APPLIED AND ENVIRONMENTAL MICROBIOLOGY, May 2003, p. 2755–2764 Vol. 69, No. 5
0099-2240/03/$08.00+0 DOI: 10.1128/AEM.69.5.2755–2764.2003

Genome Differences That Distinguish *Bacillus anthracis* from *Bacillus cereus* and *Bacillus thuringiensis*

Lyndsay Radnedge,[1] Peter G. Agron,[1] Karen K. Hill,[2] Paul J. Jackson,[2] Lawrence O. Ticknor,[2] Paul Keim,[3] and Gary L. Andersen[4]*

Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, California 94551[1]; Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545[2]; Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011[3]; and Center for Environmental Biotechnology, Lawrence Berkeley National Laboratory, Berkeley, California 94720[4]

Received 15 October 2002/Accepted 13 February 2003

The three species of the group 1 bacilli, *Bacillus anthracis*, *B. cereus*, and *B. thuringiensis*, are genetically very closely related. All inhabit soil habitats but exhibit different phenotypes. *B. anthracis* is the causative agent of anthrax and is phylogenetically monomorphic, while *B. cereus* and *B. thuringiensis* are genetically more diverse. An amplified fragment length polymorphism analysis described here demonstrates genetic diversity among a collection of non-anthrax-causing *Bacillus* species, some of which show significant similarity to *B. anthracis*. Suppression subtractive hybridization was then used to characterize the genomic differences that distinguish three of the non-anthrax-causing bacilli from *B. anthracis* Ames. Ninety-three DNA sequences that were present in *B. anthracis* but absent from the non-anthrax-causing *Bacillus* genomes were isolated. Furthermore, 28 of these sequences were not found in a collection of 10 non-anthrax-causing *Bacillus* species but were present in all members of a representative collection of *B. anthracis* strains. These sequences map to distinct loci on the *B. anthracis* genome and can be assayed simultaneously in multiplex PCR assays for rapid and highly specific DNA-based detection of *B. anthracis*.

Bacillus anthracis, *B. cereus*, and *B. thuringiensis* are genetically very closely related members of the group 1 bacilli, a fact which has led to the proposal that they should be considered a single species (15). *B. cereus* is frequently isolated as a contaminant of various foods and can occasionally be an opportunistic human pathogen (9, 15). *B. thuringiensis* has been widely exploited in agriculture as an insecticide by virtue of the presence of plasmid-borne crystal toxin genes (34). *B. anthracis* is a virulent pathogen of mammals and is the causative agent of anthrax. All three species are readily isolated from soil environments (38).

Extensive genetic diversity among environmental isolates of *B. cereus* and *B. thuringiensis* has been demonstrated by pulsed-field gel electrophoresis (6), multienzyme electrophoresis (6, 15), and amplified fragment length polymorphism (AFLP) analysis (38). AFLP analysis proved sensitive enough to classify *B. cereus* and *B. thuringiensis* into five phylogenetic groups (38). Significantly, the American Type Culture Collection (ATCC) reference strains of *B. cereus* and *B. thuringiensis* did not seem to be represented in the collection of environmental isolates used in that study. There was little correlation between species designation and the five phylogenetic groups identified by AFLP analysis, indicating significant genetic variability within and between *B. cereus* and *B. thuringiensis*. Furthermore, in two separate studies, the *B. anthracis* strains appeared to cluster together with a group containing periodontal *B. cereus* pathogens (16, 38).

* Corresponding author. Mailing address: Center for Environmental Biotechnology, Lawrence Berkeley National Laboratory, 1 Cyclotron Rd., Mail Stop 70A-3317, Berkeley, CA 94720. Phone: (510) 495-2795. Fax: (510) 486-7152. E-mail: GLAndersen@lbl.gov.

In contrast to the genomic diversity within *B. cereus* and *B. thuringiensis*, *B. anthracis* appears to be genetically clonal (19). Genetic identity among different *B. anthracis* isolates has necessitated fastidious analysis of their genomes for strain identification. Accurate discrimination is now possible by analysis of variable-number tandem repeats, which enumerates small tandem repeats at several locations in the *B. anthracis* genome (3, 20). A collection of 426 *B. anthracis* environmental isolates that contains representatives from worldwide origins was subdivided by variable-number tandem repeat analysis into just six genetically distinct groups (20).

Several strains of *B. cereus* and *B. thuringiensis* which appear to be genetically related to *B. anthracis* have been isolated (15, 16, 38). Isolation and characterization of the genome regions unique to *B. anthracis* will provide clues to its genetic relationship to these strains and ultimately direct pathogenicity studies. AFLP analysis provides a rapid method for measuring phylogenetic distances. AFLP analysis generates a strain-specific fingerprint of amplified DNA fragments that demonstrate genomic variations in a microbial population based on an analysis of a portion of their genome sequences and DNA fragment length polymorphisms. The great advantage of AFLP analysis is its ability to analyze rapidly many loci, resulting in a phylogenetic resolution higher than those obtained with other methods. However, AFLP analysis provides little information about the genetic differences responsible for these polymorphisms. Suppression subtractive hybridization (SSH) is a highly efficient technique for the isolation and characterization of the large genomic differences that often drive bacterial genome evolution (1, 2). SSH reveals DNA sequence differences that are responsible for many AFLPs.

Unique genomic differences can be exploited as "DNA sig-

natures" for the discrimination of *B. anthracis* from its closest relatives. Plasmid-encoded toxin genes have proved a useful source of targets for rapid DNA-based assays (4). However, a plasmid-based test may not detect the occurrence of non-plasmid-containing strains of *B. anthracis*, which have been isolated from the environment (39). Furthermore, plasmids can be readily engineered and can be transferred to other bacteria for heterologous gene expression (5, 10, 37), and there is concern that pXO1 and pXO2 sequences that are present in other *Bacillus* strains may be encountered in previously uncharacterized genomes (25). There are many published examples of chromosomal regions that can be examined for *B. anthracis* identification, but all require time-consuming downstream analysis (3, 7, 18, 26). Two recent examples of real-time PCR targeting of the *rpoB* gene of *B. anthracis* were useful but reported examples of false-positive results with some strains of *B. cereus* and targeted only a single locus (11, 28).

This report demonstrates the phylogenetic relationships of a collection of non-anthrax-causing *Bacillus* species to each other and to *B. anthracis*, as determined by AFLP analysis. SSH with *B. anthracis* and three of these close relatives identified a set of unique DNA regions that represent genomic differences between *B. anthracis* and these non-anthrax-causing *Bacillus* species. One immediate product of these experiments was a robust set of chromosomal DNA signatures which are capable of quickly detecting all six genetically distinct groups of *B. anthracis* (20) and which can be used with any rapid DNA-based detection platform. A multiplex PCR analysis with four separate loci of the *B. anthracis* chromosome provides a rapid and highly specific means for the identification of *B. anthracis*.

MATERIALS AND METHODS

AFLP analysis of DNA samples. AFLP analysis was accomplished as previously described (19, 40). Briefly, DNA (100 ng) was digested with *Eco*RI and *Mse*I, and the resulting fragments were ligated to double-stranded adaptors. The digested and ligated DNA was then amplified by PCR with *Eco*RI and *Mse*I +0/+0 primers. The +0/+0 PCR product was analyzed by agarose gel electrophoresis to determine the size range of amplified fragments. Three microliters was used in subsequent selective amplifications with 6-carboxyfluorescein-labeled +1/+1 primers *Eco*RI-C (5′-GTAGACTGCGTACCAATTCC-3′) and *Mse*I-G (5′-GACGATGAGTCCTGAGTAAG-3′). Selective amplifications were performed with 20-μl reaction mixtures. The resulting products (0.5 to 1.0 μl) were mixed with a solution containing the DNA size standards Genescan-500 (Applied Biosystems, Inc., Foster City, Calif.) and MapMarker-400 (BioVentures, Inc., Murfreesburo, Tenn.), both labeled with *N*,*N*,*N*,*N*-tetramethyl-6-carboxyrhodamine. Following heat denaturation at 90°C for 2 min, the reaction mixtures were loaded into a 5% Long Ranger DNA sequencing gel (BioWhittaker Molecular Applications, Rockland, Maine) and visualized with an ABI 377 automated fluorescence sequencer (Applied Biosystems). GeneScan analysis software (Applied Biosystems) was used to determine the lengths of the sample fragments by comparison to the DNA fragment length size standards included with each sample. AFLP data analysis was performed as described by Ticknor et al. (38).

