JOURNAL OF BACTERIOLOGY, Feb. 1997, p. 818–824 Vol. 179, No. 3
0021-9193/97/$04.00+0

Molecular Evolution and Diversity in *Bacillus anthracis* as Detected by Amplified Fragment Length Polymorphism Markers

PAUL KEIM,[1,2]* ABDULAHI KALIF,[3] JAMES SCHUPP,[2] KAREN HILL,[1] STEVEN E. TRAVIS,[2] KARA RICHMOND,[1] DEBRA M. ADAIR,[1] MARTIN HUGH-JONES,[3] CHERYL R. KUSKE,[1] AND PAUL JACKSON[1]

Environmental Molecular Biology, Los Alamos National Laboratory, Los Alamos, New Mexico 87545[1]; Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011-5640[2]; and Department of Epidemiology and Community Health, School of Veterinary Medicine, Louisiana State University, Baton Rouge, Louisiana 70803[3]

Received 27 June 1996/Accepted 11 November 1996

Bacillus anthracis **causes anthrax and represents one of the most molecularly monomorphic bacteria known. We have used AFLP (amplified fragment length polymorphism) DNA markers to analyze 78 *B. anthracis* isolates and six related *Bacillus* species for molecular variation. AFLP markers are extremely sensitive to even small sequence variation, using PCR and high-resolution electrophoresis to examine restriction fragments. Using this approach, we examined ca. 6.3% of the *Bacillus* genome for length mutations and ca. 0.36% for point mutations. Extensive variation was observed among taxa, and both cladistic and phenetic analyses were used to construct a phylogeny of *B. anthracis* and its closest relatives. This genome-wide analysis of 357 AFLP characters (polymorphic fragments) indicates that *B. cereus* and *B. thuringiensis* are the closest taxa to *B. anthracis*, with *B. mycoides* slightly more distant. *B. subtilis*, *B. polymyxa*, and *B. stearothermophilus* shared few AFLP markers with *B. anthracis* and were used as outgroups to root the analysis. In contrast to the variation among taxa, only rare AFLP marker variation was observed within *B. anthracis*, which may be the most genetically uniform bacterial species known. However, AFLP markers did establish the presence or absence of the pXO1 and pXO2 plasmids and detected 31 polymorphic chromosomal regions among the 79 *B. anthracis* isolates. Cluster analysis identified two very distinct genetic lineages among the *B. anthracis* isolates. The level of variation and its geographic distribution are consistent with a historically recent African origin for this pathogenic organism. Based on AFLP marker similarity, the ongoing anthrax epidemic in Canada and the northern United States is due to a single strain introduction that has remained stable over at least 30 years and a 1,000-mile distribution.**

Bacillus anthracis has been taxonomically aligned with *B. cereus*, *B. thuringiensis*, *B. megaterium*, and *B. mycoides* based on morphological and DNA similarities (reviewed in reference 4). All five species have been placed into *Bacillus* subgroup 1 based on their relatively large cell width (>1 μm) and spores that do not distend the sporangia (6). A comparison of the rRNA gene sequences is also consistent with this taxonomic arrangement (2, 3). DNA studies have further united *B. anthracis*, *B. cereus*, *B. mycoides*, and *B. thuringiensis*, which all have very AT-rich genomes (14) and high homology (16). It has even been suggested that these four species should be reorganized into a single species with four subspecies (21).

In addition to its chromosome, *B. anthracis* contains two large plasmids that carry essential genes for pathogenesis. The pXO1 plasmid is ca. 174 kb in size and carries the *pag*, *lef*, and *cya* toxin genes (12). The pXO2 plasmid is ca. 95 kb in size and carries the capsule genes *capA*, *capB*, and *capC* (12). Few other genes have been mapped to these very large plasmids. For anthrax virulence to be fully manifested, both plasmids must be present. Traditional vaccine strains (e.g., Sterne) carry only the pXO1 plasmid, while strains that carry only the pXO2 plasmid (e.g., Pasteur) have been isolated. Strains that carry neither plasmid have been isolated from the field (20) and have been created by growth conditions in the laboratory (23). Plasmid transfer between strains has been accomplished in the laboratory, and horizontal transfer among natural populations of *Bacillus* species seems very feasible.

Previous attempts at molecular characterization and DNA fingerprinting of *B. anthracis* strains have revealed a lack of diversity in this species. Exact DNA identity among strains has been observed in direct DNA sequencing of the 16S RNA and of the genic regions between the 16S and 23S rDNAs and between *gyrA* and *gyrB* (8). Restriction fragmentation patterns for 18 restriction enzymes, PCR patterns for 10 arbitrary primers, and PCR patterns for a bacterial repetitive element primer were found to be identical for 35 *B. anthracis* strains (9). Finally, use of an M13 phage primer as an arbitrary PCR primer identified one DNA fragment difference among strains (9). This primer was subsequently used to isolate and sequence a variable number tandemly repeated sequence called the *vrrA* locus from *B. anthracis* and other *Bacillus* species (1, 10). This sequence contained a perfect 12-bp repeat that was present in different copy numbers to generate the polymorphic DNA fragments. Prior to the research presented here, the *vrrA* locus was the only known variable DNA sequence among *B. anthracis* strains.

Molecular marker analysis is becoming increasingly capable of identifying informative genetic variation. Amplified fragment length polymorphism (AFLP [a trademark of KeyGene International]) markers are among the most recent innovations in genetic marker technologies and provide a greater capacity for genome coverage and more reproducible results than previous technologies (22). These markers are based on restriction fragment polymorphisms that are detected by selec-

* Corresponding author. Mailing address: Department of Biology, Northern Arizona University, Flagstaff, AZ 86011-5640. Phone: (520) 523-1078. Fax: (520) 523-7500. E-mail: Paul.Keim@NAU.EDU.

