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,

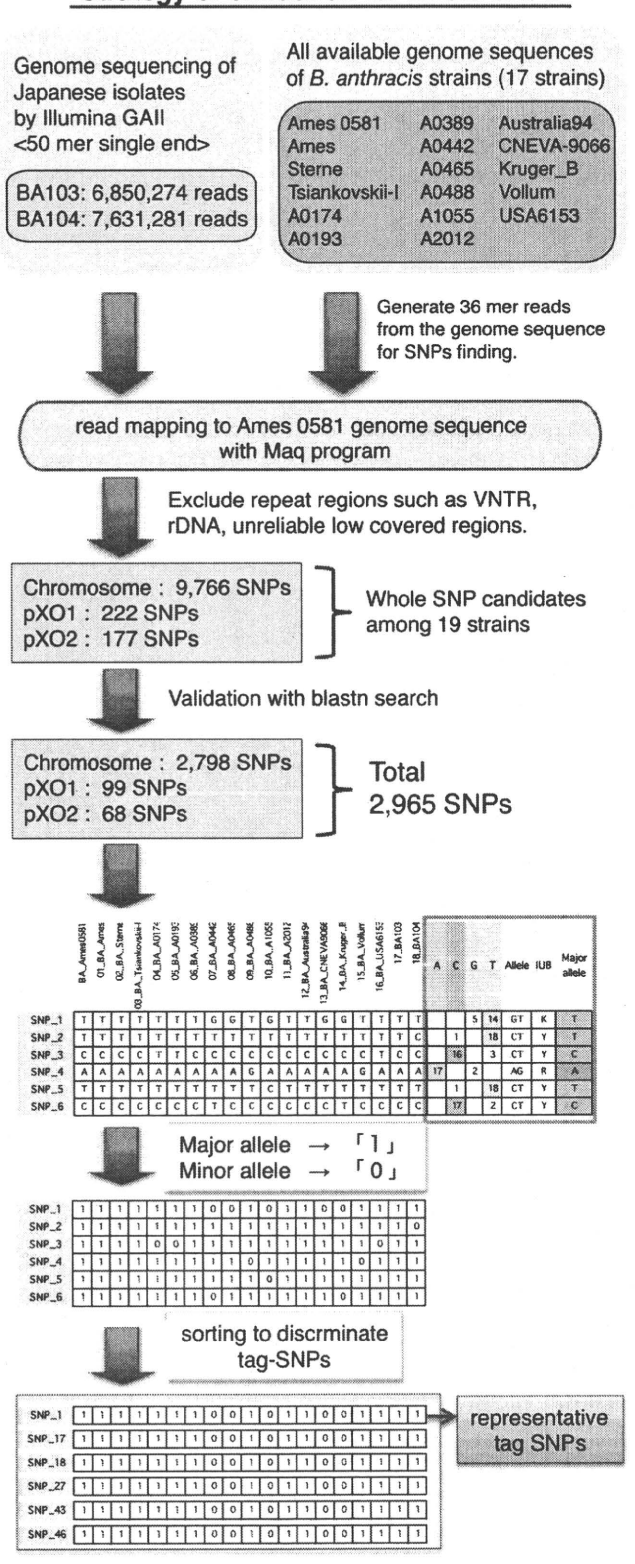

FIG. 1. Schematic representation of the extraction of whole SNPs from the genomic sequences of *B. anthracis* strains. In total, 19 strains including 2 Japanese isolates in this study and 17 available genomic sequences were used for SNP extraction.

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

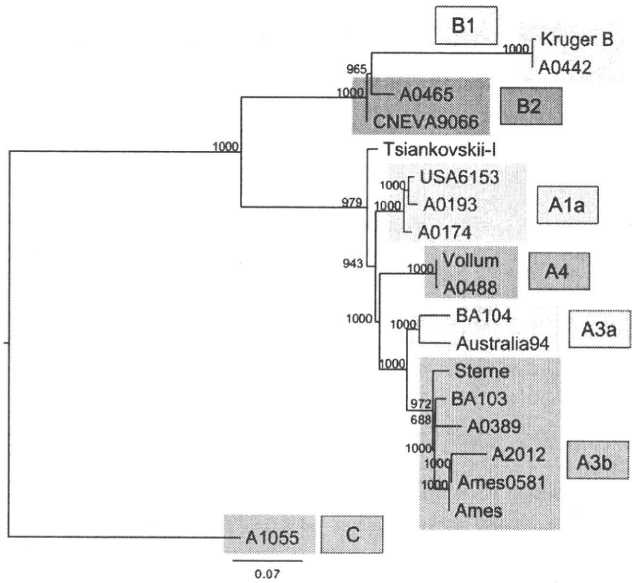

FIG. 2. Obtained whole 2,965 SNP alleles were concatenated into a single nucleotide sequence for each strain and examined by phylogenetic analysis. The indicated cluster was previously defined into a category such as A1a to C by MLVA 25 (4, 13). The scale indicates the nucleotide substitution rate per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