Preparation of genomic DNA. The bacterial strains used for genomic DNA preparation and their geographic origins, where known, are listed in Table 1. DNA was extracted from cell pellets from overnight cultures grown in L broth by using a MasterPure DNA purification kit (Epicentre, Madison, Wis.) according to the manufacturer's instructions. Genomic DNA was prepared from enteric bacteria by using Wizard Genomic DNA Preps (Promega, Madison, Wis.) according to the manufacturer's instructions. Human genomic DNA was purchased from Clontech (Palo Alto, Calif.), and bovine genomic DNA was purchased from Novagen (Madison, Wis.). Soil DNA was extracted by using an UltraClean soil DNA kit (MoBio, Solana Beach, Calif.) according to the manufacturer's instructions. DNA isolated from organisms present in the air was prepared from filters as previously described (31). Both soil and air filter DNAs were prepared from samples originating in Livermore, Calif. After ethanol precipitation, all genomic DNAs were dissolved in 10 mM Tris HCl (pH 8.0) to a concentration of approximately 0.2 μg/ml.

TABLE 1. Strains used in this study

Strain	Phylogenetic group[a]	Geographic origin(s)[b]
B. anthracis		
G3	A1.a	North America, Europe, Brazil
G20	A1.a	Italy
G25	A1.b	U.S.
G29	A2	Pakistan
G38	A3.a	Germany
G62 (Ames)	A3.b	U.S., U.K.
G67	A3	South Africa
G77 (Vollum)	A4	U.K., Spain, Zimbabwe
G80	B2	France
G87	B1	South Africa
Non-anthrax-causing *Bacillus* species		
B. cereus ATCC 14579	NA	NR
B. cereus ATCC 4342	NA	NR
B. cereus D17	NA	U.S.
B. cereus 3A	NA	U.S.
B. cereus S2-8	NA	U.S.
B. cereus F1-15	NA	U.S.
B. thuringiensis 97-27	NA	Bosnia
B. thuringiensis Al Hakam	NA	Iraq
B. thuringiensis ATCC 10792	NA	Israel
B. thuringiensis HD-571	NA	USDA

[a] Phylogenetic group designations were described by Keim et al. (20). NA, not applicable.
[b] U.S., United States; U.K., United Kingdom; USDA, U.S. Department of Agriculture; NR, the geographic origin of this strain was not recorded by ATCC.

SSH. Genome comparisons by SSH were performed as previously described, except for the differences noted below (2, 30). Briefly, the tester-specific DNA was digested with restriction endonucleases (*Rsa*I and *Mse*I; New England Biolabs, Beverly, Mass.) used according to the manufacturer's instructions. The fragments were first marked by ligation to specialized oligonucleotide adaptors. When the marked DNA was denatured and hybridized to excess unmarked driver-specific DNA that had been digested with the same enzymes, most tester-specific sequences formed heterohybrids with the driver. Some tester-specific sequences, however, self-hybridized to form amplifiable fragments that were then enriched by PCR, cloned, and sequenced. Agarose gel electrophoresis determined that the restriction endonucleases *Mse*I and *Rsa*I cut the genomic DNA to generate fragments in the optimal size range (200 to 1,000 bp). Modified adaptors were constructed to allow for subtractions with *Mse*I-digested DNA as previously described (1). Namely, adaptor 1 was formed by annealing the "adaptor 1 long" oligonucleotide (2) with the oligonucleotide 5′-TAACCTGCCCGG to form an adaptor with appropriate cohesive ends. Adaptor 2 was formed by annealing the "adaptor 2 long" oligonucleotide with the oligonucleotide 5′-TA ACCTCGGCCG. T4 DNA ligase (New England Biolabs) was inactivated by incubation at 72°C for 20 min. Three separate SSH experiments were performed with *B. anthracis* Ames as the tester and three of the non-anthrax-causing *Bacillus* strains listed in Table 1 as the drivers: *B. cereus* ATCC 14579, *B. cereus* 3A, and *B. thuringiensis* Al Hakam.

DNA sequencing. Nonpurified PCR products were cloned by using a pGEM-T Easy TA cloning kit (Promega). Recombinant clones were picked by using a BioPick automated colony picker (BioRobotics, Woburn, Mass.), and plasmid templates were prepared by boiling lysis and magnetic bead capture with a high-throughput procedure (35). Sequencing of plasmid templates was performed by using an Applied Biosystems Big-Dye Terminator system and either an ABI 377 or an ABI 3700 automated sequencer. The sequencing primers used were 5′-TGTAAAACGACGGCCAGT (forward) and 5′-CAGGAAACAGCT ATGACC (reverse). The resulting data were analyzed by using ABI sequencing analysis software, version 3.2, and then assembled and edited by using Phred, Phrap, and Consed 7.0 (14). BLAST searches with the tester-specific DNA sequences were performed by using the National Center for Biotechnology

Information website http://www.ncbi.nlm.nih.gov/. BLAST identity was considered significant only when the expected probability of a fortuitous match (E value) was less than 10^{-4}. Comparison of the sequence candidates to plasmids pXO1 and pXO2 and determination of their coordinates on the published *B. anthracis* genome (32) were performed by using the cross_match program, which is part of the Consed software package (14).

PCRs. Oligonucleotide primers were designed from the putative tester-specific sequences and were supplied by Sigma-Genosys (The Woodlands, Tex.) or Invitrogen (Carlsbad, Calif.). The primers were designed by using Primer3 software (33); they had a melting temperature of >65°C, contained no more than three identical consecutive nucleotides, and possessed a two-nucleotide 3′ GC clamp. The primers were initially screened against genomic DNAs from both the tester and the driver. To determine whether a primer pair was tester specific, 1 ng each of tester- and driver-specific DNAs was used as a template in PCRs with Accuprime polymerase, primers at 10 μM (Invitrogen), and the following cycling parameters: 94°C for 15 s, 65°C for 15 s, and 72°C for 30 s for 32 cycles. The products were visualized on a 1.5% agarose gel run in 0.5× Tris-borate-EDTA; if a product was present with tester-specific DNA as a template and absent with driver-specific DNA, then that sequence was designated tester specific. The tester-specific oligonucleotides were then used to prime PCRs with the *B. anthracis*, *B. cereus*, and *B. thuringiensis* strains listed in Table 1 and the same reaction conditions as those described above. The integrity of the genomic DNA template was tested in all PCRs with primers specific for a region of the 23S gene conserved in *Bacillus* species: 5′-CTACCTTAGGACCGTTATAGTTAC and 5′-AGGTAGGCGAGGAGAGAATCC. Multiplex PCRs were performed as described above, except that template DNA was added to a final concentration of 10 ng/μl. The primers (see Table 5) were used at the following concentrations: 23S primers at 3 μM (288 bp), dhp73.017 (241 bp) at 20 μM, dhp73.019 (196 bp) at 10 μM, dhp61.183 (163 bp) at 5 μM, and dhp77.002 (133 bp) at 20 μM.

RESULTS

Genetic relatedness of *B. anthracis* to *B. cereus* and *B. thuringiensis*, as determined by AFLP. The phylogenetic tree derived from the fluorescent AFLP data is shown in Fig. 1. The *Bacillus* isolates used in this study were chosen from a previous AFLP study of over 300 *Bacillus* isolates (K. K. Hill and P. J. Jackson, unpublished data). They included several *B. cereus* and *B. thuringiensis* isolates that were found to be the most closely related to *B. anthracis* and two more distantly related *B. cereus* and *B. thuringiensis* strains (ATCC type strains 14579 and 10792, respectively). The other *Bacillus* isolates represented on the tree include pathogenic and nonpathogenic isolates; details for these isolates are shown in Table 1. *B. anthracis* is represented in this tree by the Ames and Vollum strains. A large number of *B. anthracis* isolates were previously tested and found to possess AFLP profiles indistinguishable from those of these two strains (data not shown). The isolate most closely related to *B. anthracis* is *B. thuringiensis* 97-27. This isolate was collected from the wound of a French soldier and was shown to be capable of infecting and killing immunocompetent mice in subsequent studies (17). The three other *B. thuringiensis* isolates on the tree are HD-571, obtained from the U.S. Department of Agriculture *B. thuringiensis* collection; Al Hakam, collected by the United Nations Special Commission at a suspected bioweapons facility in Iraq; and the ATCC type strain (ATCC 10792). The *B. cereus* isolates also include pathogenic and nonpathogenic members of this species. Three *B. cereus* isolates (F1-15, 3A, and D17) were collected from food sources that caused human illness, and one (S2-8) was a soil isolate. The *B. cereus* ATCC type strain (ATCC 14579) and ATCC 4342, an isolate from milk (36), were also included.