TABLE 1. *B. anthracis* cultures used in this study[a]

Source	Country (state) of origin	Strain code
ADRI/Lethbridge	Canada	62W-8, 67C-174, 72-235C, 72-284W, 73-237L-1, 74-389C-52, 74-402C-A, 74-412C-8, 78-306W, 78-345W, 80-77C-A, 81-173C-2, 91-316C-3, 91-323W-3, 91-334W-2, 91-383C, 91-429C-2, 93-194C, 93-195C-8, 93-196C-1, 93-206C, 93-212C-2, 93-213C-1, 94-188C-1
CVL/Oslo	Norway	B1965/77, B286/76, B648/82, B6273/93, B7227/83
OBP/Onderstepoort	Namibia	SA0573, SA1189, SA1225, SA1265, SA4045
VDL/UN-Lincoln	U.S. (NE)	UNL24673.94
TAMU/College Station	U.S. (TX)	C93022281, C94275042
UC/Davis	U.S. (CA)	D9106771, D9106853, D9107191, D9106955
USAMRIID	Haiti	BA1017
	Ireland	BA1024
	Pakistan	BA1009/1023, BA1021
	Scotland	BA1087
	South Africa	BA1018, BA1031, BA1033, BA1043 (Sterne), BA1035, #33, #58, #83
	?U.K.	BA1000 (Vollum)
	U.S. (CO)	BA1040
	U.S. (IA)	BA1007, #39
	U.S. (KS)	#26
	U.S. (MD)	BA1015
	U.S. (MI)	#37
	U.S. (NH)	BA0078/1008
	U.S. (NC)	#109
	U.S. (NY)	#47
	U.S. (OH)	#20, #28
	U.S. (SD)	#25
	U.S. (TX)	#38
	U.S. (UT)	BA1002 (Vollum-1b)
	Zimbabwe	BA1086
USAMRIID/CDC	Argentina	BA0015
	Canada	BA0018
	Jamaica	BA0052
	Lebanon	BA0003
	Pakistan	BA0006
VRL/Harare	Zimbabwe	ZIM69
K. Wilson		Ames*, Sterne*, Vollum* (cured of pXO1 and pXO2 at USAMRIID)

[a] Contributing institutions: ADRI/Lethbridge, Animal Disease Research Institute, Agriculture & Agri-Food Canada, Lethbridge, Alberta, Canada; CVL/Oslo, Central Veterinary Laboratory, Royal Ministry of Agriculture, Oslo, Norway; OBP/Onderstepoort, Onderstepoort Biological Products, Onderstepoort, South Africa; VDL/UN-Lincoln, Veterinary Diagnostic Laboratory, University of Nebraska, Lincoln; TAMU/College Station, State Veterinary Medical Diagnostic Laboratory, College of Veterinary Medicine, Texas A&M University, College Station; UC/Davis, California Veterinary Diagnostic Laboratory System, University of California, Davis; USAMRIID, U.S. Army Medical Research Institute for Infectious Diseases, Fort Detrick, Frederick, Md.; USAMRIID/CDC, the Centers for Disease Control and Prevention, cultures originally from Atlanta, Ga., maintained at USAMRRIID; VRL/Harare, Veterinary Research Laboratory, Causeway, Harare, Zimbabwe; K. Wilson, Durham V.A. N.C. Medical Center/Duke University, Durham, N.C. Abbreviations: U.S., United States; U.K., United Kingdom. States are indicated by postal code abbreviations.

tive PCR. This is accomplished by ligation of DNA adapters to the ends of restriction fragments (commonly *Eco*RI and *Mse*I) followed by amplification using adapter-homologous primers (+0/+0 primers). The selectivity is accomplished by adding arbitrary nucleotides to the primer 3′ ends, which reduces the resulting fragment complexity. When one nucleotide is added to each primer (+1/+1 primers), between 50 and 100 fragments are detected from a typical bacterial genome (5). DNA fragment pattern complexity can be reduced, as necessary, by adding arbitrary nucleotides to the 3′ ends of the primers. Electrophoresis using denaturing polyacrylamide DNA sequencing gels separates these fragments and identifies length differences as small as one nucleotide and variation associated with restriction sites. Many bacterial species are highly diverse, and AFLP markers offer the potential to evaluate many polymorphic loci on a single gel (5). For a relatively monomorphic bacterial species like *B. anthracis*, this technical approach allows the evaluation of a large percentage of the genome to identify rare genetic variation.

In this paper, we report the molecular characterization of *B. anthracis* and selected related *Bacillus* species by using AFLP markers. Our analysis defines the phylogenetic relationships among these selected species and discriminates among *B. anthracis* isolates by using 31 polymorphic genomic regions. The molecular variation observed is consistent with a recent origin for *B. anthracis* or with epidemiological episodes that have reduced variation due to population bottlenecks. Two very distinct genetic *B. anthracis* lineages are identified among 33 unique types from the 79 isolates examined. This report provides a new level of genetic discrimination in a highly monomorphic pathogen.

MATERIALS AND METHODS

Strains. Seventy-eight *B. anthracis* isolates and six related *Bacillus* taxa were obtained for analysis and included isolates generously provided by many different laboratories. Type strains from the American Type Culture Collection were also included (Table 1). Three *B. anthracis* strains previously cured of their plasmids (1, 23) by the U.S. Army Research Institute for Infectious Diseases were used to complement our identification of plasmid AFLP fragments in the analysis. With the exception of the diverse taxa and the three plasmid-cured *B. anthracis* strains, all strains were placed into the DNA analysis after being given randomly assigned

codes. Consequently, AFLP marker analysis and subsequent data analyses were performed blindly by one of the authors without knowing the isolates' identities.

