyping methods (Fig. 2). Furthermore, MLVA 25 was able to show that the fragment lengths at 25 loci were identical between BA106 and BA107 (16); however, 80-tag SNP typing could discriminate between these isolates (Fig. 6). These results suggest that the increased testing factors of tag SNP typing might improve the resolution power compared to the 25 loci identified by MLVA; indeed, tag SNP typing more effectively discriminated between 12 Japanese isolates.

A recent striking report suggested that a highly virulent *B. cereus* strain carries anthrax toxins (6), and therefore primary filtering must be extended to detect such potential virulent strains. In addition to strain typing, the 80-tag SNP PCR testing could distinguish *B. anthracis* from other *B. cereus* group species carrying anthrax toxins in the primary PCR amplification step (Fig. 3 and 4). In contrast, PCR primers for 13 loci canSNPs coincide with the genomic sequences of *B. cereus* group species, indicating that the canSNP PCR method is not available for species identification among the group (25). Therefore, our extended testing may facilitate identification of the outbreak strain and allow us to conclude whether it is a local epidemic case (e.g., food poisoning by *B. cereus*) or a suspected bioterrorism case (e.g., *B. anthracis* or other anthrax-like pathogens).

In conclusion, we identified strain-specific SNPs for *B. anthracis* strains by genome-wide screening using short read mapping, and we developed a rapid species-strain typing system using 80-tag SNPs for the primary investigation of an anthrax or anthrax-like outbreak. For further identification, whole SNPs on genomic sequences would be desirable to predict the origin of the strain using entire genetic information. Recent innovations in genetic manipulation may increase the risk of a bioterrorist attack using anthrax or other biological agents;

