Differentiation of *Clostridium botulinum* Serotype A Strains by Multiple-Locus Variable-Number Tandem-Repeat Analysis[▽][†]

Thomas E. Macdonald,[1] Charles H. Helma,[1] Lawrence O. Ticknor,[2] Paul J. Jackson,[3]
Richard T. Okinaka,[1] Leonard A. Smith,[4] Theresa J. Smith,[4] and Karen K. Hill[1]*

*Divisions of Bioscience[1] and Computing, Computational and Statistical Sciences,[2] Los Alamos National Laboratory, Los Alamos,
New Mexico 87545; Defense Biology Division, Lawrence Livermore National Laboratory, Livermore, California 94551[3]; and
Integrated Toxicology Division, United States Army Medical Institute of Infectious Diseases (USAMRIID),
Fort Detrick, Maryland 21702[4]*

Ten variable-number tandem-repeat (VNTR) regions identified within the complete genomic sequence of *Clostridium botulinum* strain ATCC 3502 were used to characterize 59 *C. botulinum* strains of the botulism neurotoxin A1 (BoNT/A1) to BoNT/A4 (BoNT/A1-A4) subtypes to determine their ability to discriminate among the serotype A strains. Two strains representing each of the *C. botulinum* serotypes B to G, including five bivalent strains, and two strains of the closely related species *Clostridium sporogenes* were also tested. Amplified fragment length polymorphism analyses revealed the genetic diversity among the serotypes and the high degree of similarity among many of the BoNT/A1 strains. The 10 VNTR markers amplified fragments within all of the serotype A strains but were less successful with strains of other serotypes. The composite multiple-locus VNTR analysis of the 59 BoNT/A1-A4 strains and 3 bivalent B strains identified 38 different genotypes. Thirty genotypes were identified among the 53 BoNT/A1 and BoNT/A1(B) strains, demonstrating discrimination below the subtype level. Contaminating DNA within crude toxin preparations of three BoNT/A subtypes (BoNT/A1 to BoNT/A3) also supported amplification of all of the VNTR regions. These markers provide clinical and forensics laboratories with a rapid, highly discriminatory tool to distinguish among *C. botulinum* BoNT/A1 strains for investigations of botulism outbreaks.

Clostridium botulinum is an anaerobic spore-forming bacterium that produces the most toxic biological toxin known, botulinum neurotoxin (BoNT). Botulism poisoning can occur from ingestion of the toxin produced by the bacterium within improperly prepared food (food-borne) or by infection of the bacterium or its spores followed by growth and production of toxin under anaerobic conditions in the body, known as wound or infant botulism (5).

The species can be divided into four groups based on biochemical and biophysical parameters. Group I contains serotype A and proteolytic B and F strains, group II includes serotype E and nonproteolytic strains, group III strains produce serotype C and D toxins, and strains producing serotype G toxin are the lone members of group IV (6, 20). Strains can be further distinguished based on their ability to produce seven distinct toxin types (A to G) recognized by polyclonal antisera.

Historically, identification of a BoNT serotype has been based on neutralization of the toxin with polyclonal antisera by use of mice or immunological methods such as enzyme-linked immunosorbent assay (5, 14). These methods can detect and identify the serotype and sometimes the subtype of the organism but do not provide additional strain information. Other analytical methods applied to *C. botulinum* include pulsed-field gel electrophoresis (13, 17), randomly amplified polymorphic

DNA analysis (8), amplified fragment length polymorphism analysis (AFLP) (7, 11), repetitive-element sequence-based PCR (8), riboprinting (19), and amplified rRNA gene restriction analysis (18). These methods provide various degrees of discrimination, but they cannot provide significant resolution of individual strains beyond serotyping.

Human botulism cases are overwhelmingly due to one of three serotypes: A, B, or E. In the United States, 60% of the food-borne botulism cases and 45% of infant botulism cases are due to BoNT/A-producing *C. botulinum* strains (5). The ability to improve differentiation within serotype A strains, especially BoNT/A1 subtypes in the event of a botulism outbreak, would provide public health agencies with a valuable tool for determining the relatedness of strains within multiple outbreaks or for providing historical or geographical perspectives to previous outbreaks. Multiple-locus variable-number tandem-repeat analysis (MLVA) has been successfully used to differentiate among closely related strains of *Bacillus anthracis*, *Francisella tularensis*, *Yersinia pestis*, and *Brucella* species (4, 10, 12, 23) and have provided a powerful tool to interrogate forensic samples (9).