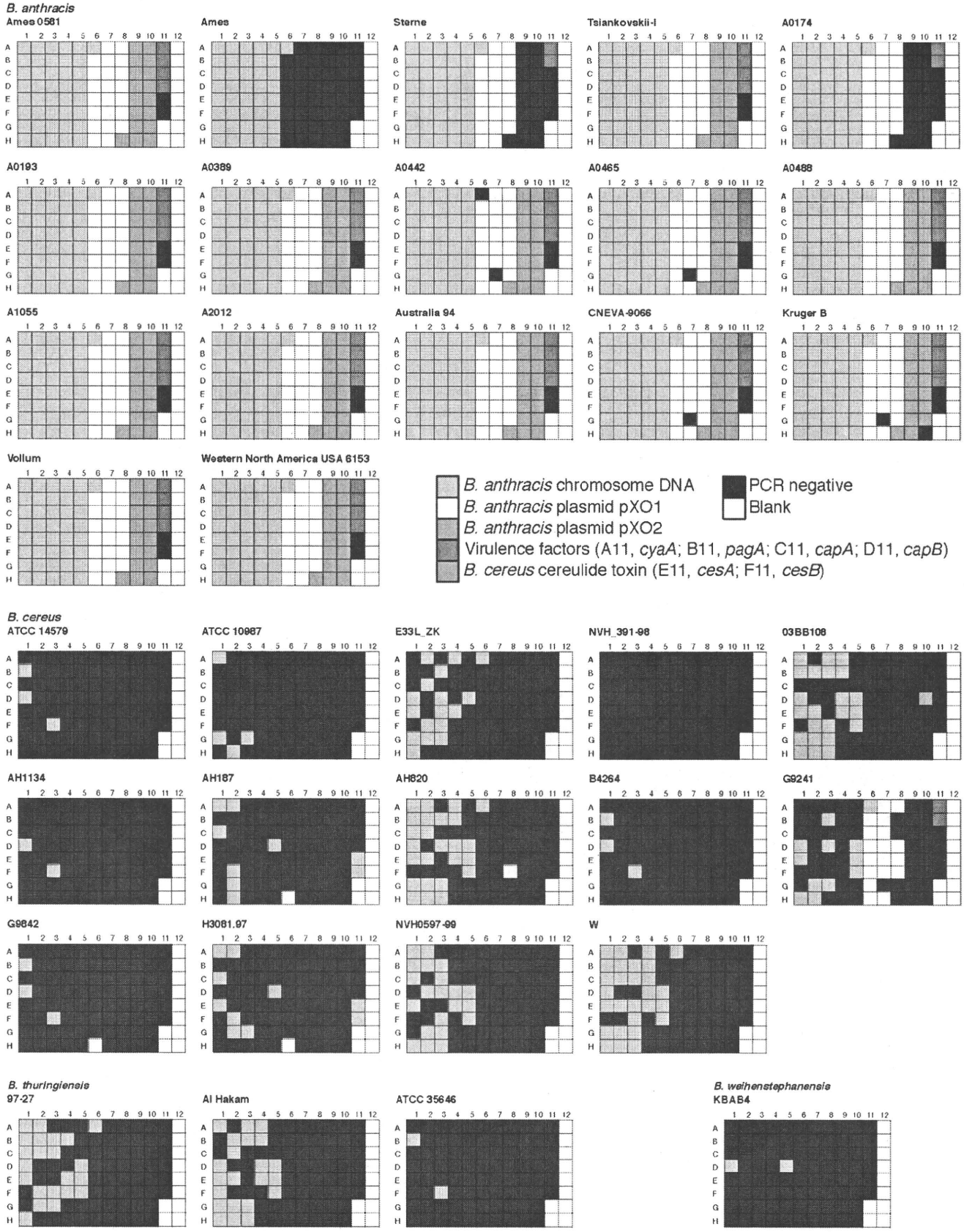

FIG. 3. *In silico* 80-tag SNP PCR amplification. The genomic sequences of all tested strains shown in Table 1 are available from the GenBank database. In addition to the 80-tag SNPs, 6 loci were included as positive controls for anthrax toxins or the *B. cereus* cereulide toxin. The predicted PCRs are 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.

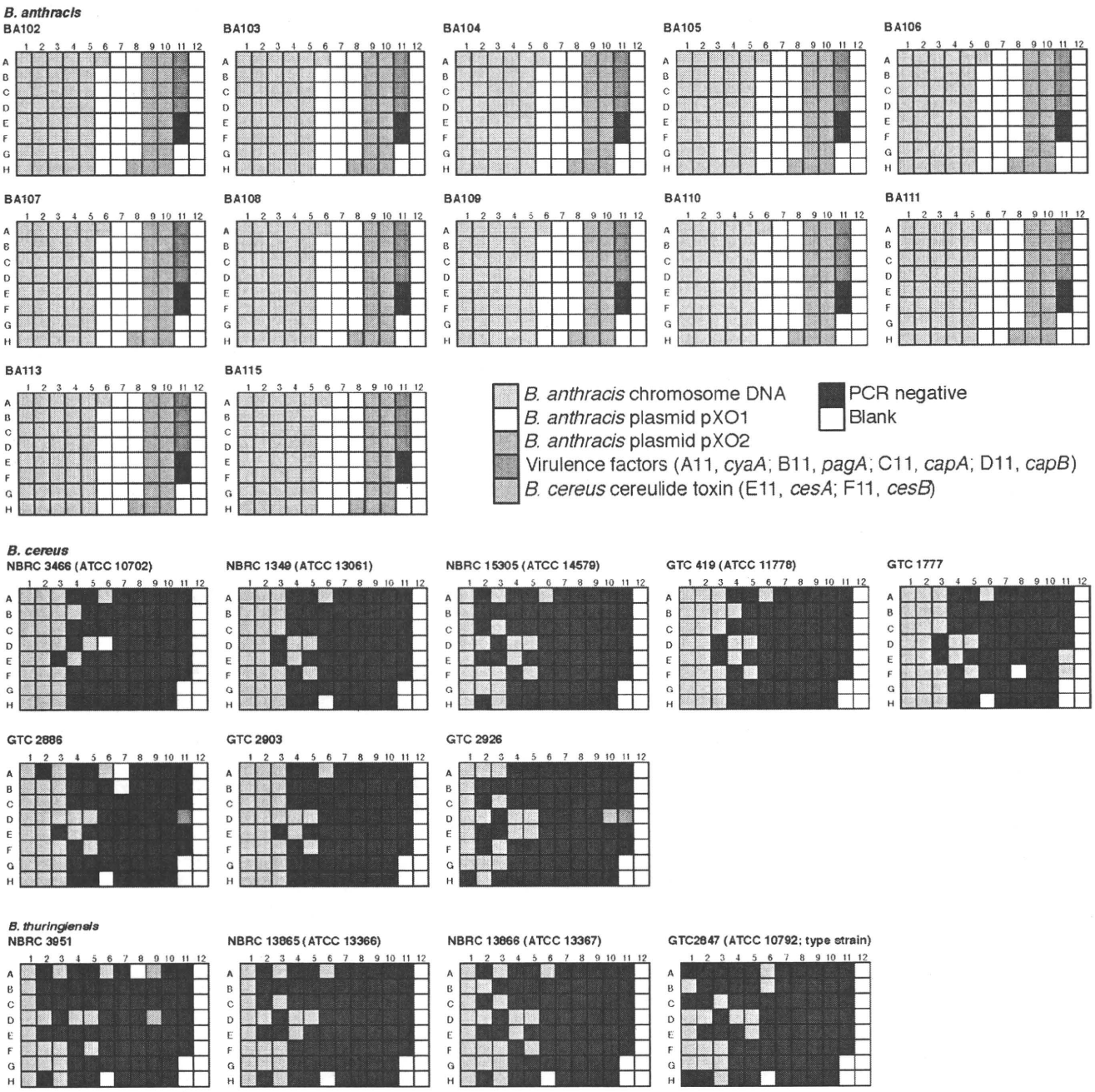

FIG. 4. PCR amplification of 80-tag SNPs in 12 Japanese isolates of *B. anthracis*, 8 *B. cereus* strains, and 4 *B. thuringiensis* strains.