Data were analyzed by using three or more replicates of the AFLP profiles to compare each isolate to the other isolates. AFLP fragments that appeared in every replicate for an isolate were used to define a fingerprint for that isolate. Each gel contained one or more control DNA samples to allow comparisons of profiles on the different gels. Jaccard coefficients for the fingerprints of the isolates being compared were analyzed by using an unweighted pair-group method with arithmetic means to produce the dendrogram. Based on an analysis of control DNA samples, the variability within the analysis was less than about 0.2 for the Jaccard coefficients. Any differences below this level therefore may be due to experimental or analytical variability and cannot be considered significant. Differences above this level are reproducible and carry a high degree of confidence (38).

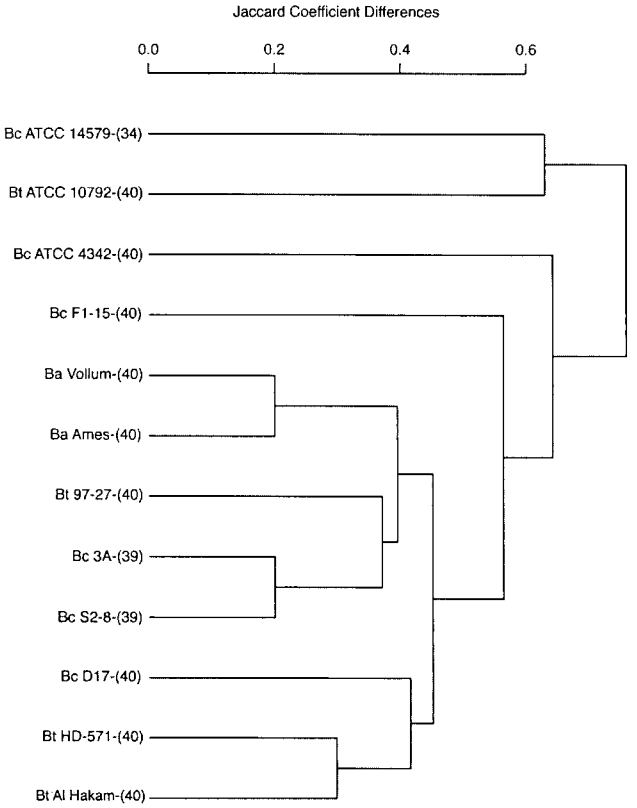

FIG. 1. Phylogenetic tree of *B. anthracis*, *B. cereus*, and *B. thuringiensis* isolates. The phylogenetic tree was derived from fluorescent AFLP analyses of 12 *B. anthracis* (Ba), *B. cereus* (Bc), and *B. thuringiensis* (Bt) isolates. The phylogenetic tree was based on 34 to 40 amplified DNA fragments per sample (the number per sample is shown in parentheses) generated from *Eco*RI/*Mse*I digestion of genomic DNAs. Jaccard coefficients for the fragments common in three replicate gels for each isolate were analyzed by using the unweighted pair-group method with arithmetic means to produce the dendrogram.

Isolation of tester-specific DNA sequences by SSH. SSH identifies DNA sequences that are specific to one genome (tester) and absent from the other genome (driver). Typically these DNA sequences range in length from 200 to 1,000 bp and are referred to as difference products. Three subtractions were undertaken to increase the yield and representation of *B. anthracis*-specific regions. *B. cereus* ATCC 14579 was selected because it represents one of the two most distantly related strains in this study and would be most likely to yield difference products. *B. cereus* 3A was chosen because it represents one of

TABLE 2. Isolation of 28 *B. anthracis*-specific DNA signatures from three SSH experiments

Tester	Driver	Candidate sequence	No. (%) of candidates remaining
B. anthracis Ames	*B. cereus* ATCC 14579	From SSH	256 (100)
		Tester specific	39 (15)
		After screening with non-anthrax-causing *Bacillus* pathogens	6 (2)
B. anthracis Ames	*B. thuringiensis* Al Hakam	From SSH	48 (100)
		Tester specific	28 (58)
		After screening with non-anthrax-causing *Bacillus* pathogens	8 (17)
B. anthracis Ames	*B. cereus* 3A	From SSH	48 (100)
		Tester specific	26 (54)
		After screening with non-anthrax-causing *Bacillus* pathogens	14 (29)

the closest relatives of *B. anthracis*, while *B. thuringiensis* Al Hakam represents intermediate relatedness.

A total of 256 candidate sequences were generated from subtractions by using genomic DNA from *B. cereus* ATCC 14579 as the driver (Table 2). Sequences that do not occur on plasmids pXO1 and pXO2 of *B. anthracis* were of primary interest, since some plasmid sequences are conserved in closely related *Bacillus* species (25). Sequences identical to pXO1 and pXO2 were eliminated by computer comparisons to published nucleotide sequences (see Materials and Methods) (24). The remaining sequences were used to design oligonucleotide primers for PCR analysis, to determine their representation in the tester and driver genomes. PCR experiments with *B. anthracis* Ames and *B. cereus* ATCC 14579 DNA as a template identified 39 *B. anthracis* Ames-specific sequences (15% of the tester-specific candidates). Six of the tester-specific candidates (2%) did not amplify a product when DNA from the collection of non-anthrax-causing pathogens was used as a template. Similarly, 28 out of 48 sequences (58%) were absent from *B. thuringiensis* Al Hakam; 8 of these (16%) did not occur in the collection of closely related non-anthrax-causing pathogens (Table 2). Finally, 48 sequences each were generated from two separate subtractions by using genomic DNAs prepared from *B. cereus* 3A and *B. thuringiensis* Al Hakam as drivers. A total of 26 out of 48 sequences (54%) were present in *B. anthracis* Ames but absent from *B. cereus* 3A; 14 of these did not occur in the collection of closely related non-anthrax-causing pathogens (Table 2).

It is crucial that DNA signatures intended for DNA-based detection of *B. anthracis* show no false-negative results with all isolates that might be encountered in environmental samples. All of the *B. anthracis*-specific candidates successfully amplified a PCR product when DNAs from the eight representatives of *B. anthracis* (Table 1) were used as templates (data not shown) (20).

Distribution of tester-specific loci in *B. cereus* and *B. thuringiensis*. Genomic variations within the collection of non-anthrax-causing pathogens were observed as the presence or absence of 39 tester-specific regions isolated from the subtraction in which genomic DNA from *B. cereus* ATCC 14579 was used as the driver and DNA from *B. anthracis* Ames was used as the tester (Table 3). The tester-specific primers are listed in Table 3 in decreasing order of the number of products amplified from this panel of non-anthrax-causing pathogens. Since, by definition, the driver will always be negative, there will be a maximum of nine loci possible in the 10 non-anthrax-causing *Bacillus* species. Similarly, the tester will always be positive, so that the minimum number of products possible will be one. The data in Table 3 show that seven primer sets are present in nine templates (excluding the driver), while six candidates (shown in bold type) are represented in *B. anthracis* only. The columns of data for non-anthrax-causing pathogens are arranged from left to right in Table 3 according to the number of tester-specific sequences seen in their respective genomes. The maximum number of candidates possible for each non-anthrax-causing *Bacillus* species is 39 for the tester (*B. anthracis*), while the minimum is 0 for the driver (*B. cereus* ATCC 14579). Strains that have more markers in common are suggested to be more closely related to *B. anthracis*. The data in Table 3 show that *B. cereus* 3A is the most closely related (31 out of 39 sequences), followed by *B. cereus* S2-8 and D17 (29 out of 39 sequences); as expected, *B. thuringiensis* ATCC 10792 is the least closely related (10 out of 39 sequences).The presence or absence of these loci is not correlated with the current species designations of *B. cereus* and *B. thuringiensis*. Table 4 shows data obtained with primers designed from the subtraction with the most closely related DNA as a driver. Far fewer primer candidates detect the non-anthrax-causing *Bacillus* species, with 14 (shown in bold type) being found only for *B. anthracis* Ames.