Culture conditions. Bacterial cultures were grown on sheep blood agar for at least 16 h at 37°C. Single colonies were selected and used to inoculate nutrient broth. Incubation was continued at 37°C until turbidity was observed. Growth was monitored by colony counts and cultures harvested at the 10^8 to 10^9 CFU/ml (24). The isolates that were not received in pure cultures were first grown on PLET selective medium (13). Single colonies were then picked, grown on sheep blood agar, and inoculated into nutrient broth as described above.

DNA isolation. Cells were harvested from the broth cultures by centrifugation at 5,000 × *g* for 10 min at 4°C. The pellet was resuspended in 8.5 ml of TE buffer (10 mM Tris [pH 8.0], 1 mM EDTA) by repeated pipetting. The resuspended pellet was frozen in liquid nitrogen for 5 min and then warmed in a 65°C water bath for 5 min. The freezing-warming process was repeated two more times. Then 450 μl of 10% sodium dodecyl sulfate and 45 μl of proteinase K (20 mg/ml) were added and mixed thoroughly, and the mixture was incubated for 1 h at 42°C. Following this, 1.5 ml of a 5 M NaCl was added to the suspension and mixed thoroughly. Then 1.2 ml of a solution of 10% (wt/vol) hexadecyltrimethlyammonium bromide (Sigma Chemical Co., St. Louis, Mo.) in 0.7 M NaCl was added and mixed thoroughly, and the mixture was incubated for 10 min at 65°C. An equal volume of chloroform-isoamyl alcohol (24:1) was added to the suspension and mixed by inversion, and the mixture was centrifuged at 5,000 × *g* for 5 min. The upper viscous phase was collected, phenol-chloroform-isoamyl alcohol (25:24:1) was added and mixed by inversion, and the mixture was centrifuged at 5,000 × *g* for 5 min. The upper phase was collected, an equal volume of chloroform-isoamyl alcohol (24:1) was added and mixed by inversion, and the mixture was centrifuged at 5,000 × *g* for 5 min. The upper phase was extracted once more with chloroform-isoamyl alcohol (24:1) and collected, and the DNA was precipitated by addition of 0.6 volumes of isopropyl alcohol. The DNA precipitate was collected with a glass rod and then washed in cold 70% ethyl alcohol prior to suspension TE buffer. Purified plasmid pXO1 and pXO2 DNAs were purchased from Donald L. Robertson (12).

AFLP markers. AFLP markers were detected as described previously by Travis et al. (19) and Vos et al. (22), with the following modifications. Briefly, 100 ng of DNA was digested with *Eco*RI and *Mse*I and then ligated to *Eco*RI and *Mse*I double-stranded adapters (5′-CTCGTAGACTGCGTACC-3′ plus 3′-CTGACG CATGGTTAA-5′ and 5′-GACGATGAGTCCTGAG-3′ plus 3′-TACTCAGGA CTCAT-5′, respectively). The digested and ligated DNA was then amplified by PCR (30 cycles) using the *Eco*RI and *Mse*I +0/+0 primers 5′-GTAGACTGCG TACCAATTC-3′ and 5′-GACGATGAGTCCTGAGTAA-3′ in a final volume of 50 μl. The +0/+0 PCR product was diluted 10-fold, and only 1 μl was used in subsequent +1/+1 reactions. Each isolate was analyzed by using all 16 possible +1/+1 primer combinations. AFLP products were separated on a 6% acrylamide DNA sequencing gel as described previously (19).

Data analysis. AFLP markers were used as genetic characters to determine the relationships among different *Bacillus* taxa and strains. Plasmid fragments from pXO1 and pXO2 were determined by comparison to purified standards. *Bacillus* species were studied by using 357 AFLP markers from two primer combinations (C/G and G/A) that were scored as present (1) or absent (0). AFLP fragments that could be attributed to pXO1 or pXO2 were not included in these analyses. Among-taxa data were analyzed phenetically by calculating a Dice genetic distance followed by unweighted pair group mathematical average (UPGMA) cluster analysis using the NTSYS software package (15). Because of the great diversity observed among *Bacillus* taxa, we used the Dice coefficient, which assumes that two taxa are the same if they share the presence of a character but not if they share the lack of a character. The genetic character data were also analyzed cladistically by using the maximum parsimony algorithm (branch and bound) in the software package PAUP (17). Likewise, *B. anthracis* strains were analyzed by scoring polymorphic fragments as present (1) or absent (0). However, in three cases the distribution of polymorphic fragments was consistent with a multiallelic locus. These were scored as A, B, C, and D to represent four different alleles at a single locus. Phenetic analysis of *B. anthracis* was calculated by using the genetic distance coefficient from PAUP, which assumes that two strains lacking a particular fragment are identical at that locus (percent shared fragments). This assumption will be valid in almost all cases due to the very low diversity among *B. anthracis* strains. Maximum parsimony analysis (PAUP, heuristic) of *B. anthracis* strains resulted in >1,400 equally parsimonious trees and is not reported here.