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

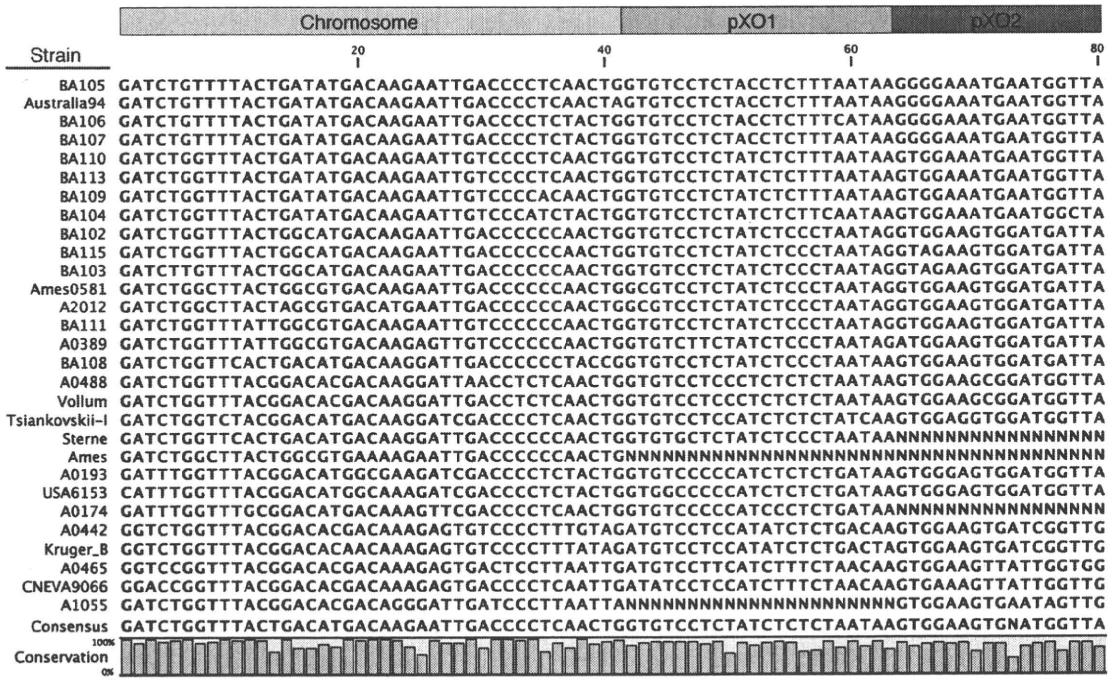

FIG. 5. Alignment of the concatenated sequences from 80-tag SNP alleles of each *B. anthracis* strain. These 80 SNPs consist of 41, 22, and 17 alleles located on the chromosome, pXO1, and pXO2, respectively. The "N" nucleotide indicates no SNP alleles due to the lack of the pXO1 or pXO2 plasmid.

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;

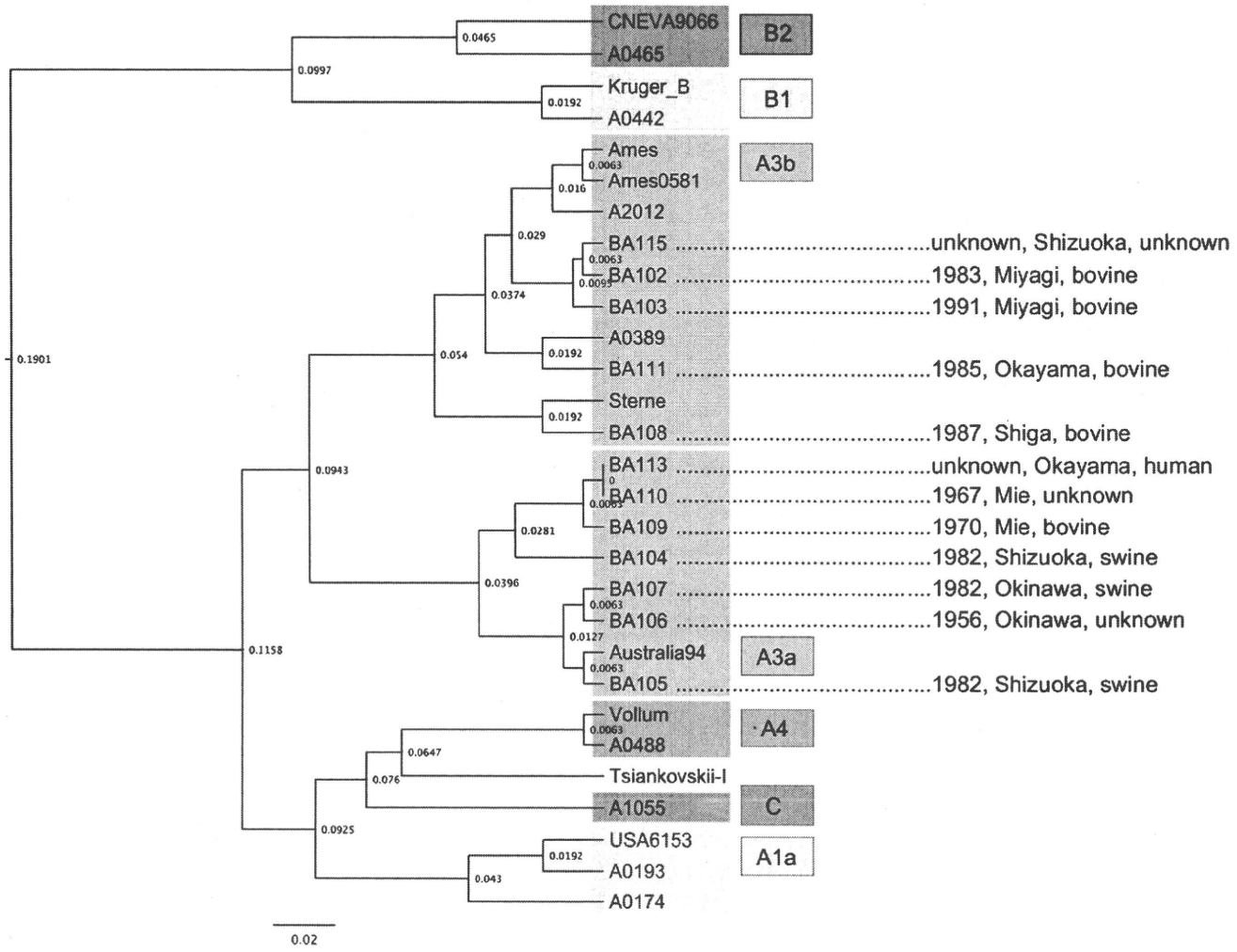

FIG. 6. Phylogenetic analysis of the alignment sequence shown in Fig. 5 using the UPGMA method. The indicated cluster was previously defined into a category such as A1a to C by MLVA 25 (4, 13). Information on the 12 Japanese isolates is shown on the right side of the branch: year isolated, prefecture 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. Field-able 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.

ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, July 2010, p. 2787–2792 Vol. 54, No. 7
0066-4804/10/$12.00 doi:10.1128/AAC.01405-09

Genomewide Screening for Novel Genetic Variations Associated with Ciprofloxacin Resistance in *Bacillus anthracis*▽†

Masakuni Serizawa,[1] Tsuyoshi Sekizuka,[1] Akiko Okutani,[2] Satomi Banno,[1] Tetsutaro Sata,[3] Satoshi Inoue,[2] and Makoto Kuroda[1]*

Laboratory of Bacterial Genomics, Pathogen Genomics Center,[1] and Departments of Veterinary Science[2] and Pathology,[3] National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan

Received 6 October 2009/Returned for modification 28 February 2010/Accepted 19 March 2010

Fluoroquinolone (FQ) resistance of *Bacillus anthracis* is a serious concern in the fields of biodefense and bioterrorism since FQs are very effective antibiotics and are recommended as first-line treatment against this lethal bacterium. In this study, we obtained 2 strains of *B. anthracis* showing resistance or intermediate resistance to ciprofloxacin (CIP) by a stepwise selection procedure with increasing CIP concentrations. Fifteen genetic variations were identified between the parental and CIP-resistant strains by next-generation sequencing. Nonsynonymous mutations in the quinolone resistance-determining region (QRDR) of type II DNA topoisomerase were identified in the resistant strain but not in the intermediate-resistant strain. The GBAA0834 (TetR-type transcriptional regulator) locus was also revealed to be a novel "mutation hot spot" that leads to the increased expression of multidrug efflux systems for CIP resistance. As an initial step of CIP resistance in *B. anthracis*, such disruptive mutations of GBAA0834 appear to be more easily acquired than those in an essential gene, such as that encoding type II DNA topoisomerase. Such an intermediate-resistant phenotype could increase a cell population under CIP-selective pressure and might promote the emergence of highly resistant isolates. Our findings reveal, in addition to QRDR, crucial genetic targets for the investigation of intermediate resistance of *B. anthracis* to FQs.