***B. anthracis*-specific DNA sequences and primers.** From the original collection of 352 fragments isolated from three subtractions, 28 were specific for *B. anthracis* and did not amplify a product with the near neighbors. All 93 tester-specific primers, some of which can be used to distinguish the *B. cereus* and *B. thuringiensis* strains used in this study, can be found at http://bbrp.llnl.gov/html/BAspc.html. All primers successfully amplified a PCR product of the predicted size, showing no evidence of variations at these loci in the *B. anthracis* strains (data not shown). These 28 *B. anthracis*-specific candidates were used to screen DNAs from common enteric pathogens—*Yersinia pestis*, *Y. pseudotuberculosis*, *Y. enterocolitica*, and *Escherichia coli*—species that may be encountered in environmental samples in suspected anthrax cases (soil, air, human, and bo-

TABLE 3. Presence or absence of *B. anthracis* Ames-specific nucleotide sequences in the genomes of 10 non-anthrax-causing *Bacillus* pathogens as determined by PCR with *B. cereus* ATCC 14579 as the driver

Primer[a]	Presence (+) or absence (−) of *B. anthracis* Ames-specific nucleotide sequences in:										
	B. anthracis Ames (tester)	*B. cereus* 3A	*B. cereus* S2-8	*B. cereus* D17	*B. thuringiensis* 97-27	*B. thuringiensis* HD-571	*B. thuringiensis* Al Hakam	*B. cereus* ATCC 4342	*B. cereus* F1-15	*B. thuringiensis* ATCC 10792	*B. cereus* ATCC 14579 (driver)
M.Ctg015	+	+	+	+	+	+	+	+	+	+	−
R.Ctg013	+	+	+	+	+	+	+	+	+	+	−
R.Ctg048	+	+	+	+	+	+	+	+	+	+	−
R.Ctg102	+	+	+	+	+	+	+	+	+	+	−
R.Ctg107	+	+	+	+	+	+	+	+	+	+	−
R.Ctg137	+	+	+	+	+	+	+	+	+	+	−
R.Ctg173	+	+	+	+	+	+	+	+	+	+	−
M.Ctg011	+	+	+	+	+	+	+	+	+	−	−
M.Ctg013	+	+	+	+	+	+	+	+	+	−	−
M.Ctg033	+	+	+	+	+	+	+	+	+	−	−
M.Ctg038	+	+	+	+	+	+	+	+	+	−	−
M.Ctg039	+	+	+	+	+	+	+	+	+	−	−
R.Ctg070	+	+	+	+	+	+	+	+	+	−	−
R.Ctg116	+	+	+	+	+	+	+	+	+	−	−
R.Ctg153	+	+	+	+	+	+	+	+	+	−	−
R.Ctg168	+	+	+	+	+	+	+	+	+	−	−
R.Ctg018	+	+	+	+	+	+	+	+	−	−	−
R.Ctg036	+	+	+	+	+	+	+	+	−	−	−
R.Ctg087	+	+	+	+	+	+	+	+	−	−	−
M.Ctg025	+	+	+	+	+	+	+	−	+	−	−
R.Ctg079	+	+	+	+	+	+	+	−	+	−	−
R.Ctg108	+	+	+	+	+	+	+	−	−	+	−
M.Ctg003	+	+	+	+	+	+	+	−	−	−	−
R.Ctg119	+	+	+	+	+	+	+	−	−	−	−
R.Ctg156	+	+	+	−	+	+	+	+	−	−	−
M.Ctg056	+	+	−	+	+	+	+	−	+	−	−
R.Ctg100	+	+	+	+	+	−	−	−	−	−	−
M.Ctg016	+	+	+	+	−	−	−	−	−	−	−
R.Ctg185	+	+	+	−	−	−	−	+	−	−	−
R.Ctg197	+	+	+	−	−	−	−	+	−	−	−
M.Ctg007	+	+	−	−	−	−	−	+	−	+	−
M.Ctg031	+	−	−	+	−	−	−	+	−	+	−
M.Ctg012	+	−	−	+	−	−	−	+	−	−	−
M.Ctg032	+	−	−	−	−	−	−	−	−	−	−
M.Ctg037	+	−	−	−	−	−	−	−	−	−	−
R.Ctg132	+	−	−	−	−	−	−	−	−	−	−
R.Ctg139	+	−	−	−	−	−	−	−	−	−	−
R.Ctg152	+	−	−	−	−	−	−	−	−	−	−
R.Ctg177	+	−	−	−	−	−	−	−	−	−	−

[a] Bold type indicates primers that are specific for *B. anthracis*.

vine samples). All 28 candidates failed to amplify a product from these bacterial DNAs (data not shown). The average G+C content of these sequences was 35% (28 to 44%), a value consistent with that of *B. anthracis*.

All but 1 of the 28 *B. anthracis*-specific DNA sequences can be mapped to the 5.23-Mb genome of *B. anthracis* A2012 (GenBank accession number NC_003995) (32). BLASTX analysis showed that 24 of the remaining 27 DNA sequences map to the open reading frames defined therein, while the remainder map to intergenic regions. The coordinates for nucleotide identity to *B. anthracis* A2012, open reading frame identity, and gene identity are listed in Table 5. BLASTX identities with E values of less than 10^{-3} are also listed. BLAST data for the candidates from the *B. cereus* ATCC 14579 subtraction show that three of the six DNA sequences have no previously ascribed function, based on similarity searches of the GenBank database. One of the remaining three shows identity with the S-layer protein of *B. anthracis*, and the other two show identity with a hypothetical phage protein from *Streptococcus pyogenes* and a galactosyltransferase-related protein from *Clostridium acetobutylicum*. BLAST data for the candidates from the *B. thuringiensis* Al Hakam subtraction show that four of the eight DNA sequences have no previously ascribed function. The remaining four share sequence identity with a hypothetical protein in *B. halodurans*, a penicillin binding protein of *B. cereus*, a putative phage terminase of *C. perfringens*, and a cytosine-specific methyltransferase of *B. halodurans*. BLAST data for the candidates from the *B. cereus* 3A subtraction show that 9 of the 14 DNA signatures have no previously ascribed function. The remaining five show sequence identity with the *Staphylococcus aureus* terminase large subunit, a hypothetical protein from the *nanH* region and an ATP binding cassette transporter of *C. perfringens*, a glucosamine synthetase of *B. subtilis*, and an unknown conserved protein of *B. halodurans*.

Figure 2 provides a visualization of the distribution of the tester-specific sequences for each of the three subtractive hybridization experiments. There are five regions that have more than two *B. anthracis*-specific loci that lie within 50 kb of each other, suggesting genomic islands that are found only in *B.*

TABLE 4. Presence or absence of *B. anthracis* Ames-specific nucleotide sequences in the genomes of 10 non-anthrax-causing *Bacillus* pathogens as determined by PCR with *B. cereus* 3A as the driver

Primer[a]	Presence (+) or absence (−) of *B. anthracis* Ames-specific nucleotide sequences in:										
	B. anthracis Ames (tester)	*B. cereus* D17	*B. thuringiensis* HD-571	*B. thuringiensis* Al Hakam	*B. cereus* F1-15	*B. cereus* ATCC 4342	*B. thuringiensis* 97-27	*B. thuringiensis* ATCC 10792	*B. cereus* ATCC 14579	*B. cereus* S2-8	*B. cereus* 3A (driver)
dhp77.46	+	+	+	+	−	+	−	+	−	−	−
dhp73.18	+	+	+	+	−	−	+	−	−	−	−
dhp73.03	+	+	+	−	+	+	−	−	−	−	−
dhp77.38	+	−	+	+	−	−	−	+	−	−	−
dhp73.04	+	−	−	+	−	−	+	−	−	−	−
dhp73.21	+	+	−	−	−	−	−	−	−	−	−
dhp77.50	+	+	−	−	−	−	−	−	−	−	−
dhp77.21	+	−	−	−	+	−	−	−	−	−	−
dhp77.27	+	−	−	−	+	−	−	−	−	−	−
dhp77.42	+	−	−	−	+	−	−	−	−	−	−
dhp77.36	+	−	−	−	−	+	−	−	−	−	−
dhp73.12	+	−	−	−	−	−	+	−	−	−	−
dhp73.01	+	−	−	−	−	−	−	−	−	−	−
dhp73.09	+	−	−	−	−	−	−	−	−	−	−
dhp73.11	+	−	−	−	−	−	−	−	−	−	−
dhp73.16	+	−	−	−	−	−	−	−	−	−	−
dhp73.17	+	−	−	−	−	−	−	−	−	−	−
dhp73.19	+	−	−	−	−	−	−	−	−	−	−
dhp73.22	+	−	−	−	−	−	−	−	−	−	−
dhp73.26	+	−	−	−	−	−	−	−	−	−	−
dhp73.29	+	−	−	−	−	−	−	−	−	−	−
dhp77.02	+	−	−	−	−	−	−	−	−	−	−
dhp77.03	+	−	−	−	−	−	−	−	−	−	−
dhp77.04	+	−	−	−	−	−	−	−	−	−	−
dhp77.47	+	−	−	−	−	−	−	−	−	−	−
dhp77.56	+	−	−	−	−	−	−	−	−	−	−

[a] Bold type indicates primers that are specific for *B. anthracis*.

anthracis. Similar genomic islands have been observed when genomes of different strains of the same bacterial species have been compared (e.g., *Y. pestis* [8, 29] and *E. coli* [27]).