RESULTS

Species-level AFLP variation. AFLP markers were used to characterize molecular diversity among five *B. anthracis* strains and seven strains from six closely related taxa. These 13 *Bacillus* strains were analyzed with all 16 possible +1/+1 *Eco*RI-*Mse*I primer combinations. Purified DNAs from the pXO1 and pXO2 plasmids were also analyzed to identify plasmid-specific AFLP fragments in the various strains by comparison to the three strains cured of these plasmids (Fig. 1). Plasmid fragments from pXO1 and pXO2 were not included in subsequent analyses to focus on chromosome-based relationships and to avoid complications from the loss of these plasmids from some strains. We found that the intensity of the fragments decreased as the number of detectable fragments increased for particular +1/+1 primer combinations. Two of the best-quality primer combinations were G/C and G/A. These primer combinations generated 357 scorable fragments across the seven taxa studied (Fig. 1). The quality of the AFLP fragments was lowest with high-A+T primer combinations, where the number of fragments observed was also the highest (Table 2).

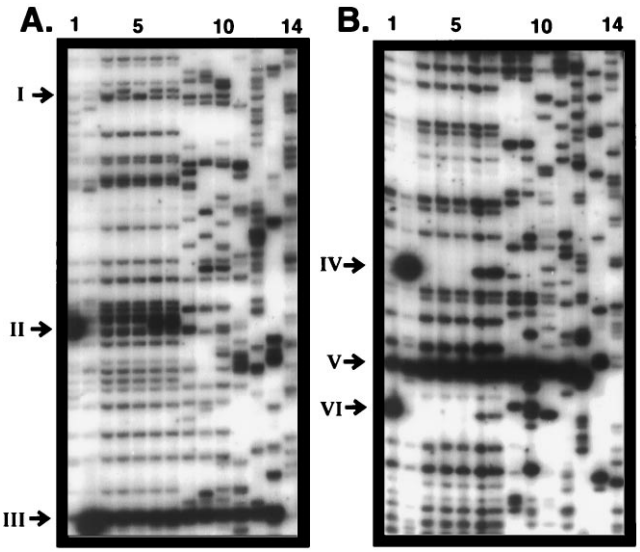

FIG. 1. AFLP autoradiogram of *Bacillus* taxa related to *B. anthracis*. (A) AFLP primer combination C/G; (B) AFLP primer combination G/G. In both panels, lanes contain, from left to right, pXO1, pXO2, *B. anthracis* Ames*, *B. anthracis* Vollum*, *B. anthracis* Sterne*, *B. anthracis* B1965/77, *B. anthracis* B648/82, *B. cereus*, *B. thuringiensis* ATCC 10792, *B. thuringiensis* ATCC 35646, *B. mycoides*, *B. polymyxa*, *B. subtilis*, and *B. stearothermophilus*. I, the only *B. anthracis* AFLP polymorphism observed in the initial study; II and VI, pXO1-specific AFLP fragments; III and V, intense AFLP fragments conserved in many *Bacillus* taxa. III also marks a pXO2-specific AFLP fragment that is identical in size to an intense conserved AFLP fragment.

Very low AFLP diversity was observed among the five *B. anthracis* strains in this initial study. Once AFLP differences from the presence and absence of the pXO1 and pXO2 plasmids were discounted, the fragment monomorphism within *B. anthracis* is quite apparent (Fig. 1). The only polymorphic fragments observed within the five *B. anthracis* strains for any of the 16 different primer combinations (ca. 1,200 *B. anthracis* fragments [Table 2]) are shown in Fig. 1B. The low diversity in *B. anthracis* is in contrast to the diversity between *B. thuringiensis* strains ATCC 35646 and ATCC 10792, where 42% of the fragments differed (Fig. 2A). The low AFLP diversity of *B. anthracis* is consistent with previous molecular diversity studies (8, 9).

Very great AFLP diversity was observed among the *Bacillus* taxa in this study. *B. cereus*, *B. thuringiensis*, and *B. mycoides* are thought to be close relatives of *B. anthracis* based on 16S RNA sequences and phenotypic data (4). The AFLP similarities are consistent with these expected close relationships, though *B. cereus* and *B. thuringiensis* seem to be more closely related to each other than to *B. anthracis* (Fig. 2). The AFLP marker data are consistent with *B. mycoides* being slightly more distant from *B. anthracis* than *B. cereus* and *B. thuringiensis*, though this conclusion is based on only a single *B. mycoides* strain. *B.*

TABLE 2. Numbers and types of polymorphisms observed in this study

Primer[a]	No. of polymorphisms				
	All fragments			Polymorphic fragments	
	pXO1	pXO2	Total	Length	Null allele
A/A	11	1	112	0	0
A/C	3	1	65	1	1
A/G	3	1	51	0	1
A/T	9	5	100	4	2
C/A	3	2	114	1	0
C/C	2	1	70	0	1
C/G	2	1	52	3	2
C/T	11	4	81	2	1
G/A	7	3	92	2	0
G/C	1	0	55	0	0
G/G	1	1	39	1	0
G/T	1	0	81	0	2
T/A	9	5	107	0	2
T/C	2	2	48	1	2
T/G	2	2	58	0	0
T/T	6	3	96	1	1
Total	73	32	1,221	16	15

[a] *Eco*RI-*Mse*I +1/+1.

subtilis, *B. polymyxa*, and *B. stearothermophilus* were quite different from *B. anthracis* and from each other, sharing less than 20% of the AFLP fragments with other taxa (Fig. 2A). A few highly conserved fragments, shown in Fig. 1, are shared by most of the taxa. We have noted that some of these highly conserved fragments also appear more intense. This could be due to a higher genomic copy number than for other fragments (bands III and V in Fig. 1). The large number of copies of the evolutionary conservative rRNA genes may be the source of these fragments. The infrequency of shared AFLP fragments among the more distant taxa in this study provides little phylogenetic information about these relationships.