Anthrax caused by the spore-forming bacterium *Bacillus anthracis* is one of the most severe zoonoses and poses a serious threat to both public and animal health (9, 15). *B. anthracis* belongs to the *Bacillus cereus* group of bacteria, which comprises closely related Gram-positive organisms with highly divergent virulent properties (15, 20). Infection with this bacterium can occur through the skin, gastrointestinal tract, or respiratory apparatus following contact, ingestion, or inhalation of spores, respectively (9, 15). The recent "postal anthrax attack" in the United States that aimed to intentionally release *B. anthracis* spores underlines the growing importance of the identification of *B. anthracis* at the strain level in forensic and epidemiological investigations (3, 21).

Fluoroquinolone (FQ) resistance is a major concern in medical treatment following anthrax bioterrorism because FQs are first-line antibiotics for the treatment of *B. anthracis* infection (9, 27). FQs act as broad-spectrum bactericidal antibiotics by inhibiting type II DNA topoisomerases, DNA gyrases (GyrA and GyrB), and type IV DNA topoisomerases (ParC and ParE). The mechanism responsible for FQ resistance has been well documented with bacteria, in which frequent mutations of topoisomerase genes have been identified and designated the quinolone resistance-determining region (QRDR) (22, 23), the most fundamental region for a primary investigation of FQ resistance.

Recent studies have reported several mutations in QRDR in *in vitro*-selected FQ-resistant strains of *B. anthracis* (2, 8, 18). Aside from QRDR, multidrug resistance proteins are reportedly involved in the development of FQ resistance by scavenging intracellular FQ (13, 22, 23). In fact, recent reports have also suggested a possible contribution of multidrug efflux pumps to FQ resistance in *B. anthracis* (2, 18). However, the responsible efflux pump has not yet been identified, and thus, mutations responsible for additional resistance mechanisms remain to be elucidated. Recently, next-generation sequencing has facilitated the identification of genomewide single nucleotide polymorphisms (SNPs) and insertions-deletions (indels) (4, 11, 28). Using such a method, the present study reveals novel genetic variations responsible for FQ resistance in *B. anthracis*.

MATERIALS AND METHODS

Selection and isolation of CIP-resistant mutants. Japanese isolates of *B. anthracis* BA103 or BA104 were used as the parental strains. The BA103 strain was isolated from beef cattle in 1991, while BA104 was isolated from pigs in a sporadic incident in 1982. CIP-resistant mutants were isolated by serial passage on Muller-Hinton (MH) agar supplemented with 2-fold stepwise increasing concentrations of ciprofloxacin (CIP) (Wako, Osaka, Japan), from 0.0625 to 16 mg/liter CIP at 37°C. The MIC to antibiotics was determined using Etest (AB Biodisk, Solna, Sweden) according to the manufacturer's instructions.

The handling of live *B. anthracis* was performed at biosafety level 3 (BSL-3) according to the Regulations on the Safety Control of Laboratories Handling Pathogenic Agents (SCLHPA) authorized by the biorisk management committee of the National Institute of Infectious Diseases (NIID), Japan. The management of group 2 agents, such as *B. anthracis*, was also regulated and required obligatory special permission, special biosafety facilities, and other measures for its possession and use as stated in the Law Concerning the Prevention of Infectious

* 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://aac.asm.org/.

▽ Published ahead of print on 12 April 2010.

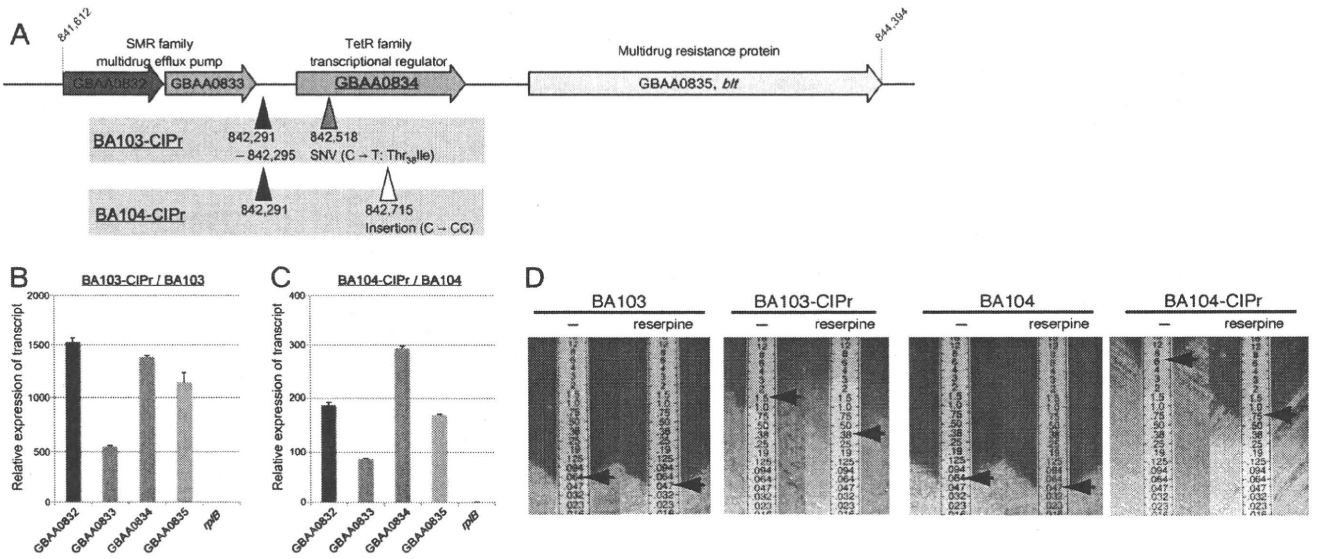

FIG. 1. (A) Schematic gene organization around the GBAA0834 locus. Vertical arrowheads indicate positions of the mutations (black, deletion mutation; white, insertion mutation; gray, SNV). The detailed sequence is shown in Fig. S4 in the supplemental material. Numbers indicate the genomic position of the Ames 0581 strain. Relative transcriptional expression around GBAA0834 in BA103 (B) and BA104 (C) derivatives. (D) Testing of susceptibility to CIP on Muller-Hinton agar containing 10 mg/liter reserpine, which acts by blocking the multidrug efflux protein. Dimethyl sulfoxide was used as the control solvent (−). Arrows indicate MICs for CIP.

Diseases and Medical Care for Patients of Infections (the Infectious Diseases Control Law) by the Ministry of Health, Labor, and Welfare.

Short sequencing reads using GA II. Library preparation was performed using a genomic DNA sample prep kit (Illumina, San Diego, CA), and DNA clusters were generated on a slide using the cluster generation kit (version 2) on an Illumina cluster station (Illumina), according to the manufacturer's instructions. To obtain ~10 million clusters for one lane, the general procedure as described in the standard recipe (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 Illumina genome analyzer II (GA II) using the Illumina sequencing kit (version 3). Fluorescent images were analyzed with the Illumina base-calling pipeline 1.3.2 to obtain FASTQ formatted sequence data.