To further distinguish among BoNT/A1-producing strains, variable-number tandem-repeat (VNTR) regions were identified and used to distinguish among members of a collection of 59 BoNT/A strains, including representatives of the different BoNT/A1 to BoNT/A4 subtypes and two strains representing each of the serotypes B to G. Five bivalent strains and two nontoxic *C. sporogenes* strains were also included in the study. Here we describe the use of these VNTR markers to differentiate among these clostridial strains.

---

* Corresponding author. Mailing address: Bioscience Division, MS: M888, Los Alamos National Laboratory, Los Alamos, NM 87545. Phone: (505) 667-1309. Fax: (505) 665-3024. E-mail: khill@lanl.gov.
† Supplemental material for this article may be found at http://aem .asm.org/.
▽ Published ahead of print on 14 December 2007.

TABLE 1. Clostridial strains analyzed

Name[a]	Serotype	Strain	Source and/or year
A142	BoNT/A1	Schantz	
A143ˆ	BoNT/A1	ATCC 3502 (Hall 174)	Canned peas, Berkeley, California
A144	BoNT/A1	ATCC 17862	
A146	BoNT/A1	ATCC 25763	
A147*	BoNT/A1	CDC 1757 infant	Oregon, 1977
A150	BoNT/A1	Hall	
A256ˆ	BoNT/A1	Hall 5675	Canned cauliflower, Colorado, 1932
A312ˆ	BoNT/A1	Prevot Ppois	Peas, 1947
A384	BoNT/A1	CDC 297	
A385	BoNT/A1	Prevot P146	
A386~	BoNT/A1	VPI 7124	Soil, Virginia
A388ˆ	BoNT/A1	CDC 4997	Chicken livers
A389	BoNT/A1	ATCC 449	
A401	BoNT/A1	Hall 11481	
A405	BoNT/A1	McClung 447	1935
A412ˆ	BoNT/A1	McClung 465	Spinach, 1935
A413	BoNT/A1	Prevot 792	
A414	BoNT/A1	Prevot 910	Bovine, Crepy en Valois, France, 1953
A415	BoNT/A1	Prevot 969	1953
A416	BoNT/A1	Prevot 62NCA	1953
A417	BoNT/A1	Prevot P179 pike intestine	1953
A418	BoNT/A1	Prevot 878	
A419	BoNT/A1	Prevot 62	
A420	BoNT/A1	Prevot 865	Bovine liver, 1953
A421	BoNT/A1	Prevot F18	Bovine liver, 1953
A422	BoNT/A1	Prevot F16	Bovine liver, 1953
A423	BoNT/A1	Prevot F57	Liver, 1947
A424	BoNT/A1	Prevot Dewping	VPI 15,019
A425	BoNT/A1	Prevot F60	Liver, 1954
A427	BoNT/A1	Prevot 697B	Catgut, Stockholm, 1952
A428	BoNT/A1	Prevot F5G	Liver, 1954
A429	BoNT/A1	Prevot 892	
A487	BoNT/A1	ATCC 17916	
A503	BoNT/A1	McClung 844	
A504	BoNT/A1	McClung 450	1935
A505	BoNT/A1	McClung 457	1935
A674	BoNT/A1	ATCC 7948	
A148*	BoNT/A1(B)	CDC 1744 infant	Pennsylvania, 1977
A391ˆ	BoNT/A1(B)	Hall 183	Home-canned corn, Fort Collins, Colorado, 1922
A393	BoNT/A1(B)	Hall 3676	Human feces, November, 1929
A394ˆ	BoNT/A1(B)	Hall 3685a	Home-canned string beans, Pueblo, Colorado, 1929
A395ˆ	BoNT/A1(B)	Hall 4934Aa	Home-canned beans, Purcell, Colorado, 1931
A396ˆ	BoNT/A1(B)	Hall 4834	Home-canned spinach, Scottsbluff, Nebraska, 1931
A397ˆ	BoNT/A1(B)	Hall 8388A	Home-canned chile peppers, Springer, New Mexico, 1935
A398ˆ	BoNT/A1(B)	Hall 8857Ab	Home-canned corn, Scottsbluff, Nebraska, 1935
A402	BoNT/A1(B)	Hall 11569	
A403	BoNT/A1(B)	Hall 17544	
A404	BoNT/A1(B)	Hall 6581Ae	
A406	BoNT/A1(B)	McClung 452	1935
A407ˆ	BoNT/A1(B)	CDC 2084	Food-borne, New Mexico
A408	BoNT/A1(B)	CDC 7243	Culture, Indiana
A410	BoNT/A1(B)	CDC 8701	
A411	BoNT/A1(B)	CDC 2357	Stool, Colorado
A693*	BoNT/A2	FRI honey	
A694*	BoNT/A2	Kyoto-F	
Af695	BoNT/A2	Strain 84	
Ab149*	BoNT/A2	CDC 1436 infant	Utah, 1977
A254ˆ	BoNT/A3	Loch Maree	Wild-duck paste, United Kingdom, 1922
Ba207*	BoNT/A4	CDC 657 infant	1988
B152ˆ	BoNT/B	NCTC 7273	Beans
Bf258*	BoNT/B	An436	Sweden
B512ˆ	BoNT/B	Prevot 1687	Ham, 1957, nonproteolytic
Bf698	BoNT/B	CDC 3281	New Mexico, 1980