Strain-level AFLP variation. In a second study, 79 *B. anthracis* isolates were analyzed by using 16 different +1/+1 AFLP primer combinations. The goal of this study was to identify molecular characters to distinguish subgroups within this species and to understand its evolution. Again, we used purified pXO1 and pXO2 DNAs as standards to distinguish between plasmid and chromosomal AFLP fragments. All 16 *Eco*RI-*Mse*I +1/+1 primer combinations were used to detect variation in all 79 isolates. Over 1,200 AFLP fragments, including 73 pXO1-specific and 32 pXO2-specific fragments, were observed (Table 2). Given that the average AFLP fragment is ca. 225 bp long and that there are 12 nucleotides associated with the *Eco*RI and *Mse*I sites and +1 arbitrary nucleotides of each fragment, we have surveyed ca. 6.3% of the *B. anthracis* genome for length mutation and ca. 0.36% of the genome for point mutations. Likewise, about 8.5% of the pXO1 and pXO2 plasmids were examined for length mutations, given the 105 plasmid-specific fragments observed. About 0.48% of the plasmid sequences were examined for point mutations associated with the restriction sites. Overall, this genomic analysis included ca. 270,000 nucleotides for length changes and ca. 14,400 nucleotides for point changes. If summed across all 79 isolates analyzed, this amounts to >21 Mb of sequence examined for length changes and >1.1 Mb examined for point changes. AFLP markers provide a powerful tool to rapidly identify genetic variation in a large number of isolates in a monomorphic species.

While most of the *B. anthracis* strains in this study contained both pXO1 and pXO2, we used the 105 plasmid-specific AFLP fragments to determine which strains did not (Fig. 3). We confirmed that three strains (Sterne*, Ames*, and Vollum*) had been cured of both pXO1 and pXO2 (Fig. 1 and 3). Six additional strains were contained pXO1 but not pXO2 (Fig. 3). The pXO1 and pXO2 differences among strains may be related to function in the case of vaccine strains such as Sterne but may also be due to loss during laboratory maintenance. The presence or absence of these plasmids was not used as a genetic character for further discrimination among strains in this study.

In general, there was very little AFLP variation among the 79 *B. anthracis* isolates. There were a total of 1,221 fragments observed in *B. anthracis*, and 1,184 (97%) were monomorphic (Table 2). In contrast, nearly 60% of the fragments differed between *B. anthracis* and its nearest relatives, *B. cereus* and *B. thuringiensis* (Fig. 2). The 31 polymorphic chromosomal regions observed could be categorized as two different types: those whose polymorphic distributions were inversely correlated with other polymorphic fragments and those that were either present or absent in a particular isolate. For the former type, which we have designated length mutations, we scored the inversely correlated fragments as a single variable (i.e., a single locus with two or more observable alleles). The latter type we have designated a polymorphism with a null allele (i.e., a single locus with only one observable allele). The observed AFLP frequency is 3.0% per fragment across all 79 isolates, or <0.04% per fragment per isolate, with about half of these

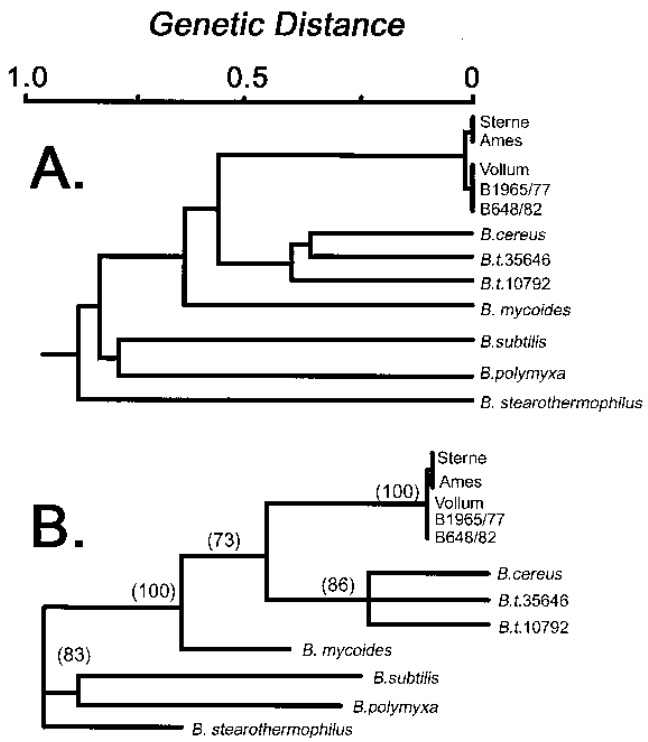

FIG. 2. UPGMA dendrogram of genetic distance among *Bacillus* taxa. (A) UPGMA cluster analysis based on 357 polymorphic AFLP fragments; (B) maximum parsimony analysis (branch and bound) based on the same 357 AFLP characters. *B. subtilis*, *B. polymyxa*, and *B. stearothermophilus* were used as outgroups in this analysis. Two equally parsimonious trees (77 steps; confidence interval = 0.78) that differed only in the arrangement of the three *B. cereus* and *B. thuringiensis* strains were found. The consensus tree is presented with bootstrap values (10,000 replications) in parentheses.