Identification of single nucleotide variations (SNVs) and indels specific to *in vitro*-selected CIP-resistant *B. anthracis* mutants. A schematic flow chart of the data processing procedure is shown in Fig. S2 in the supplemental material. To identify specific SNVs/indels compared with the reference sequence of *B. anthracis* Ames 0581 (GenBank accession number NC_007530), Maq software (version 0.7.1) (12), a mapping assembler for short reads generated by the next-generation sequencer, was used with the easyrun Perl command as a default parameter. Read alignment for the validation of SNVs/indels was performed using the MapView graphical alignment viewer (1). Strain-specific SNVs/indels were extracted from the "cns.final.snp" or "cns.indelse" file, respectively, by comparison of the parental or CIP-resistant strains to Ames 0581. SNVs located in repetitive sequence regions (e.g., variable-number tandem repeat [VNTR], rRNA, and insertion sequence [IS]) were excluded from the analysis. Regarding the reliable detection of indels, the cns.indelse file, which includes potential indels showing abnormal alignment patterns, was processed by the recommended filtering technique to reduce the number of false positives, as described in the manual. Furthermore, we analyzed the 100 nucleotides surrounding the indels with a BlastN search against all sequenced short reads (parameter −F F −e 1.0E−10 −m 3). Putative SNVs/indels were finally verified by Sanger sequencing using the BigDye terminator version 3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) (see Fig. S6 in the supplemental material). PCR and sequence primers are listed in Table S2 in the supplemental material.

qRT-PCR analysis. For the preparation of total RNA from *B. anthracis*, bacterial cells were cultured in brain heart infusion broth at 37°C to the mid-log phase, and then total RNA was extracted with Trizol reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. Total RNA was treated with the turbo DNA-free kit (Ambion, TX). Quantitative reverse transcription-PCR (qRT-PCR) was performed using 100 ng of total RNA, gene-specific primers designed by PrimerQuest (see Table S3 in the supplemental material), and the SuperScript III Platinum SYBR green one-step qRT-PCR kit with ROX (Invitrogen) and analyzed using the ABI Prism 7900HT real-time PCR system (Applied Biosystems). We used the following qRT-PCR program: RT reaction, 50°C for 3 min; initial denaturation, 95°C for 5 min; and 2 steps of amplification (40 cycles), 95°C for 15 s and 60°C for 30 s. The *rplB* gene encoding the 50S ribosomal protein L2 was used as the internal control.

Inhibition assay of multidrug resistance efflux system with reserpine. The MIC was determined using a CIP Etest (AB Biodisk) on MH agar containing 10 mg/liter reserpine (Sigma-Aldrich, St. Louis, MO) (14). Dimethyl sulfoxide was used as the control solvent.

Nucleotide sequence accession numbers. The short-read archives have been deposited in the DNA Data Bank of Japan (DDBJ; accession numbers DRA000067, DRA000068, DRA000069 and DRA000070 for BA103, BA104, BA103-CIPr, and BA104-CIPr, respectively).

RESULTS

Susceptibility testing for *in vitro*-selected CIP-resistant mutants of *B. anthracis*. Both BA103 and BA104 were susceptible to CIP at 0.064 mg/liter, the MIC as determined using Etest (Fig. 1D). Resistant mutants were identified by repeating the selection procedure until a final concentration of 16.0 μg/ml CIP. Two CIP-resistant mutants, BA103-CIPr (MIC, 1.5 mg/liter) and BA104-CIPr (MIC, 6 mg/liter), were isolated from BA103 and BA104, respectively. Their CIP resistance levels increased by 23- and 94-fold, respectively (Fig. 1D). The two selected mutants also showed reduced susceptibility to another quinolone antibiotic, levofloxacin (see Fig. S1 in the supplemental material): for BA103 and BA104, the MIC was 0.047 mg/liter; for BA103-CIPr, 0.125 mg/liter; and for BA104-CIPr, 0.50 mg/liter. Cross-resistance to other antibiotics was not observed with the selected mutants: for penicillin, the MIC was 0.023 mg/liter; for tetracycline, <0.016 mg/liter; and for amikacin, 0.25 mg/liter (see Fig. S1 in the supplemental material).

Summary of sequence reads and coverage. The Illumina GA II sequencer produced 6.85 to 10.31 million 50-base-long reads per strain after applying the quality filter of the Illumina base-

TABLE 1. SNVs specific to the derivative, CIP-resistant mutants of BA103 and BA104

Derivative[a]	Genomic position of SNV[b]	Detected SNV[c]			Type of DNA substitution	Amino acid substitution	Read depth[d,e]		Genetic information about SNVs		
		Ames 0581	Parental strain[d]	Mutant strain[d]			Parental strain	Mutant strain	Locus tag[f]	Product	Strand[g]
BA103	842518	A**C**T	A**C**T	A**T**T	Transition	Thr38Ile	88	173	GBAA0834	TetR family transcriptional regulator	+
BA104	6850	**G**CT	**G**CT	**C**CT	Transversion	Ala86Pro	170	183	GBAA0006 (*gyrA*)	DNA gyrase subunit A	+
	2138798	**A**AA	**A**AA	**G**AA	Transition	Lys80Glu	5	8	GBAA2291	Sensor histidine kinase	+
	4595343	T	T	C	Transition		75	25	Intergenic region (GBAA5072–GBAA5074)		

[a] MICs (mg/liter) of ciprofloxacin for parent strains and CIP-resistant mutants were 0.064 and 1.5, respectively, for BA103 and 0.064 and 6, respectively, for BA104.
[b] Numbers indicate the genomic position in the sequence of the *B. anthracis* Ames 0581 strain (GenBank accession number NC_007530).
[c] Bold and underlined nucleotide indicates the SNV position.
[d] "Mutant strain" indicates CIP-resistant mutant generated from each "parent strain" by *in vitro* selection in this study.
[e] Read depth indicates the number of short reads generated by the Illumina genome analyzer II, which mapped to each SNV position.
[f] Locus tags are described as shown in the Comprehensive Microbial Resource (CMR) and Entrez Gene. As for the intergenic region, genes located adjacent to the SNV are shown.
[g] Clockwise and counterclockwise gene directions are indicated by "+" and "−," respectively.

calling pipeline 1.32 (see Table S1 in the supplemental material). More than 94.4% of the filter-passed reads were aligned to the reference sequences of the Ames 0581 strain using Maq software, resulting in more than a 56-fold coverage depth on average and a 92.7 to 99.34% coverage of the sequence, excluding ambiguous repetitive regions (Table S1 in the supplemental material). Since genomic analysis of laboratory strains of *B. subtilis* using Illumina GA II was performed at a maximum coverage of 51.7-fold (26), the coverage in the present study would be sufficient for the identification of SNVs and indels.

Detection of specific SNVs associated with quinolone resistance. At the first screening step, with reads mapping to Ames 0581, 93 possible SNVs were found in BA103 and 85 in BA103-CIPr (see Fig. S3A in the supplemental material). The Venn diagram in Fig. S3A indicates that 16, 6, and 75 positions could be SNV candidates for BA103-specific, BA103-CIPr-specific, and common SNVs, respectively. Incidentally, these SNV candidates appeared to be wrongly identified due to a misalignment with Maq, because 5 of the 6 BA103-CIPr-specific SNV candidates were found to be incorrect by Sanger sequencing methods. In most cases of such incorrect extractions, mapped reads showed a lower coverage depth at a 50% coefficient of variation (CV) at SNV candidates than those around the SNV position. Thus, the original filtering method was performed by calculating the CV with a coverage depth value around 100 bp at each SNV candidate. This filtering enabled us to refine the potential SNVs and led to the identification of one BA103-CIPr-specific SNV (Fig. S3A).