TABLE 1—*Continued*

Name[a]	Serotype	Strain	Source and/or year
C167	BoNT/C	Stockholm	1990
C210	BoNT/C	468	Continental Can Co., Chicago, Illinois, 1971
D175	BoNT/D	1873	1989
D177	BoNT/D	Schantz	
E185	BoNT/E	Alaska E43	
E213^	BoNT/E	Beluga	Fermented whale flippers, United States, 1992
F188^	BoNT/F	Langeland	Liver paste, Denmark, 1992
F550~	BoNT/F	Eklund202F	Marine sediments, Pacific Ocean, 1992, nonproteolytic
G194	BoNT/G	1354	
G196	BoNT/G	2740	Human, Switzerland, 1981
C. sporogenes 719		ATCC 19404	
C. sporogenes 720		ATCC 3584	

[a] Symbols: *, infant cases; ^, food-borne cases; ~, soil or marine sediment.

MATERIALS AND METHODS

Strains. Strains were kindly provided by the Integrated Toxicology Division, United States Army Medical Institute of Infectious Diseases (USAMRIID), Fort Detrick, MD, and Eric Johnson at the Department of Food Microbiology and Toxicology, University of Wisconsin, Madison, WI. Many of the strains formed part of the Virginia Polytechnic Institute Anaerobe Laboratory collection and were a kind gift of Myron Sasser. DNA preparations from the 73 *C. botulinum* strains and 2 *C. sporogenes* strains listed in Table 1 were prepared as previously described (7). These 73 strains are a subset of the 174 clostridial strains previously analyzed by AFLP and represent the seven different serotypes (A to G) (7). In addition, two non-toxin-producing *C. sporogenes* strains were included.

These *C. botulinum* strains were collected from outbreaks occurring as early as 1922 and represent food-borne and infant botulism cases as well as environmental samples. The serotype A strains tested included 53 BoNT/A1 strains, 4 BoNT/A2 strains, 1 BoNT/A3 strain, and 1 BoNT/A4 strain. Sixteen of the 53 BoNT/A1 isolates were A1(B) strains that contain an unexpressed, or silent, BoNT/B gene. The collection included five bivalent strains (Ab149, Ba207, Bf258, Bf698, and Af695), where the predominant toxin produced is designated

by the capital letter. One of the bivalent strains, Ba207, produces the unique subtype BoNT/A4.