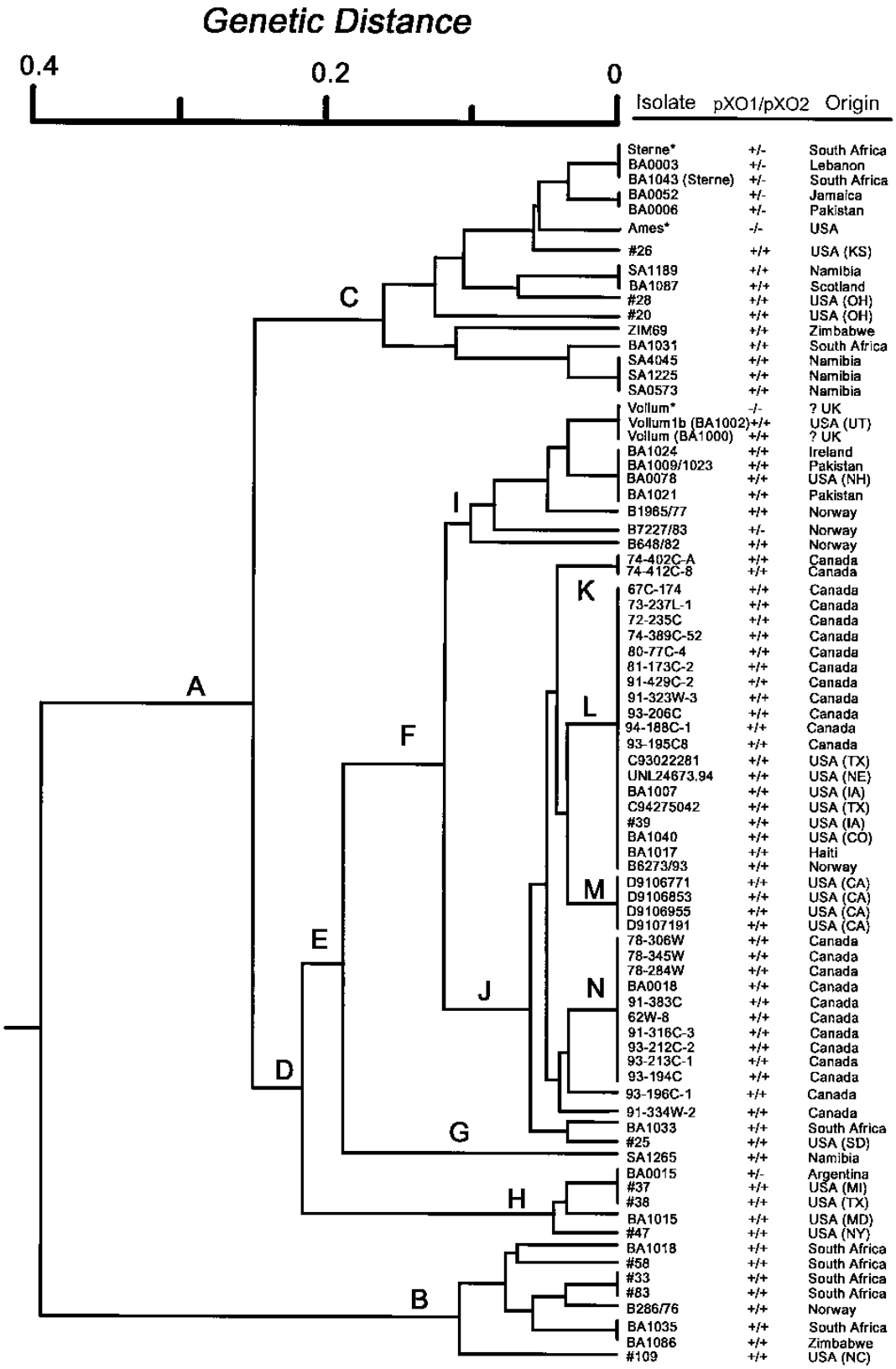

FIG. 3. UPGMA dendrogram of genetic distance among *B. anthracis* strains. The presence (+) or absence (−) of pXO1 and pXO2 plasmids is indicated. Particular nodes are labeled with letters to facilitate their discussion in the text. Note that the genetic distance values in Fig. 2 are not directly comparable to those on this figure because monomorphic fragments were not included in the genetic distance calculations shown in this figure.

being length mutations and half having null alleles (Table 2). Many of the variable regions in the 79 isolates studied were observed in only one strain and thus contain little information for discrimination among strains. Other variable region alleles were more equally distributed among the 79 isolates and provide greater discrimination power.

Genetic relationships among *B. anthracis* strains. Chromosomal AFLP markers were analyzed phenetically (UPGMA) to determine genetic relationships among *B. anthracis* strains. Cladistic analysis of these same data was consistent with the UPGMA results and is not reported. Perhaps the most notable of the genetic relationships identified are the two well-separated strain groups labeled A and B (Fig. 3). These represent a major subdivision in this species, with an average genetic distance between these groups of 0.4. The larger group A also contains distinct genetic clusters within it. Sterne and Ames strains are found within cluster C, while the three identical Vollum representatives are found in cluster I. The large number of very similar North American isolates reside in cluster J. These distinctive clusters are consistent with known strain variation and epidemiological relationships.

DISCUSSION

In this study, we have achieved a high degree of discrimination among isolates of *B. anthracis* by using AFLP molecular markers despite the relatively monomorphic nature of this species. Distinguishing strains is important for epidemiological analysis, clinical identification of particular strains, vaccine development and monitoring, and understanding the evolution of this pathogenic organism. The identification of 31 polymorphic genomic regions is the first step in development of diagnostic tools and in our molecular characterization of this organism.

Origin and evolution of *B. anthracis*. The extremely low level of molecular variation in the *B. anthracis* isolates examined in this study is consistent with three different hypotheses. (i) It is possible that *B. anthracis* has a greatly reduced rate of evolution that restricts the generation of molecular diversity compared to other *Bacillus* species. This seems unlikely due to the genomewide strategy that we used to detect variation and the presumption that AFLP markers will sample neutral as easily as nonneutral genomic regions. Therefore, we argue that *B. anthracis* molecular monomorphism must be due to either (ii) a very recent origin or (iii) population size constrictions (bottlenecks) associated with its worldwide spread.