This BA103-CIPr-specific SNV was at genomic position 842518 and was a C-to-T base substitution in the GBAA0834 gene, encoding the TetR family transcriptional regulator. This substitution resulted in an amino acid substitution of threonine to isoleucine at the 38th amino acid ($Thr_{38}Ile$) (Table 1; Fig. 1A). SNV was not detected in the pXO1 and pXO2 plasmids between BA103 derivatives.

Using the same filtering methods described above, we identified 3 BA104-CIPr-specific SNVs, which were verified by Sanger sequencing (see Fig. S3A in the supplemental material). Among these 3 BA104-CIPr-specific SNVs, an SNV at genomic position 6850 was a G-to-C nucleotide substitution resulting in an amino acid substitution ($Ala_{86}Pro$) in the QRDR of the *gyrA* gene (GBAA0006) (Table 1). There are no previous reports of this SNV associated with CIP resistance in *B. anthracis*; however, the substitution has been found as $Ala_{84}Pro$ in *Escherichia coli* (25) and $Ser_{85}Pro$ in *Staphylococcus aureus* (10). The second SNV was located at genomic position 2138798 and was an A-to-G nucleotide substitution resulting in an amino acid substitution ($Lys_{80}Glu$) in GBAA2291 encoding the major sporulation sensor histidine kinase. The third SNV was located at genomic position 4595343 and was a substitution of T to C in the intergenic region between GBAA5072 and GBAA5074. SNV was not detected in the pXO1 and pXO2 plasmids between BA104 derivatives.

Detection of specific indels associated with quinolone resistance. At the first screening step, with reads mapping to *B. anthracis* Ames 0581, 7 possible short indels were identified in BA103 and 10 in BA103-CIPr (see Fig. S3B in the supplemental material). The Venn diagram in Fig. S3B shows that 4, 7, and 3 positions could be candidates for BA103-specific indels, BA103-CIPr-specific indels, and common indels, respectively. Further verification with a BlastN search around the positions of these indels refined 2 potential BA103-CIPr-specific indels (Table 2). These were validated by Sanger sequencing (see Fig. S6 in the supplemental material). The first short indel, located at genomic positions 842291 to 842295, was a deletion of 5 nucleotides (TAACA) in BA103-CIPr. This indel was located in the 132-bp intergenic region between GBAA0833 and GBAA0834 (Fig. 1A; see also Fig. S4A in the supplemental material). The second short indel was located at genomic positions 5215555 to 5215556 and was an insertion of 2 nucleotides (CA) within GBAA5724. This indel caused a frameshift in the coding sequence, resulting in the truncation of the amino acid sequence from amino acids (aa) 366 to 197 (see Fig. S5 in the supplemental material). Indel was not detected in the pXO1 and pXO2 plasmids between BA103 derivatives.

Using the filtering methods described above, we identified 9 BA104-CIPr-specific indels (Table 2). The first short indel was located at genomic positions 336869 to 336873 and was a deletion of 5 nucleotides (ACTTA) in GBAA0328, encoding a small multidrug resistance (SMR) family multidrug efflux

TABLE 2. Indels specific to the derivative CIP resistant mutants of BA103 and BA104

Derivative[a]	Genomic position[b]		Detected indels[c]			Genetic information on the indels		
	Insertion	Deletion	Ames 0581	Parental strain	Mutant strain	Locus tag[d]	Product	Strand[e]
BA103		842291–842295	TAACA	TAACA		Intergenic region (GBAA0833–0834)		
	5215555–5215556		CA	CA	CACA	GBAA5724 (*yhcF*)	Translation-associated GTPase	−
BA104		336869–336873	ACTTA	ACTTA		GBAA0328	SMR family multidrug efflux pump	+
		842291	T	T		Intergenic region (GBAA0833–0834)		
	842715		C	C	CC	GBAA0834	TetR family transcriptional regulator	+
	1059593–1059594		GT	GT	GTGT	GBAA1077	Hypothetical protein	−
	2609500–2609503		CGGC	CGGC	CGGCCGGC	GBAA2814	UvrD/Rep helicase family protein	−
	3363521–3363526		GTCACC	GTCACC	GTCACCGTCACC	GBAA3656 (*parC*)	DNA topoisomerase IV subunit A	−
	3796958–3796959		CT	CT	CTCT	GBAA4143 (*ylbF*)	ComK regulator	−
		4286011–4286021	TACCTTACAGT	TACCTTACAGT		Intergenic region (GBAA4707–4709)		
	4831099–4831104		GTCCCC	GTCCCC	GTCCCCGTCCCC	GBAA5329	Iron compound ABC transporter, permease protein	−

[a] MICs (mg/liter) of ciprofloxacin for parent strains and CIP-resistant mutants were 0.064 and 1.5, respectively, for BA103 and 0.064 and 6, respectively, for BA104.
[b] Numbers indicate the genomic position of the *B. anthracis* Ames 0581 strain (GenBank accession number, NC_007530).
[c] "Mutant strain" indicates CIP-resistant mutant generated from each "parental strain" by *in vitro* selection in this study.
[d] Locus tags are described as shown in the Comprehensive Microbial Resource (CMR) and Entrez Gene. As for the intergenic region, genes located adjacent to the SNV are shown.
[e] Clockwise and counterclockwise gene direction are indicated by "+" and "−" respectively.

pump. This indel caused a frameshift mutation leading to an increase in the coding sequence from aa 107 to aa 114 (see Fig. S5 in the supplemental material). The second short indel was located at genomic position 842291 and was a single nucleotide (T) deletion in the 132-bp intergenic region between GBAA0833 and GBAA0834 (Fig. 1A). Intriguingly, this deletion overlapped the first BA103-CIPr indel (Table 2; see also Fig. S4A in the supplemental material). Moreover, the third indel was located at genomic position 842715 and was a single nucleotide (C) insertion within the GBAA0834 TetR family regulator, the same as that mentioned above, leading to a truncation of the amino acid sequence from aa 191 to aa 117 (Fig. 1A; see also Fig. S4C in the supplemental material). The other 6 indels appeared to be generated by tandem duplications of 2 to 6 nucleotide bases, as summarized in Table 2. These indels resulted in a frameshift causing the truncation of the coding sequences of GBAA1077, GBAA2814, and GBAA4143 (see Fig. S5 in the supplemental material). On the other hand, the tandem duplications of 6 bases in GBAA3656 (*parC*) and GBAA5329 resulted in the incorporation of an additional 2 amino acids without frameshift mutation (Fig. S5).

Disruptive mutations of GBAA0834 are associated with overexpression of multidrug efflux pumps, leading to quinolone resistance. Three genes of the putative multidrug efflux system, GBAA0832, GBAA0833, and *blt* (GBAA0835), are located adjacent to GBAA0834 (Fig. 1A). We performed qRT-PCR to determine whether the TetR-type regulator (GBAA0834) is involved in the regulation of the multidrug efflux systems. All 4 genes were significantly upregulated in both BA103-CIPr and BA104-CIPr compared with their respective parental strains (Fig. 1B and C).