Multiplex PCR analysis for the simultaneous detection of four *B. anthracis*-specific loci. Multiplex PCR is a powerful tool for the simultaneous detection of multiple loci within bacterial genomes. The simultaneous detection of four separate loci (A, C, D, and E) on the *B. anthracis* genome was achieved here by selecting primers (shown in bold type in Table 5) that target these loci while providing sufficient size discrimination for resolution by gel electrophoresis on 4% agarose (163, 133, 196, and 241 bp, respectively). An internal positive control (288 bp) was designed from a region of the 23S gene conserved in *Bacillus* species (see Material and Methods) and confirmed the integrity of the DNA template. Figure 3 shows that the *Bacillus* 23S control primer yielded a 288-bp PCR product for all DNA templates (10 ng) tested. Furthermore, the four predicted *B. anthracis*-specific bands were seen for all strains (Fig. 3, lanes 13 to 23) and were absent from *B. cereus* and *B. thuringiensis* strains (lanes 2 to 11). This multiplex PCR was capable of detecting all four loci with as little as 100 pg of template DNA (data not shown).

DISCUSSION

B. anthracis is a potent mammalian pathogen and bioterrorist agent. This pathogenic *Bacillus* species shares so much genetic material with *B. cereus* and *B. thuringiensis* that its discrimination from the other species can be problematic. All three *Bacillus* species are prevalent in many environments, and it is important to define the genetic differences specific to *B. anthracis* in order to design specific DNA-based identification protocols. We surmised that the most efficient approach to finding *B. anthracis*-specific DNA sequences would be to find the *Bacillus* species that are most closely related to *B. anthracis* and then to compare their genomes in vitro by using SSH.

AFLP analysis was used to reveal genetic diversity among non-anthrax-causing *Bacillus* strains and to determine which are most closely related to the highly pathogenic *B. anthracis*. The phylogenetic tree shown in Fig. 1 indicates that *B. cereus* 3A, *B. cereus* S2-8, and *B. thuringiensis* 97-27 are the strains most genetically similar to *B. anthracis* of the 10 strains examined. *B. cereus* ATCC 4342 was shown previously by multienzyme electrophoresis (16) and AFLP analysis (38) to be very closely related to *B. anthracis*. The AFLP analysis presented here shows that seven strains are even more closely related to *B. anthracis*. The relationships do not correlate with the species designation for *B. cereus* or *B. thuringiensis*, providing another example of how the species designations for *B. cereus* and *B. thuringiensis* appear to cross the boundaries of various phylogenetic analyses (15, 38). It may be necessary to develop new criteria for species designations within this group, as information about more *Bacillus* genomes becomes available.

Although AFLP analysis provides a sensitive method for defining the genetic relationships between bacterial genomes, it provides no information regarding the genetic rearrangements responsible for these differences. Such information can be attained by in vitro genome comparison by SSH, a proven, efficient method for the identification of nucleotide sequences that differ between two genomes (1). The complementary tech-

TABLE 5. Primers designed from 28 *B. anthracis*-specific DNA signatures and their coordinates on the *B. anthracis* A2012 genome (GenBank accession number NC_003995)

Tester	Driver	Fragment no.	Fragment name[a]	Length of difference product (bp)	% G+C content	Primer sequence		Predicted PCR product (bp)	Results of BLASTX search against:								
									B. anthracis A2012 genome						GenBank database		
									Strand[b]	Coordinate		Accession no.	Gene synonym	E value	Description	Accession no.	E value
						Forward	Reverse			5′	3′						
B. anthracis Ames	*B. cereus* ATCC 14579	1	R.Ctg132	94	36.7	ACCTTTTGCTAAAGTGATTTCACCTTTTG	ACTGAAGTTGTTTCACCAGAGGGTATTGA	90	C	139668	1396592	NP_654829	BA_1469	3.0E-09	S-layer protein (*B. anthracis*)	P49051	3.0E-09
		2	R.Ctg139	268	35.6	CAAAGTGGCAAAAGGGGAATTTAGA	CGCCCAACAACTTGAGAGCCTACAA	157	C	3923298	3923035	NP_657605	BA_4245	6.0E-44	Hypothetical phage protein (*Streptococcus pyogenes*)	NP_606634	9.0E-05
		3	R.Ctg152	425	32.8	TCCCAATACATATGAGCGATTCGCC	GGTGGTATGTAAAAATCAAATAGGACGGA	155		4227290	4227710	NP_657938	BA_4578	4.0E-12	Novel		
		4	R.Ctg177	612	31.9	TTGGATCAGCGTTTCTGAATTCAGC	TCCCCATATCGCTCAATTCCATCTA	160		996045	996652	NP_654415	BA_1055	4.0E-88	Novel		
		5	M.Ctg032	342	34.7	GGCCCCATAACTTGATTTCTCTTCA	ATCTTGCTCCAGCTGAAGGAATGGC	151		344770	345109	NP_653725	BA_0365	2.0E-63	Galactosyltransferase (*Clostridium acetobutylicum*)	NP_349903	7.0E-08
		6	M.Ctg037	151	36.2	GAAATGAATGCAGAAAAATTAGAGCTTGCG	TACAAATGCTTGTTTACCTTCTTCTCCACC	135		990508	990656	NP_654413	BA_1053	1.0E-06	Novel		
B. anthracis Ames	*B. thuringiensis* Al Hakam	7	dhp61.159	420	31.7	AATTCTCATCCTACAGCATCTCC	TGAAGAAGCACCCTCTGTTG	151		4870037	4870457	NP_658648	BA_5288	1.0E-29	Novel		
		8	dhp61.178	479	31.3	GGTGGGTAAATGATGAGTGTTTC	TTTTTGATTTGACCAGCCTTTTTAG	145	C	349630	349152	NP_653729	BA_0369	1.0E-76	Novel		
		9	dhp61.181	1,597	36.8	TGTGGAAGAATCCACTGTCC	TCTGTCTCTTTCATTTCCTCAAC	110	C	4209359	4207762	NP_657910	BA_4550	1.0E-179	BH0961: unknown conserved protein (*Bacillus halodurans*)	NP_241827	4.0E-04
		10	**dhp61.183**	645	31.8	GAAGGACGATACAGACATTTATTGG	ACCGCAAGTTGAATAGCAAG	163	C	186609	185965	NP_653563	BA_0203	4.0E-42	Novel		
		11	dhp64.177	482	43.6	TCGCAAGATAAATGGGAAG	TCCTCCCGTACCAATCAC	101	C	3427212	3426731	NP_657087	BA_3727	8.0E-82	Penicillin binding protein (*Bacillus cereus*)	BAB69972	1.0E-67
		12	dhp64.188	531	35.8	TTGGTAAAGTAGAAGCATCTTGG	CCATCTTATCTGACCCTTGG	134	C	4203832	4203302	NP_657908	BA_4548	1.0E-100	Novel		
		13	dhp64.202	451	36.6	ATATGTGGTTAGGAGCCAATG	ACAAATCGAGTCCAATAATCACC	137		192470	192922	NP_653575	BA_0215	9.0E-87	Putative phage terminase (*Clostridium perfringens*)	NP_562034	5.0E-08
		14	dhp64.208	596	42.0	ACTACTCGACAAGGACGTTTGTATTTCC	CCATGTTTCCCATTTGCTTC	152	C	3951710	3951115	NP_657641	BA_4281	7.0E-83	Cytosine-specific methyltransferase (*Bacillus halodurans*)	NP_244375	2.0E-03
B. anthracis Ames	*B. cereus* 3A	15	**dhp77.002**	240	33.9	TGATATTTATGACCAAGATTCAATATACG	GCCATAGCTCAAGGTCAATAGG	133		980947	981187	NP_654400	BA_1040	5.0E-42	ϕSLT terminase large subunit (*Staphylococcus aureus*)	NP_075501	1.0E-06
		16	dhp77.003	229	41.6	CCAGGTCCATCCTATGTTGC	ACTACAATCCGCGTGTCTCC	178		3701825	3702023	Intergenic			Novel		
		17	dhp77.004	300	38.1	CACGTGAGTTTACACCATACGC	CTAGCACTTGCTCTCATTTCG	243		1292142	1292441	NP_654715	BA_1355	7.0E-52	Novel		
		18	dhp77.047	175	42.4	GGAACAACTTTCCCACAAGC	TCATTCGTGAACCCAATAAGC	146	C	965870	965695	NP_654367	BA_1007	1.0E-26	Hypothetical protein 3 *nanH* region (*Clostridium perfringens*)	I40868	2.0E-06
		19	dhp77.056	179	42.0	AGGAACGTGCGACGACTAGC	CACCCGATATGTTCAACAGG	167	C	3927774	3927596	Intergenic			Novel		
		20	dhp73.001	490	32.2	TTTCTTATCCAACAGGTGTATGC	GCGAGTGTGACGTTGATAAGC	211	C	699664	699175	NP_654100	BA_0740	2.0E-15	Glucosamine synthetase (*Bacillus subtills*)	BAA04741	2.1E-12
		21	dhp73.009	438	30.6	TCAAACGTTGCCGAATAAGC	GACGATTCGACACAACATATCG	166	C	3223428	322991	NP_656851	BA_3491	3.0E-22	Novel		
		22	dhp73.011	387	33.3	TTGGCATTGTGTTAGATGAGG	TGTGATAAAGCCAAAGCAACC	204		2472346	2472732	NP_655966	BA_2626	1.0E-67	ATP binding cassette transporter (ATP binding protein) (*Clostridium perfringens*)	NP_561360	5.0E-32
		23	dhp73.016	467	28.9	CAGAGCGTTTCTTAAATGAAGAGG	TTCATCATATGCCTCAATTTGG	166	C	3945592	3945125	NP_657634	BA_4274	8.0E-72	Novel		
		24	**dhp73.017**	497	33.8	AAAGGCGGTTTAGAATTTGG	TGCTGCTCTTTACCCATGC	241	C	4198535	4198039	NP_657898	BA_4538	1.0E-82	Novel		
		25	**dhp73.019**	327	38.5	TGTAAATGAACGCCTTGACC	CCGACTCCTTCTATCAATTCC	196		3920569	3920882	NP_657604	BA_4244	8.0E-45	Novel		
		26	dhp73.022	244	39.8	AAACAATGTTGGATGGTAAATGG	CCATTCGCAAATTTCACTCC	144	C	4205247	4205004	NP_657908	BA_4548	1.0E-42	Novel		
		27	dhp73.026	548	27.7	CCGAGCTTTAATGGAAATGG	CCTTTACCTTCAGCCTCTAGC	236		1472742	1473289	Intergenic			BH2522: unknown conserved protein (*Bacillus halodurans*)	NP_243388	1.0E-13
		28	dhp73.029	242	30.6	GGAGAAACAGATAGTGGTGAAAGC	ACTCTAAAGCCTTGTATTAGATTTCC	133							Novel		