A relatively monomorphic species would have resulted if *B. anthracis* was derived from a single *B. cereus* or *B. thuringiensis* strain in the recent past. This hypothesis revolves around the acquisition of the two virulence plasmids, pXO1 and pXO2, since both of these plasmids are required for *B. anthracis* to be fully pathogenic. A single horizontal transfer event where one plasmid was moved into a host cell containing the other plasmid would have created a genetic combination with a selective advantage over what was previously only a mildly pathogenic organism. The high lethality of anthrax, with a concurrent massive and efficient multiplication over a few days, and the absence of chronic anthrax infections provide little opportunity for host immunological selection for altered types. When *B. anthracis* is not growing in a host organism, it is thought to be quiescent as spores and, hence, perhaps evolving at a reduced rate. The worldwide expansion of a strain resulting from a recent single plasmid transfer would have resulted in the relative monomorphic pathogen that we observe today.

The third hypothesis is that anthrax is an ancient disease that has undergone successive epidemic episodes that emanate from successive infective foci. These foci representing population bottlenecks could be as small as a single anthrax death. These single infective events could contain all of the diversity found in the subsequent epidemic even after it had spread over great geographic distances. This appears to be representative of most North American isolates in this study, where many collections were nearly identical though they were collected across some 1,000 km and 20 years (see below). The recent origin and population bottleneck models are not mutually exclusive. Both may have shaped the current populations of *B. anthracis*, with all of the currently available isolates emanating from one or two historically recent anthrax outbreaks.

Africa may represent the anthrax center of diversity and, hence, the origin of *B. anthracis*. This is a continent of great herbivore populations where anthrax is currently endemic, and our analysis indicates great *B. anthracis* diversity. The A and B nodes in Fig. 3 are the most genetically distinct of any in our analysis, and they emanate from Africa. Currently, the strains isolated from North America, the Middle East, and Europe seem to be subsets of these two distinct lineages. These two groups may represent two independent epidemic foci from which the individual isolates have spread around Africa and the world (independent bottlenecks). More extensive sampling of Africa and Asia is currently under way to test this hypothesis. Independent and distinct *B. anthracis* lineages should radiate from the most common focus which could also be the origin of a relatively recently derived pathogen.

Molecular basis of the observed variation. Many of the variable AFLP fragments observed may be inherently unstable. Two types of polymorphic fragments (null alleles and length mutations) were observed in about equal numbers (Table 2). Null AFLP alleles may be due to several different molecular mechanisms, including point mutations, allele masking by other monomorphic fragments, and sequence rearrangements that create an unamplifiable distance between sites. In contrast, the length mutations are most probably due only to insertion or deletion events between the restriction sites. The polymorphic *B. anthracis vrrA* region (1, 10) is associated with a 12-nucleotide tandem arrayed sequence that expands and contracts over evolutionary time. A large percentage of the variation in this study may be due to a similar phenomenon. We have examined three AFLP markers by sequencing both alleles and have observed short repeated sequences in differing numbers in all cases (12a). Similar hypervariable phenomena have been observed in a wide range of organisms ranging from bacteria to humans (7, 11, 18).

Geographic correlation. We have noted that many of the isolates from Canada and United States are very similar (Fig. 3, node J). These samples were isolated from a variety of different animals (e.g., moose, bison, and bovine) separated by >1,000 km and collected over a 25-year period. This close genetic similarity suggests that this ongoing epidemic is the result of a single introduction and that *B. anthracis* is changing very little with the spread of the disease. Likewise, four isolates from California are similar to those found in node J but group together in node M. This is consistent with a common origin of the Californian outbreak and the ongoing Canadian-central U.S. epidemic. Even though we have argued for the rapid evolution of these AFLP markers, the alleles appear to be stable in the time periods important to tracking in an epidemic. The North American isolate similarity is contrasted by the great differences in Norwegian isolates. Three similar Norwegian isolates are in cluster I, but one other is found in the very distinct cluster B (Fig. 3). The great dissimilarity between these two types of Norwegian isolates suggests independent introduction of anthrax spores (possibly through commodity impor-

tation) rather than an ongoing epidemic emanating from a single introduction.

Polymorphic plasmid fragments. The most common type of variant AFLP fragments were those associated with the presence and absence of the pXO1 and pXO2 plasmids (105 fragments). These were manifested as null allele polymorphisms and identified by contrasting purified plasmid standards with plasmid-cured strains. In this analysis, we detected one AFLP fragment for about every 3.0 kb of DNA. This was true for the 87-kb pXO2 (32 fragments), the 175-kb pXO1 (73 fragments), and the ca. 4-Mb chromosome (1,116 fragments). This frequency indicates that the presence or absence of an unknown plasmid of >10 kb should also be readily detectable. Therefore, one hypothesis for the great difference between the A and B lineages (Fig. 3) could be the presence or absence of an unknown plasmid. However, the differences between the A and B clusters were a mixture of both length mutation and null allele markers (data not shown), arguing against an unknown plasmid-based difference. The polymorphism associated with the presence or absence of pXO1 and pXO2 plasmid fragments was ignored for much of this study because of the ephemeral nature of these plasmids.

ACKNOWLEDGMENTS

This work was partly supported by a sabbatical salary grant from Northern Arizona University and the Cowden Endowment for Food Microbiology to Paul Keim.