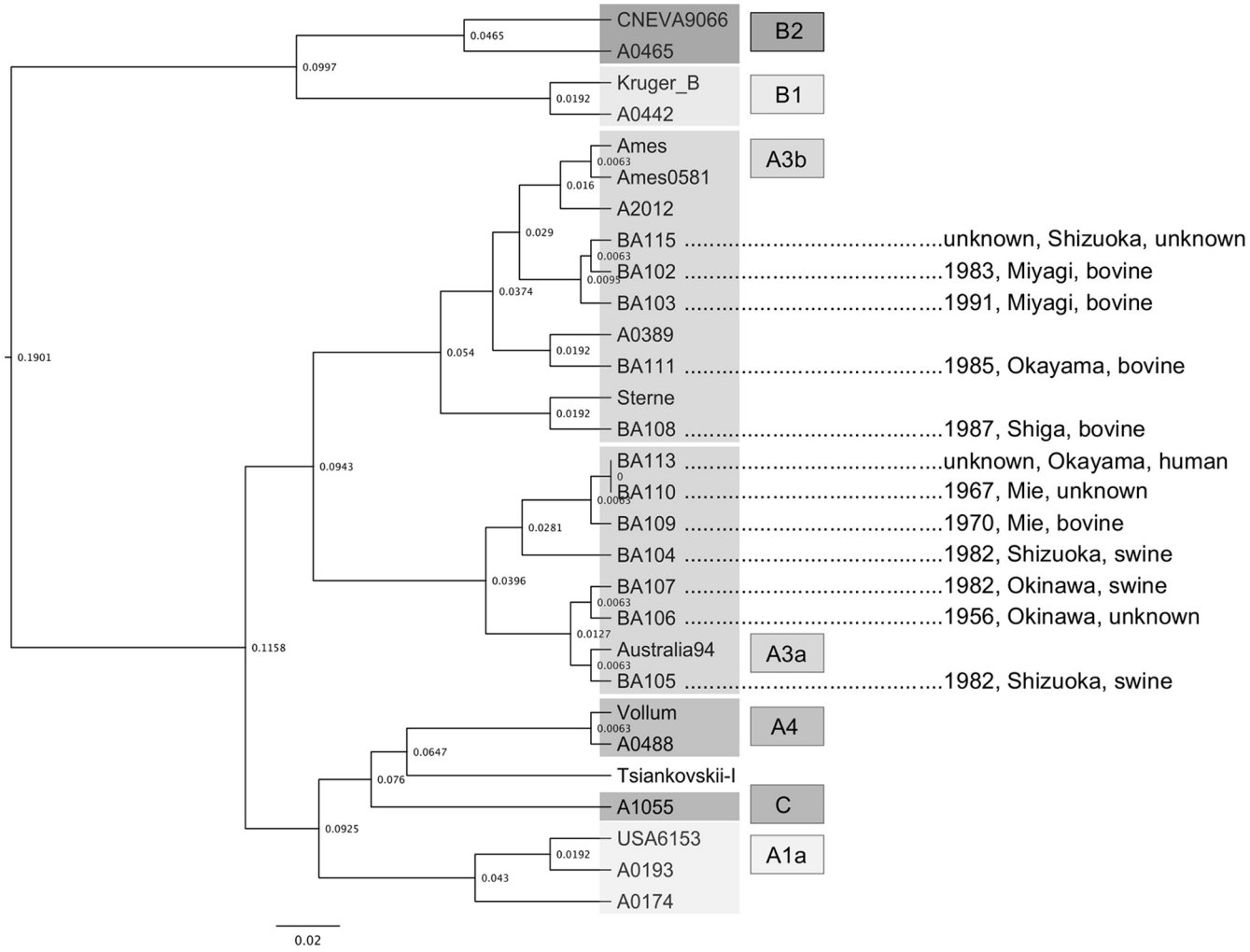

FIG. 6. Phylogenetic analysis of the alignment sequence shown in Fig. 5 using the UPGMA method. The indicated cluster was previously defined into a category such as A1a to C by MLVA 25 (4, 13). Information on the 12 Japanese isolates is shown on the right side of the branch: year isolated, prefecture isolated in Japan, and source. The scale indicates the nucleotide substitution rate per site.

thus, such a rapid and comprehensive analysis system would be indispensable for dealing with bioterrorism attacks and characterizing emerging infectious diseases.

ACKNOWLEDGMENTS

We thank Tadahito Kanda and Akio Yamada for valuable suggestions.

This work was supported by a grant for Research on Emerging and Re-emerging Infectious Diseases (H20 & 21 Shinko-Ippan-6) from the Ministry of Health, Labor, and Welfare, Japan.

REFERENCES

1. **Bao, H., H. Guo, J. Wang, R. Zhou, X. Lu, and S. Shi.** 2009. MapView: visualization of short reads alignment on a desktop computer. Bioinformatics **25:**1554–1555.
2. **Beecher, D. J.** 2006. Forensic application of microbiological culture analysis to identify mail intentionally contaminated with *Bacillus anthracis* spores. Appl. Environ. Microbiol. **72:**5304–5310.
3. **Challacombe, J. F., M. R. Altherr, G. Xie, S. S. Bhotika, N. Brown, D. Bruce, C. S. Campbell, M. L. Campbell, J. Chen, O. Chertkov, C. Cleland, M. Dimitrijevic, N. A. Doggett, J. J. Fawcett, T. Glavina, L. A. Goodwin, L. D. Green, C. S. Han, K. K. Hill, P. Hitchcock, P. J. Jackson, P. Keim, A. R. Kewalramani, J. Longmire, S. Lucas, S. Malfatti, D. Martinez, K. McMurry, L. J. Meincke, M. Misra, B. L. Moseman, M. Mundt, A. C. Munk, R. T. Okinaka, B. Parson-Quintana, L. P. Reilly, P. Richardson, D. L. Robinson, E. Saunders, R. Tapia, J. G. Tesmer, N. Thayer, L. S. Thompson, H. Tice, L. O. Ticknor, P. L. Wills, P. Gilna, and T. S. Brettin.** 2007. The complete genome sequence of *Bacillus thuringiensis* Al Hakam. J. Bacteriol. **189:**3680–3681.