[a] Bold type indicates primers used in the multiplex PCR assays (see Fig. 3).
[b] C, identity to the complementary strand.

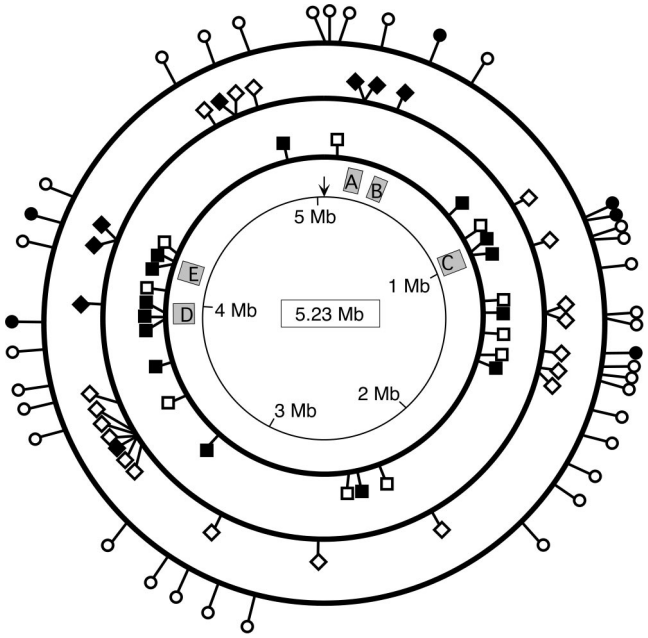

FIG. 2. Graphic representation of the locations of the tester-specific clones on the *B. anthracis* A2012 genome. The outermost circle maps the locations of the clones isolated from the subtraction by using DNA from *B. cereus* ATCC 14579 as the driver (circles), the second circle maps clones isolated from the subtraction by using DNA from *B. thuringiensis* Al Hakam as the driver (diamonds), and the third circle maps clones isolated from the subtraction by using DNA from *B. cereus* 3A as the driver (squares). The innermost circle shows the coordinates of the *B. anthracis* genome (5.23 Mb); the arrow shows the location of the first nucleotide (32). The sequences that are seen in *B. anthracis* but that are not seen in any non-anthrax-causing pathogens are represented by closed symbols. Gray boxes A to E, adjacent to the innermost circle, indicate the five regions that contain more than two *B. anthracis*-specific DNA sequences within 50 kb of each other; these represent putative *B. anthracis*-specific genomic islands.

niques of AFLP analysis and SSH provide a powerful means of defining accurate phylogenetic models and characterizing their underlying genetic components. Three subtractions were performed to identify *B. anthracis*-specific nucleotide sequences that were absent from the non-anthrax-causing *Bacillus* species. *B. cereus* ATCC 14579 represents the most distantly related strain in this study. *B. cereus* 3A was chosen because it represents one of the closest relatives of *B. anthracis*. The third strain, of intermediate relatedness, was *B. thuringiensis* Al Hakam. It would be expected that strains with more sequence identity to *B. anthracis* (in this situation, *B. cereus* 3A) would yield more sequences that would not be found in the closely related *B. cereus* and *B. thuringiensis*. Indeed, the subtraction with *B. cereus* 3A yielded 29% *B. anthracis*-specific sequences, compared to 17 and 2% for *B. thuringiensis* Al Hakam and *B. cereus* ATCC 14579, respectively (Table 2), confirming the prediction from AFLP analysis.

The genetic diversity demonstrated within this strain collection by AFLP analysis was mirrored by the results of subsequent PCR analysis. Bacterial evolution is driven by rearrangements of large genomic islands associated with lateral gene transfer (21, 23). Analysis of the G+C contents of the difference products reveals an average of 35.4%, typical of *B. anthracis*. This finding suggests that any lateral gene transfer has ameliorated the G+C content over a long period of evolution or has been received from species with a similar G+C content. The presence or absence of each of the tester-specific sequences in the non-anthrax-causing pathogens was determined by PCR amplification with primers designed for the difference products. If the number of loci shared by *B. anthracis* indicates relatedness among the whole panel of non-anthrax-causing *Bacillus* strains, then *B. cereus* 3A is the most closely related (Table 3). This result is in agreement with the phylogenetic analysis shown in Fig. 1. The phylogenetic tree shows two pairs of isolates that cannot be distinguished by AFLP analysis: *Bacillus* strains Vollum and Ames and *B. cereus* strains 3A and S2-8. SSH with *B. cereus* ATCC 14579 as a driver yielded two primer sets (M.Ctg056 and M.Ctg007) that can distinguish *B. cereus* 3A from S2-8 (Table 3). Primers from the same subtraction cannot distinguish *B. thuringiensis* HD-571 from Al Hakam. Table 4 demonstrates that the number of loci that are shared between the non-anthrax-causing *Bacillus* species and *B. anthracis* decreases when the driver strain is much more closely related to *B. anthracis*. This subtraction also yielded two primer sets (dhp73.03 and dhp73.04) that can distinguish *B. thuringiensis* HD-571 from Al Hakam, which are indistinguishable by the primer sets listed in Table 3.