We thank Kenneth H. Wilson for a critical review of the manuscript. It is a pleasure to acknowledge the great generosity of all those who freely provided isolates: the Animal Disease Research Institute (Lethbridge, Alberta, Canada), Central Veterinary Institute (Oslo, Norway), Onderstepoort Biological Products (Onderstepoort, South Africa), Texas State Veterinary Medical Diagnostic Laboratory, California Veterinary Diagnostic Laboratory System, U.S. Army Medical Research Institute for Infectious Diseases, Veterinary Diagnostic Laboratory (Lincoln, Neb.), and Veterinary Research Laboratory (Harare, Zimbabwe), and especially Jack Burchak, John Ezell, Art Friedlander, and Huck Jaeger.

REFERENCES

1. **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.
2. **Ash, C., and M. D. Collins.** 1992. Comparative analysis of 23S ribosomal RNA gene sequences of *Bacillus anthracis* and emetic *Bacillus cereus* determined by PCR-direct sequencing. FEMS Microbiol. Lett. **73(1–3):**75–80.
3. **Ash, C., J. A. Farrow, M. Dorsch, E. Stackebrandt, and M. D. Collins.** 1991. Comparative analysis of *Bacillus anthracis*, *B. cereus*, and related species on the basis of reverse transcriptase sequencing of the 16S rRNA. Int. J. Syst. Bacteriol. **41:**343–346.
4. **Drobniewski, F. A.** 1993. *Bacillus* and related species. Clin. Microbiol. Rev. **6:**324–338.
5. **Geert, H., R. Coopman, P. Jansseen, and K. Kersters.** 1996. High-resolution genotypic analysis of the genus *Aeromonas* by AFLP fingerprinting. Int. J. Sys. Bacteriol. **46:**572–580.
6. **Gordon, R. E., W. C. Haynes, and C. H.-N. Pang.** 1973. The genus Bacillus. U.S. Department of Agriculture Agricultural Handbook no. 427. U.S. Government Printing Office, Washington, D.C.
7. **Greonen, O. N. A., A. E. Bunschoten, D. van Soolingen, and J. D. A. van Embden.** 1993. Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*: application for strain differentiation by a novel typing method. Mol. Microbiol. **10:**1057–1065.
8. **Harrell, L. J., G. L. Andersen, and K. H. Wilson.** 1995. Genetic variability of *Bacillus anthracis* and related species. J. Clin. Microbiol. **33:**1847–1850.
9. **Henderson, I., C. J. Duggleby, and P. C. B. Turnbull.** 1994. Differentiation of *Bacillus anthracis* from other *B. cereus* group bacteria with PCR. Int. J. Syst. Bacteriol. **44:**99–105.
10. **Jackson, P. J., K. L. Richmond, D. M. Adair, K. K. Hill, C. R. Kuske, E. Walthers, G. L. Andersen, K. H. Wilson, M. E. Hugh-Jones, A. S. Kalif, and P. Keim.** Characterization of the *vrrA* region that varies in different isolates of *Bacillus anthracis*. Submitted for publication.
11. **Jeffreys, A. J., V. Wilson, and S. L. Thein.** 1985. Individual-specific "fingerprints" of human DNA. Nature **316:**76–79.
12. **Kaspar, R. L., and D. L. Robertson.** 1987. Purification and physical analysis of *Bacillus anthracis* plasmids pXO1 and pXO2. Biochem. Biophys. Res. Comm. **149:**362–368.

12a. **Keim, P.** Unpublished data.

13. **Knisely, R. F.** 1966. Selective medium for *Bacillus anthracis*. J. Bacteriol. **92:**784–786.
14. **MacDonald, W. C., I. C. Felkner, A. Turetsky, and T. S. Matney.** 1963. Similarity in base composition of deoxyribonucleates from several strains of *Bacillus anthracis* and *B. cereus*. J. Bacteriol. **85:**1071–1073.
15. **Rohlf, F. J.** 1993. NTSYS-PC—numerical taxonomy and multivariate analysis system, version 1.8. Exeter Software, Setauket, N.Y.
16. **Seki, T., C. Chang, H. Mikami, and Y. Oshima.** 1978. Deoxyribonucleic acid homology of the genus *Bacillus*. Int. J. Syst. Bacteriol. **28:**182–189.
17. **Swofford, D. L.** 1991. PAUP: phylogenetic analysis using parsimony, verson 3.0s. Computer program distributed by the Illinois Natural History Survey, Champaign.
18. **Taylor, T. K., J. B. Bashiruddin, and A. R. Gould.** 1992. Application of a dynamic DNA probe for the differentiation of two types of *Mycoplasma mycoides* subspecies *mycoides*. Res. Vet. Sci. **53:**154–159.
19. **Travis, S. E., J. Maschinski, and P. Keim.** 1996. An analysis of genetic variation in *Astragalus cremnophylax* var. *cremnophylax*, a critically-endangered plant, using AFLP markers. Mol. Ecol. **5:**1–11.
20. **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.
21. **Turnbull, P. C. B., and J. M. Kramer.** 1991. *Bacillus*, p. 296–303. *In* A. Balows, W. J. Hausler, Jr., K. L. Herrmann, H. D. Isenberg, and H. J. Shadomy (ed.), Manual of clinical microbiology, 5th ed. American Society for Microbiology, Washington, D.C.
22. **Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van der Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau.** 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. **23:**4407–4424.
23. **Welkos, S. L.** 1991. Plasmid-associate virulence factors of non-toxigenic (pXO1-) *Bacillus anthracis*. Microb. Pathog. **10:**183–198.
24. **World Health Organization.** 1986. World Health Organization manual for the production of anthrax vaccine. World Health Organization, Geneva, Switzerland.