VNTR identification. The genomic sequence of *C. botulinum* BoNT/A1 strain ATCC 3502 (http://www.sanger.ac.uk/Projects/Microbes/) was analyzed using the software Tandem Repeats Finder to identify VNTR regions (1). Primer sequences in Table 2 were designed flanking each VNTR region using Oligo 6.8 (Molecular Biological Insights, Inc., Cascade, CO). The primers were synthesized (Invitrogen Corp., Carlsbad, CA) with a fluorescent 6-carboxyfluorescein label on the forward primer.

VNTR analysis of samples. Each of the 10 VNTR primer sets was used individually, not multiplexed, to amplify each template. PCR amplification was performed in a final volume of 10 μl containing 1× PCR buffer, 0.25 mM $MgCl_2$, 0.25 mM deoxynucleoside triphosphates, 0.4 U Amplitaq Gold (Applied Biosystems Inc., Foster City, CA), 90 nM forward and reverse primers, and 1 ng of template DNA. Thermocycling conditions included an initial melt at 95°C for 10 min followed by 35 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min followed by a final extension step of 72°C for 5 min. After thermocycling, PCR products were diluted 1:60 in water and 1 μl of the diluted product was added to

TABLE 2. VNTR characteristics

VNTR no.	Genomic location[a]	Primer pair (forward and reverse)[b]	Repeat unit	Diversity index	Repeat unit size (bp)	No. of alleles	Amplicon size range
1	8769–8849	GCAATAAGAATAAATGTTTCG TACTGGCTGTTACAAGAATGT	GAAGAAAATTTAAAT	0.778	15	8	526–631
2	435471–435540	GAGGACGAGGGTAACTTAGT GGTACTTATGATTTCCGGTA	AGAGTAT	0.860	7	15	327–618
3	589115–589195	GTTATATTTGCTGTATTCTGTTT CCATTTATTTATGGGATATCT	TAGAAC	0.863	6	16	367–517
4	794502–794559	GATGCGGATATAGAGCTTT GTAGAATATAGTTTGCACCCT	GTAAAATAAATGTTAAA AAATAAGGAG	0.086	27	2	374–401
5	1192118–1192194	CATATAAGCCACAACAAGAAT CTTGATAAACCTTTCCTTTG	GATGGTAAAGATG TTGA	0.331	18	2	300–318
6	1452591–1452707	GAGGTGTAGTTATGAGAGATGG CTTTCATATGCTTCTCTTTCA	TATGATGGGAATGG ATGA	0.686	18	8	359–500
7	1620690–1620753	CAAATCCGTTAAAGAATGTAT CCATTGCTTTAGAAACTGAT	TAGATCTATAATAAAG AATT	0.187	20	2	505–545
8	1996310–1996377	CTACTCAAACAATAGACAGGC GTATTCCTGCTAATGTTCAAG	CTTGAATTTTACTATC CTTA	0.468	21	2	421–457
9	2271633–2271699	GTGCTAATACAGAACCTGCTA GAAGGTAAGATTATTGCATCA	TTGACCCTGGTTTG TCTG	0.653	18	4	370–424
10	594262–594371	GGTGCTTTTTCTGATACATTG GCGCTGAAGTAGTAAAGTCTAA	CTTCAGCTTCTTTTCTTT GAGCTT	0.593	24	5	268–436