To evaluate whether multidrug efflux systems are involved in CIP resistance, the MIC was determined with reserpine, which acts by blocking efflux ability (14). Reserpine reduced the MIC by 4-fold in BA103-CIPr and 8-fold in BA104-CIPr (Fig. 1D), whereas it reduced the MIC by 1.3-fold in the parental strains.

DISCUSSION

Genetic variations in the QRDR of type II topoisomerases, the most common factors involved in CIP resistance, have been well characterized. It has been suggested, however, that a certain efflux system should contribute to the intermediate resistance to FQs. We performed a genomewide search using next-generation sequencing technology and identified novel genetic elements for resistance, in addition to the action of the QRDR.

BA103-CIPr and BA104-CIPr were classified as "nonsusceptible" isolates for CIP resistance according to the MIC interpretive standards (6); however, the CIP MICs for these strains are notable values for resistance in clinical treatment. A nonsynonymous mutation ($Ala_{86}Pro$) of the *gyrA* gene in BA104-CIPr appeared to contribute to resistance, because the substitution was also reported as $Ala_{84}Pro$ in *E. coli* (25) and $Ser_{85}Pro$ in *S. aureus* (10). No variation in the QRDR was found in BA103-CIPr, suggesting that other novel genetic variations were involved in its resistance.

Three genetic variations, including 1 SNV and 2 indels, were found in the intermediate-resistant strain BA103-CIPr; in fact, 2 of the 3 mutations were located around the TetR-type transcriptional regulator (GBAA0834). Multidrug efflux pumps are highly expressed mainly as a result of mutations of either their regulatory *cis* element or their transcriptional regulator genes (7). It appears that these 2 mutations could be associated with the CIP resistance of BA103-CIPr. A nonsynonymous mutation ($Thr_{38}Ile$) of GBA0834 was located in the conserved region of the "HTH_TETR_2" domain involved in DNA binding, in which the corresponding amino acid residue Thr_{40} of TetR in *E. coli* comes directly into contact with the *cis* element (17, 19). This threonine residue is well conserved in 34% of TetR family proteins, and thus, the substitution could cause a reduced DNA-binding affinity to a regulatory element.

The TetR family regulator is likely to modulate genes adjacent to the regulator gene (19). GBAA0832 and GBAA0833 are possibly involved with the small multidrug resistance protein (SMR family), and Blt (GBAA0835) shows a high amino acid sequence similarity with NorA of *S. aureus*, which is well characterized as a major efflux pump involved in FQ resistance (16). Indeed, our qRT-PCR data suggest the increased expression of Blt, GBAA0832, and GBAA0833 (Fig. 1B and C); in addition, the reserpine inhibition test indicates the contribution of the multidrug efflux system in CIP resistance (Fig. 1D). Thus, the functional disruption of GBAA0834 increased the transcription of these multidrug resistance genes, leading to a reduced susceptibility to CIP.

Twelve genetic variations, including 3 SNVs and 9 indels, were identified in the potentially resistant strain BA104-CIPr; 1 SNV was located in the *gyrA* gene and was discussed above. Two of the 9 indels were located around GBAA0834; one was an insertion of a C nucleotide in the "tetracycline repressor-like, C-terminal" domain involved in dimerization and ligand binding, and another was a deletion located upstream of GBAA0834. We propose that this insertion resulted in the disruption of GBAA0834, leading to the increased transcription of the 3 adjacent putative multidrug resistance genes, as was observed for BA103-CIPr.

Regarding the synonymous or nonsynonymous SNVs, all 3 SNVs in the coding sequence of the genes caused nonsynonymous mutations (Table 1), while surprisingly, synonymous mutations were not found throughout the whole genome. A transversional substitution, which is supposed to occur at a lower frequency than transitional substitutions, was found in the *gyrA* gene of BA104-CIPr, implying that the detection of such a rare mutation might be the result of stringent selection pressure, such as CIP selection for adaptive mutations (5, 24).

Regarding the detection of short indels, it is recommended to use paired-end sequencing reads rather than single-end reads. Although single-end sequencing reads were used in this study, 11 novel short indels could be identified with a combination of Maq mapping, BlastN search, and Sanger sequencing. Our strategy of using short single-end reads is applicable to the identification of SNVs and short indels as strain-specific genetic markers. Variable-number tandem repeats (VNTRs) were too long to determine indels with short-read mapping. Thus, 25 VNTR loci were analyzed by multiple-locus VNTR analysis (MLVA), and a single locus (pXO2) of BA104-CIPr was found to be extended by 2 repeat units (4 bp) compared with that of BA104 (see Table S4 in the supplemental material). FQs affect DNA integrity by blocking DNA topoisomerases, suggesting that they might cause an increase in indel

variation. Such unstable VNTR loci may be an unsuitable genetic marker for the traceability of *B. anthracis* strains.

Taken together, our findings with next-generation DNA sequencing technology enabled us to identify 15 novel genetic variations associated with CIP resistance. At this point, we can conclude that the GBAA0834 locus is a novel "mutation hot spot" related to resistance, in addition to the QRDR. Type II topoisomerase is essential; thus, genetic variation around GBAA0834 could be easily acquired as a primary step toward the development of resistance in *B. anthracis*. Moreover, such an intermediate resistance could increase the probability of the emergence of highly resistant isolates. Although the QRDR is the most reliable target for investigating FQ resistance, our findings will provide crucial genetic information for the detection of nonsusceptible isolates between isolates intermediate and fully resistant to CIP.

ACKNOWLEDGMENTS

We thank Tadahito Kanda and Akio Yamada for valuable suggestions.