4. **Ciammaruconi, A., S. Grassi, R. De Santis, G. Faggioni, V. Pittiglio, R. D'Amelio, A. Carattoli, A. Cassone, G. Vergnaud, and F. Lista.** 2008. Fieldable genotyping of *Bacillus anthracis* and *Yersinia pestis* based on 25-loci multi locus VNTR analysis. BMC Microbiol. **8:**21.
5. **Han, C. S., G. Xie, J. F. Challacombe, M. R. Altherr, S. S. Bhotika, N. Brown, D. Bruce, C. S. Campbell, M. L. Campbell, J. Chen, O. Chertkov, C. Cleland, M. Dimitrijevic, N. A. Doggett, J. J. Fawcett, T. Glavina, L. A. Goodwin, L. D. Green, K. K. Hill, P. Hitchcock, P. J. Jackson, P. Keim, A. R. Kewalramani, J. Longmire, S. Lucas, S. Malfatti, K. McMurry, L. J. Meincke, M. Misra, B. L. Moseman, M. Mundt, A. C. Munk, R. T. Okinaka, B. Parson-Quintana, L. P. Reilly, P. Richardson, D. L. Robinson, E. Rubin, E. Saunders, R. Tapia, J. G. Tesmer, N. Thayer, L. S. Thompson, H. Tice, L. O. Ticknor, P. L. Wills, T. S. Brettin, and P. Gilna.** 2006. Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*. J. Bacteriol. **188:**3382–3390.
6. **Hoffmaster, A. R., J. Ravel, D. A. Rasko, G. D. Chapman, M. D. Chute, C. K. Marston, B. K. De, C. T. Sacchi, C. Fitzgerald, L. W. Mayer, M. C. Maiden, F. G. Priest, M. Barker, L. Jiang, R. Z. Cer, J. Rilstone, S. N. Peterson, R. S. Weyant, D. R. Galloway, T. D. Read, T. Popovic, and C. M. Fraser.** 2004. Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. Proc. Natl. Acad. Sci. U. S. A. **101:**8449–8454.
7. **Inglesby, T. V., T. O'Toole, D. A. Henderson, J. G. Bartlett, M. S. Ascher, E.**