The most important criteria for effective DNA signatures are the absence of false-positive results with closely related organisms and their representation in all isolates of the target (i.e., no false-negative results). Given the monomorphic nature of the *B. anthracis* genome, it was not surprising to find that there was no variation in the signatures within the strains of the collection. Twenty-eight *B. anthracis*-specific candidates isolated in these experiments fulfilled these criteria and are listed

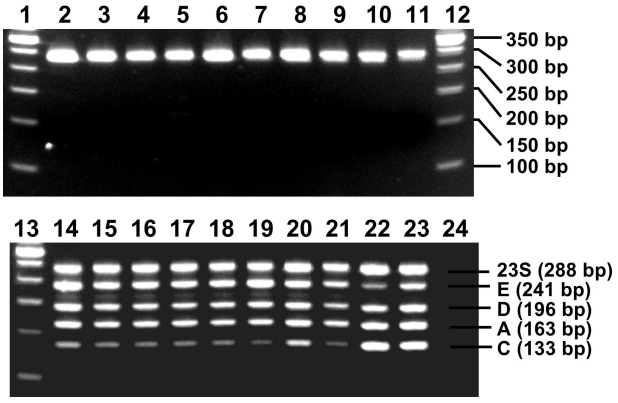

FIG. 3. Multiplex analysis of four separate loci on the *B. anthracis* genome. Four *B. anthracis*-specific primers yielded predicted products of 133, 163, 196, and 241 bp. One internal positive control yielded a predicted product of 288 bp. The DNA templates used for the multiplex analysis were as follows: lanes 1, 12, and 13, size markers as indicated; lane 2, *B. cereus* ATCC 14579; lane 3, *B. cereus* ATCC 4342; lane 4, *B. cereus* D17; lane 5, *B. cereus* 3A; lane 6, *B. cereus* S2-8; lane 7, *B. cereus* F1-15; lane 8, *B. thuringiensis* 97-27; lane 9, *B. thuringiensis* Al Hakam; lane 10, *B. thuringiensis* ATCC 10792; lane 11, *B. thuringiensis* HD-571; lane 14, *B. anthracis* G3; lane 15, *B. anthracis* G20; lane 16, *B. anthracis* G25; lane 17, *B. anthracis* G29; lane 18, *B. anthracis* G38; lane 19, *B. anthracis* G62 (Ames); lane 20, *B. anthracis* G67; lane 21, *B. anthracis* G77 (Vollum); lane 22, *B. anthracis* G80; lane 23, *B. anthracis* G87; and lane 24, no-template negative control.

in Table 5. None amplified a PCR product from any of the non-anthrax-causing *Bacillus* pathogens used in this study (no false-positive results). We were also able to exploit a collection of genetically distinct and geographically diverse isolates of *B. anthracis*. The twenty-eight DNA signatures amplified a PCR product of the predicted size for every isolate. A comparison of the DNA sequences against the completed genomes of *B. halodurans* and *B. subtilis* and the unfinished *Bacillus* genomes showed no significant sequence identity.

The non-anthrax-causing *Bacillus* species described here are so closely related to *B. anthracis* that they would be highly likely to give false-positive results in DNA-based identification assays based on chromosomal loci. The isolation of multiple *B. anthracis*-specific chromosomal regions allowed the development of a single multiplex assay for the rapid and highly specific detection of *B. anthracis*. The DNA signatures presented here have the advantage over previous detection methods that require time-consuming analysis (3, 7), are prone to false-positive results, and are based on few nucleotide differences at a single chromosomal locus (11, 28).

BLAST analysis of the nucleotide sequences of these DNA signatures shows that many are not represented in current DNA databases, other than the previously reported *B. anthracis* A2012 genome (Table 5) (32). The strongest identity seen was to the penicillin binding protein of *B. cereus* (dhp64.177; E value, 10^{-67}). R.Ctg122 showed identity to the S-layer protein of *B. anthracis* (13). This cell surface protein is also seen in some isolates of *B. cereus* and *B. thuringiensis* (22). However, there is sufficient nucleotide sequence divergence at the oligonucleotide primer binding sites to allow for successful discrimination of *B. anthracis* from the other two species.

The genomes of several strains of *B. anthracis* are currently being sequenced (12), and these data will be extremely useful for strain attribution in forensic analyses. We envisage that these DNA signatures can be used for real-time specific detection of *B. anthracis*, the source of which may then be attributed by monitoring the small nucleotide differences identified by these sequencing projects. The *B. anthracis*-specific DNA sequences identified in this work provide the largest collection of chromosomal markers that distinguish *B. anthracis* from other closely related *Bacillus* species. There are five regions in the *B. anthracis* genome where several of the specific DNA sequences are located within 50 kb of each other. Such genomic islands may define *B. anthracis* as a species and distinguish it from the closely related species *B. cereus* and *B. thuringiensis*. Future detailed analysis of these *B. anthracis*-specific regions may ultimately identify chromosome-encoded virulence factors, provide starting points for possible vaccine candidates, and help to reveal the mode of pathogenicity of this important pathogen.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy at the University of California (Lawrence Livermore National Laboratory, under contract no. W-7405-Eng-48). The Chemical and Biological National Security Program of the U.S. Department of Energy funded this work.

We are extremely grateful to Kimothy Smith for assistance with the *B. anthracis* collection used in this study and Nancy K. Montgomery for preparing genomic DNA. We also appreciate the technical contributions of Aubree Hubbell, Anne M. Erler, Silvia Gamez-Chin, Cheryl Strout, Julie R. Avila, and Linda L. Ott at Lawrence Livermore National Laboratory.