This work was supported by a Research on Emerging and Re-emerging Infectious Diseases grant (H20 and 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. **Bast, D. J., A. Athamna, C. L. Duncan, J. C. de Azavedo, D. E. Low, G. Rahav, D. Farrell, and E. Rubinstein.** 2004. Type II topoisomerase mutations in *Bacillus anthracis* associated with high-level fluoroquinolone resistance. J. Antimicrob. Chemother. **54:**90–94.
3. **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.
4. **Bentley, D. R., S. Balasubramanian, H. P. Swerdlow, G. P. Smith, J. Milton, C. G. Brown, K. P. Hall, D. J. Evers, C. L. Barnes, H. R. Bignell, J. M. Boutell, J. Bryant, R. J. Carter, R. Keira Cheetham, A. J. Cox, D. J. Ellis, M. R. Flatbush, N. A. Gormley, S. J. Humphray, L. J. Irving, M. S. Karbelashvili, S. M. Kirk, H. Li, X. Liu, K. S. Maisinger, L. J. Murray, B. Obradovic, T. Ost, M. L. Parkinson, M. R. Pratt, I. M. Rasolonjatovo, M. T. Reed, R. Rigatti, C. Rodighiero, M. T. Ross, A. Sabot, S. V. Sankar, A. Scally, G. P. Schroth, M. E. Smith, V. P. Smith, A. Spiridou, P. E. Torrance, S. S. Tzonev, E. H. Vermaas, K. Walter, X. Wu, L. Zhang, M. D. Alam, C. Anastasi, I. C. Aniebo, D. M. Bailey, I. R. Bancarz, S. Banerjee, S. G. Barbour, P. A. Baybayan, V. A. Benoit, K. F. Benson, C. Bevis, P. J. Black, A. Boodhun, J. S. Brennan, J. A. Bridgham, R. C. Brown, A. A. Brown, D. H. Buermann, A. A. Bundu, J. C. Burrows, N. P. Carter, N. Castillo, E. C. M. Chiara, S. Chang, R. Neil Cooley, N. R. Crake, O. O. Dada, K. D. Diakoumakos, B. Dominguez-Fernandez, D. J. Earnshaw, U. C. Egbujor, D. W. Elmore, S. S. Etchin, M. R. Ewan, M. Fedurco, L. J. Fraser, K. V. Fuentes Fajardo, W. Scott Furey, D. George, K. J. Gietzen, C. P. Goddard, G. S. Golda, P. A. Granieri, D. E. Green, D. L. Gustafson, N. F. Hansen, K. Harnish, C. D. Haudenschild, N. I. Heyer, M. M. Hims, J. T. Ho, A. M. Horgan, et al.** 2008. Accurate whole human genome sequencing using reversible terminator chemistry. Nature **456:**53–59.
5. **Chattopadhyay, S., S. J. Weissman, V. N. Minin, T. A. Russo, D. E. Dykhuizen, and E. V. Sokurenko.** 2009. High frequency of hotspot mutations in core genes of *Escherichia coli* due to short-term positive selection. Proc. Natl. Acad. Sci. U. S. A. **106:**12412–12417.
6. **CLSI.** 2008. Performance standards for antimicrobial susceptibility testing; 18th informational supplement. M100-S18. Clinical and Laboratory Standards Institute, Wayne, PA.
7. **Grkovic, S., M. H. Brown, and R. A. Skurray.** 2002. Regulation of bacterial drug export systems. Microbiol. Mol. Biol. Rev. **66:**671–701.
8. **Grohs, P., I. Podglajen, and L. Gutmann.** 2004. Activities of different fluoroquinolones against *Bacillus anthracis* mutants selected in vitro and harboring topoisomerase mutations. Antimicrob. Agents Chemother. **48:**3024–3027.
9. **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.
10. **Ito, H., H. Yoshida, M. Bogaki-Shonai, T. Niga, H. Hattori, and S. Nakamura.** 1994. Quinolone resistance mutations in the DNA gyrase *gyrA* and *gyrB* genes of *Staphylococcus aureus*. Antimicrob. Agents Chemother. **38:** 2014–2023.
11. **Ley, T. J., E. R. Mardis, L. Ding, B. Fulton, M. D. McLellan, K. Chen, D. Dooling, B. H. Dunford-Shore, S. McGrath, M. Hickenbotham, L. Cook, R. Abbott, D. E. Larson, D. C. Koboldt, C. Pohl, S. Smith, A. Hawkins, S. Abbott, D. Locke, L. W. Hillier, T. Miner, L. Fulton, V. Magrini, T. Wylie, J. Glasscock, J. Conyers, N. Sander, X. Shi, J. R. Osborne, P. Minx, D. Gordon, A. Chinwalla, Y. Zhao, R. E. Ries, J. E. Payton, P. Westervelt, M. H. Tomasson, M. Watson, J. Baty, J. Ivanovich, S. Heath, W. D. Shannon, R. Nagarajan, M. J. Walter, D. C. Link, T. A. Graubert, J. F. DiPersio, and R. K. Wilson.** 2008. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature **456:**66–72.
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. **Markham, P. N., and A. A. Neyfakh.** 2001. Efflux-mediated drug resistance in Gram-positive bacteria. Curr. Opin. Microbiol. **4:**509–514.
14. **Markham, P. N., E. Westhaus, K. Klyachko, M. E. Johnson, and A. A. Neyfakh.** 1999. Multiple novel inhibitors of the NorA multidrug transporter of *Staphylococcus aureus*. Antimicrob. Agents Chemother. **43:**2404–2408.
15. **Mock, M., and A. Fouet.** 2001. Anthrax. Annu. Rev. Microbiol. **55:**647–671.
16. **Ng, E. Y., M. Trucksis, and D. C. Hooper.** 1994. Quinolone resistance mediated by *norA*: physiologic characterization and relationship to *flqB*, a quinolone resistance locus on the *Staphylococcus aureus* chromosome. Antimicrob. Agents Chemother. **38:**1345–1355.
17. **Orth, P., D. Schnappinger, W. Hillen, W. Saenger, and W. Hinrichs.** 2000. Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. Nat. Struct. Biol. **7:**215–219.
18. **Price, L. B., A. Vogler, T. Pearson, J. D. Busch, J. M. Schupp, and P. Keim.** 2003. In vitro selection and characterization of *Bacillus anthracis* mutants with high-level resistance to ciprofloxacin. Antimicrob. Agents Chemother. **47:**2362–2365.
19. **Ramos, J. L., M. Martinez-Bueno, A. J. Molina-Henares, W. Teran, K. Watanabe, X. Zhang, M. T. Gallegos, R. Brennan, and R. Tobes.** 2005. The TetR family of transcriptional repressors. Microbiol. Mol. Biol. Rev. **69:**326–356.
20. **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.
21. **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.
22. **Ruiz, J.** 2003. Mechanisms of resistance to quinolones: target alterations, decreased accumulation and DNA gyrase protection. J. Antimicrob. Chemother. **51:**1109–1117.
23. **Schmitz, F. J., P. G. Higgins, S. Mayer, A. C. Fluit, and A. Dalhoff.** 2002. Activity of quinolones against gram-positive cocci: mechanisms of drug action and bacterial resistance. Eur. J. Clin. Microbiol. Infect. Dis. **21:**647–659.
24. **Sokurenko, E. V., D. L. Hasty, and D. E. Dykhuizen.** 1999. Pathoadaptive mutations: gene loss and variation in bacterial pathogens. Trends Microbiol. **7:**191–195.
25. **Sreedharan, S., M. Oram, B. Jensen, L. R. Peterson, and L. M. Fisher.** 1990. DNA gyrase *gyrA* mutations in ciprofloxacin-resistant strains of *Staphylococcus aureus*: close similarity with quinolone resistance mutations in *Escherichia coli*. J. Bacteriol. **172:**7260–7262.
26. **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.
27. **Swartz, M. N.** 2001. Recognition and management of anthrax—an update. N. Engl. J. Med. **345:**1621–1626.
28. **Wang, J., W. Wang, R. Li, Y. Li, G. Tian, L. Goodman, W. Fan, J. Zhang, J. Li, Y. Guo, B. Feng, H. Li, Y. Lu, X. Fang, H. Liang, Z. Du, D. Li, Y. Zhao, Y. Hu, Z. Yang, H. Zheng, I. Hellmann, M. Inouye, J. Pool, X. Yi, J. Zhao, J. Duan, Y. Zhou, J. Qin, L. Ma, G. Li, G. Zhang, B. Yang, C. Yu, F. Liang, W. Li, S. Li, P. Ni, J. Ruan, Q. Li, H. Zhu, D. Liu, Z. Lu, N. Li, G. Guo, J. Ye, L. Fang, Q. Hao, Q. Chen, Y. Liang, Y. Su, A. San, C. Ping, S. Yang, F. Chen, L. Li, K. Zhou, Y. Ren, L. Yang, Y. Gao, G. Yang, Z. Li, X. Feng, K. Kristiansen, G. K. Wong, R. Nielsen, R. Durbin, L. Bolund, X. Zhang, and H. Yang.** 2008. The diploid genome sequence of an Asian individual. Nature **456:**60–65.