Eitzen, A. M. Friedlander, J. Gerberding, J. Hauer, J. Hughes, J. McDade, M. T. Osterholm, G. Parker, T. M. Perl, P. K. Russell, and K. Tonat. 2002. Anthrax as a biological weapon, 2002: updated recommendations for management. JAMA **287:**2236–2252.

8. **Ivanova, N., A. Sorokin, I. Anderson, N. Galleron, B. Candelon, V. Kapatral, A. Bhattacharyya, G. Reznik, N. Mikhailova, A. Lapidus, L. Chu, M. Mazur, E. Goltsman, N. Larsen, M. D'Souza, T. Walunas, Y. Grechkin, G. Pusch, R. Haselkorn, M. Fonstein, S. D. Ehrlich, R. Overbeek, and N. Kyrpides.** 2003. Genome sequence of *Bacillus cereus* and comparative analysis with *Bacillus anthracis*. Nature **423:**87–91.
9. **Jones, S. W., M. E. Dobson, S. C. Francesconi, R. Schoske, and R. Crawford.** 2005. DNA assays for detection, identification, and individualization of select agent microorganisms. Croat. Med. J. **46:**522–529.
10. **Keim, P., K. L. Smith, C. Keys, H. Takahashi, T. Kurata, and A. Kaufmann.** 2001. Molecular investigation of the Aum Shinrikyo anthrax release in Kameido, Japan. J. Clin. Microbiol. **39:**4566–4567.
11. **Larkin, M. A., G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins.** 2007. Clustal W and Clustal X version 2.0. Bioinformatics **23:**2947–2948.
12. **Li, H., J. Ruan, and R. Durbin.** 2008. Mapping short DNA sequencing reads and calling variants using mapping quality scores. Genome Res. **18:**1851–1858.
13. **Lista, F., G. Faggioni, S. Valjevac, A. Ciammaruconi, J. Vaissaire, C. le Doujet, O. Gorge, R. De Santis, A. Carattoli, A. Ciervo, A. Fasanella, F. Orsini, R. D'Amelio, C. Pourcel, A. Cassone, and G. Vergnaud.** 2006. Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. BMC Microbiol. **6:**33.
14. **Mock, M., and A. Fouet.** 2001. Anthrax. Annu. Rev. Microbiol. **55:**647–671.
15. **Okinaka, R. T., M. Henrie, K. K. Hill, K. S. Lowery, M. Van Ert, T. Pearson, J. Schupp, L. Kenefic, J. Beaudry, S. A. Hofstadler, P. J. Jackson, and P. Keim.** 2008. Single nucleotide polymorphism typing of *Bacillus anthracis* from Sverdlovsk tissue. Emerg. Infect. Dis. **14:**653–656.
16. **Okutani, A., T. Sekizuka, B. Boldbaatar, A. Yamada, M. Kuroda, and S. Inoue.** 2010. Phylogenetic typing of *Bacillus anthracis* isolated in Japan by multiple locus variable-number tandem repeats and the comprehensive single nucleotide polymorphism. J. Vet. Med. Sci. **72:**93–97.
17. **Pearson, T., J. D. Busch, J. Ravel, T. D. Read, S. D. Rhoton, J. M. U'Ren, T. S. Simonson, S. M. Kachur, R. R. Leadem, M. L. Cardon, M. N. Van Ert, L. Y. Huynh, C. M. Fraser, and P. Keim.** 2004. Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing. Proc. Natl. Acad. Sci. U. S. A. **101:**13536–13541.
18. **Rasko, D. A., M. R. Altherr, C. S. Han, and J. Ravel.** 2005. Genomics of the *Bacillus cereus* group of organisms. FEMS Microbiol. Rev. **29:**303–329.
19. **Rasko, D. A., J. Ravel, O. A. Okstad, E. Helgason, R. Z. Cer, L. Jiang, K. A. Shores, D. E. Fouts, N. J. Tourasse, S. V. Angiuoli, J. Kolonay, W. C. Nelson, A. B. Kolsto, C. M. Fraser, and T. D. Read.** 2004. The genome sequence of *Bacillus cereus* ATCC 10987 reveals metabolic adaptations and a large plasmid related to *Bacillus anthracis* pXO1. Nucleic Acids Res. **32:**977–988.
20. **Ravel, J., L. Jiang, S. T. Stanley, M. R. Wilson, R. S. Decker, T. D. Read, P. Worsham, P. S. Keim, S. L. Salzberg, C. M. Fraser-Liggett, and D. A. Rasko.** 2009. The complete genome sequence of *Bacillus anthracis* Ames "Ancestor." J. Bacteriol. **191:**445–446.
21. **Read, T. D., S. N. Peterson, N. Tourasse, L. W. Baillie, I. T. Paulsen, K. E. Nelson, H. Tettelin, D. E. Fouts, J. A. Eisen, S. R. Gill, E. K. Holtzapple, O. A. Okstad, E. Helgason, J. Rilstone, M. Wu, J. F. Kolonay, M. J. Beanan, R. J. Dodson, L. M. Brinkac, M. Gwinn, R. T. DeBoy, R. Madpu, S. C. Daugherty, A. S. Durkin, D. H. Haft, W. C. Nelson, J. D. Peterson, M. Pop, H. M. Khouri, D. Radune, J. L. Benton, Y. Mahamoud, L. Jiang, I. R. Hance, J. F. Weidman, K. J. Berry, R. D. Plaut, A. M. Wolf, K. L. Watkins, W. C. Nierman, A. Hazen, R. Cline, C. Redmond, J. E. Thwaite, O. White, S. L. Salzberg, B. Thomason, A. M. Friedlander, T. M. Koehler, P. C. Hanna, A. B. Kolsto, and C. M. Fraser.** 2003. The genome sequence of *Bacillus anthracis* Ames and comparison to closely related bacteria. Nature **423:**81–86.
22. **Read, T. D., S. L. Salzberg, M. Pop, M. Shumway, L. Umayam, L. Jiang, E. Holtzapple, J. D. Busch, K. L. Smith, J. M. Schupp, D. Solomon, P. Keim, and C. M. Fraser.** 2002. Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis*. Science **296:**2028–2033.
23. **Simonson, T. S., R. T. Okinaka, B. Wang, W. R. Easterday, L. Huynh, J. M. U'Ren, M. Dukerich, S. R. Zanecki, L. J. Kenefic, J. Beaudry, J. M. Schupp, T. Pearson, D. M. Wagner, A. Hoffmaster, J. Ravel, and P. Keim.** 2009. *Bacillus anthracis* in China and its relationship to worldwide lineages. BMC Microbiol. **9:**71.
24. **Srivatsan, A., Y. Han, J. Peng, A. K. Tehranchi, R. Gibbs, J. D. Wang, and R. Chen.** 2008. High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. PLoS Genet. **4:**e1000139.
25. **Van Ert, M. N., W. R. Easterday, L. Y. Huynh, R. T. Okinaka, M. E. Hugh-Jones, J. Ravel, S. R. Zanecki, T. Pearson, T. S. Simonson, J. M. U'Ren, S. M. Kachur, R. R. Leadem-Dougherty, S. D. Rhoton, G. Zinser, J. Farlow, P. R. Coker, K. L. Smith, B. Wang, L. J. Kenefic, C. M. Fraser-Liggett, D. M. Wagner, and P. Keim.** 2007. Global genetic population structure of *Bacillus anthracis*. PLoS One **2:**e461.
26. **Van Ert, M. N., W. R. Easterday, T. S. Simonson, J. M. U'Ren, T. Pearson, L. J. Kenefic, J. D. Busch, L. Y. Huynh, M. Dukerich, C. B. Trim, J. Beaudry, A. Welty-Bernard, T. Read, C. M. Fraser, J. Ravel, and P. Keim.** 2007. Strain-specific single-nucleotide polymorphism assays for the *Bacillus anthracis* Ames strain. J. Clin. Microbiol. **45:**47–53.