REFERENCES

1. **Agron, P. G., M. Macht, L. Radnedge, E. W. Skowronski, W. Miller, and G. L. Andersen.** 2002. Use of subtractive hybridization for comprehensive surveys of prokaryotic genome differences. FEMS Microbiol. Lett. **211:**175–182.
2. **Akopyants, N. S., A. Fradkov, L. Diatchenko, J. E. Hill, P. D. Siebert, S. A. Lukyanov, E. D. Sverdlov, and D. E. Berg.** 1998. PCR-based subtractive hybridization and differences in gene content among strains of *Helicobacter pylori*. Proc. Natl. Acad. Sci. USA **95:**13108–13113.
3. **Andersen, G. L., J. M. Simchock, and K. H. Wilson.** 1996. Identification of a region of genetic variability among *Bacillus anthracis* strains and related species. J. Bacteriol. **178:**377–384.
4. **Bell, C. A., J. R. Uhl, T. L. Hadfield, J. C. David, R. F. Meyer, T. F. Smith, and F. R. Cockerill III.** 2002. Detection of *Bacillus anthracis* DNA by LightCycler PCR. J. Clin. Microbiol. **40:**2897–2902.
5. **Bradley, K. A., J. Mogridge, M. Mourez, R. J. Collier, and J. A. Young.** 2001. Identification of the cellular receptor for anthrax toxin. Nature **414:**225–229.
6. **Carlson, C. R., D. A. Caugant, and A.-B. Kolstø.** 1994. Genotypic diversity among *Bacillus cereus* and *Bacillus thuringiensis* strains. Appl. Environ. Microbiol. **60:**1719–1725.
7. **Daffonchio, D., S. Borin, G. Frova, R. Gallo, E. Mori, R. Fani, and C. Sorlini.** 1999. A randomly amplified polymorphic DNA marker specific for the *Bacillus cereus* group is diagnostic for *Bacillus anthracis*. Appl. Environ. Microbiol. **65:**1298–1303.
8. **Deng, W., V. Burland, G. Plunkett III, A. Boutin, G. F. Mayhew, P. Liss, N. T. Perna, D. J. Rose, B. Mau, S. Zhou, D. C. Schwartz, J. D. Fetherston, L. E. Lindler, R. R. Brubaker, G. V. Plano, S. C. Straley, K. A. McDonough, M. L. Nilles, J. S. Matson, F. R. Blattner, and R. D. Perry.** 2002. Genome sequence of *Yersinia pestis* KIM. J. Bacteriol. **184:**4601–4611.
9. **Drobniewski, F. A.** 1993. *Bacillus cereus* and related species. Clin. Microbiol. Rev. **6:**324–338.
10. **Drum, C. L., S. Z. Yan, J. Bard, Y. Q. Shen, D. Lu, S. Soelaiman, Z. Grabarek, A. Bohm, and W. J. Tang.** 2002. Structural basis for the activation of anthrax adenylyl cyclase exotoxin by calmodulin. Nature **415:**396–402.
11. **Ellerbrok, H., H. Nattermann, M. Ozel, L. Beutin, B. Appel, and G. Pauli.** 2002. Rapid and sensitive identification of pathogenic and apathogenic *Bacillus anthracis* by real-time PCR. FEMS Microbiol. Lett. **214:**51–59.
12. **Enserink, M.** 2002. Microbial genomics. TIGR begins assault on the anthrax genome. Science **295:**1442–1443.
13. **Etienne-Toumelin, I., J. C. Sirard, E. Duflot, M. Mock, and A. Fouet.** 1995. Characterization of the *Bacillus anthracis* S-layer: cloning and sequencing of the structural gene. J. Bacteriol. **177:**614–620.
14. **Gordon, D., C. Abajian, and P. Green.** 1998. Consed: a graphical tool for sequence finishing. Genome Res. **8:**195–202.
15. **Helgason, E., D. A. Caugant, I. Olsen, and A.-B. Kolstø.** 2000. Genetic structure of population of *Bacillus cereus* and *B. thuringiensis* isolates associated with periodontitis and other human infections. J. Clin. Microbiol. **38:**1615–1622.
16. **Helgason, E., O. A. Okstad, D. A. Caugant, H. A. Johansen, A. Fouet, M. Mock, I. Hegna, and A.-B. Kolstø.** 2000. *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*—one species on the basis of genetic evidence. Appl. Environ. Microbiol. **66:**2627–2630.
17. **Hernandez, E., F. Ramisse, J. P. Ducoureau, T. Cruel, and J. D. Cavallo.** 1998. *Bacillus thuringiensis* subsp. *konkukian* (serotype H34) superinfection: case report and experimental evidence of pathogenicity in immunosuppressed mice. J. Clin. Microbiol. **36:**2138–2139.
18. **Jackson, P. J., E. A. Walthers, A. S. Kalif, K. L. Richmond, D. M. Adair, K. K. Hill, C. R. Kuske, G. L. Andersen, K. H. Wilson, M. Hugh-Jones, and P. Keim.** 1997. Characterization of the variable-number tandem repeats in *vrrA* from different *Bacillus anthracis* isolates. Appl. Environ. Microbiol. **63:**1400–1405.
19. **Keim, P., A. Kalif, J. Schupp, K. Hill, S. E. Travis, K. Richmond, D. M. Adair, M. Hugh-Jones, C. R. Kuske, and P. Jackson.** 1997. Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers. J. Bacteriol. **179:**818–824.
20. **Keim, P., L. B. Price, A. M. Klevytska, K. L. Smith, J. M. Schupp, R. Okinaka, P. J. Jackson, and M. E. Hugh-Jones.** 2000. Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J. Bacteriol. **182:**2928–2936.
21. **Lawrence, J. G., and H. Ochman.** 1997. Amelioration of bacterial genomes: rates of change and exchange. J. Mol. Evol. **44:**383–397.
22. **Mesnage, S., M. Haustant, and A. Fouet.** 2001. A general strategy for identification of S-layer genes in the *Bacillus cereus* group: molecular characterization of such a gene in *Bacillus thuringiensis* subsp. galleriae NRRL 4045. Microbiology **147:**1343–1351.
23. **Ochman, H., J. G. Lawrence, and E. A. Groisman.** 2000. Lateral gene transfer and the nature of bacterial innovation. Nature **405:**299–304.
24. **Okinaka, R., K. Cloud, O. Hampton, A. Hoffmaster, K. Hill, P. Keim, T. Koehler, G. Lamke, S. Kumano, D. Manter, Y. Martinez, D. Ricke, R.**

Svensson, and P. Jackson. 1999. Sequence, assembly and analysis of pX01 and pX02. J. Appl. Microbiol. **87:**261–262.
25. **Pannucci, J., R. T. Okinaka, R. Sabin, and C. R. Kuske.** 2002. *Bacillus anthracis* pXO1 plasmid sequence conservation among closely related bacterial species. J. Bacteriol. **184:**134–141.
26. **Patra, G., P. Sylvestre, V. Ramisse, J. Therasse, and J. L. Guesdon.** 1996. Isolation of a specific chromosomic DNA sequence of *Bacillus anthracis* and its possible use in diagnosis. FEMS Immunol. Med. Microbiol. **15:**223–231.
27. **Perna, N. T., G. Plunkett III, V. Burland, B. Mau, J. D. Glasner, D. J. Rose, G. F. Mayhew, P. S. Evans, J. Gregor, H. A. Kirkpatrick, G. Posfai, J. Hackett, S. Klink, A. Boutin, Y. Shao, L. Miller, E. J. Grotbeck, N. W. Davis, A. Lim, E. T. Dimalanta, K. D. Potamousis, J. Apodaca, T. S. Anantharaman, J. Lin, G. Yen, D. C. Schwartz, R. A. Welch, and F. R. Blattner.** 2001. Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. Nature **409:**529–533.
28. **Qi, Y., G. Patra, X. Liang, L. E. Williams, S. Rose, R. J. Redkar, and V. G. DelVecchio.** 2001. Utilization of the *rpoB* gene as a specific chromosomal marker for real-time PCR detection of *Bacillus anthracis*. Appl. Environ. Microbiol. **67:**3720–3727.
29. **Radnedge, L., P. G. Agron, P. L. Worsham, and G. L. Andersen.** 2002. Genome plasticity in *Yersinia pestis*. Microbiology **148:**1687–1698.
30. **Radnedge, L., S. Gamez-Chin, P. M. McCready, P. L. Worsham, and G. L. Andersen.** 2001. Identification of nucleotide sequences for the specific and rapid detection of *Yersinia pestis*. Appl. Environ. Microbiol. **67:**3759–3762.
31. **Radosevich, J. L., W. J. Wilson, J. H. Shinn, T. Z. DeSantis, and G. L. Andersen.** 2002. Development of a high-volume aerosol collection system for the identification of air-borne micro-organisms. Lett. Appl. Microbiol. **34:** 162–167.
32. **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.
33. **Rozen, S., and H. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132:**365–386.
34. **Schnepf, E., N. Crickmore, J. Van Rie, D. Lereclus, J. Baum, J. Feitelson, D. R. Zeigler, and D. H. Dean.** 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. **62:**775–806.
35. **Skowronski, E. W., N. Armstrong, G. Andersen, M. Macht, and P. M. McCready.** 2000. Magnetic microplate-format plasmid isolation protocol for high-yield, sequencing grade DNA. BioTechniques **29:**786–790.
36. **Smith, N. R., R. E. Gordon, and F. E. Clark.** 1952. Aerobic spore forming bacteria. U. S. Dep. Agric. Monogr. **16:**1–148.
37. **Thwaite, J. E., L. W. Baillie, N. M. Carter, K. Stephenson, M. Rees, C. R. Harwood, and P. T. Emmerson.** 2002. Optimization of the cell wall microenvironment allows increased production of recombinant *Bacillus anthracis* protective antigen from *B. subtilis*. Appl. Environ. Microbiol. **68:**227–234.
38. **Ticknor, L. O., A.-B. Kolstø, K. K. Hill, P. Keim, M. T. Laker, M. Tonks, and P. J. Jackson.** 2001. Fluorescent amplified fragment length polymorphism analysis of Norwegian *Bacillus cereus* and *Bacillus thuringiensis* soil isolates. Appl. Environ. Microbiol. **67:**4863–4873.
39. **Turnbull, P. C., R. A. Hutson, M. J. Ward, M. N. Jones, C. P. Quinn, N. J. Finnie, C. J. Duggleby, J. M. Kramer, and J. Melling.** 1992. *Bacillus anthracis* but not always anthrax. J. Appl. Bacteriol. **72:**21–28.
40. **Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, et al.** 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23:**4407–4414.