[a] Location within *C. botulinum* ATCC 3502 (http://www.sanger.ac.uk/Projects/Microbes/).
[b] For each VNTR region, the forward primer is listed on the upper row and the reverse primer is listed on the lower row.
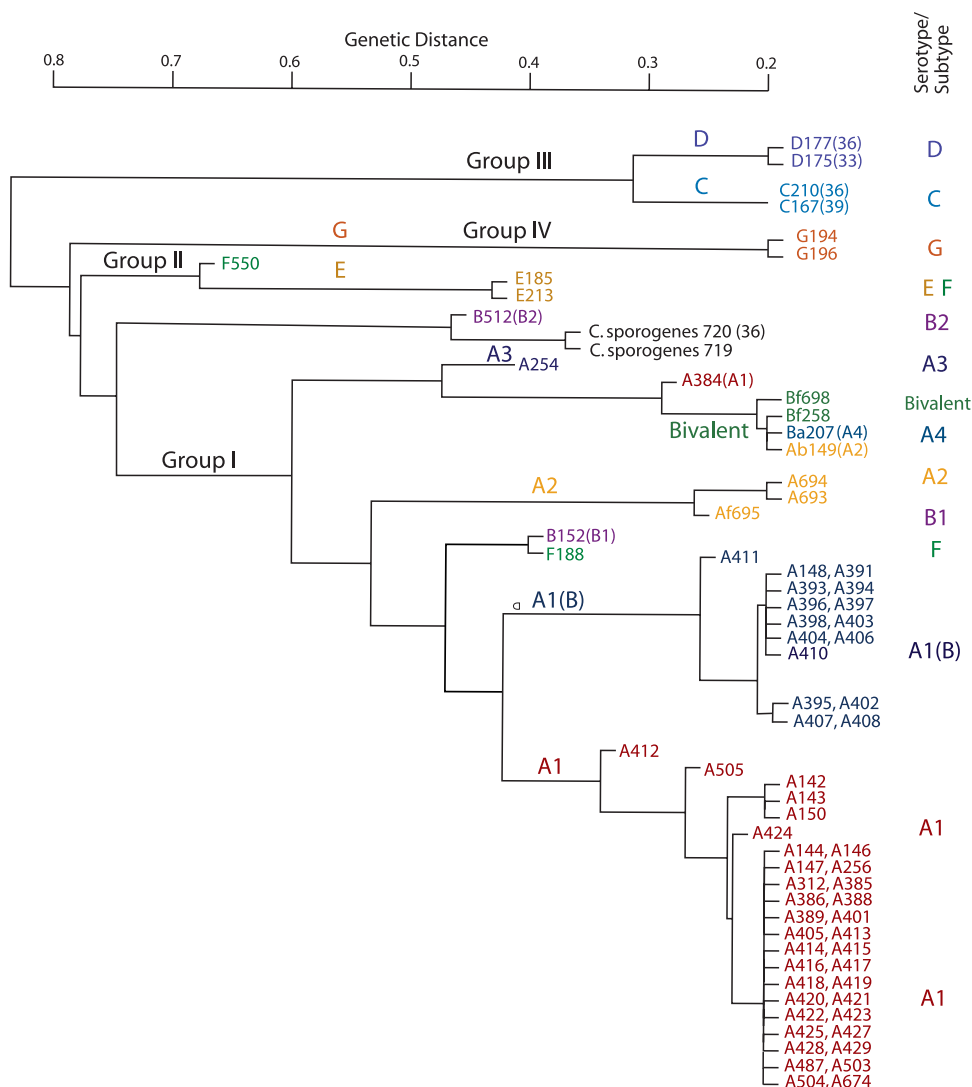
FIG. 1. AFLP-based dendrogram of 73 *C. botulinum* strains representing the seven serotypes A to G and two *C. sporogenes* strains. The dendrogram shows significant genetic diversity among the strains representing the different serotypes and illustrates the genetic similarity among many of the BoNT/A1 and A1(B) strains. Fragments generated from restriction endonuclease digestion of each of the strain DNAs were ligated into linkers and selectively amplified as previously described (7). Forty DNA fragments were used as a fingerprint to represent each strain. If forty fragments did not exist, fewer fragments were used, as noted in parentheses. The genetic distance between two samples is the proportion of the fragments that two samples do not have in common.

0.5 μl MapMarker 1000 molecular weight markers (Bioventures Inc., Murfreesboro, TN), labeled with X-rhodamine and 8.5 μl formamide. Following heat denaturation at 95°C for 12 min, the reactions were analyzed on an ABI 3130 automated fluorescent DNA sequencer (Applied Biosystems Inc., Foster City, CA).

Representatives of each of the different fragment sizes amplified by each VNTR primer pair were sequenced to verify the fragment length and tandem repeat unit. The PCR amplicons were purified using a QIAquick PCR purification kit (Qiagen, Valencia, CA) and then sequenced on an ABI 3730 automated fluorescent sequencer. Alignments were performed using Sequencher 4.7 (Gene Codes Corp., Ann Arbor, MI). Repeat unit sizes are reported in Table 2 in whole numbers.

The diversity index for each of the 10 VNTR loci was calculated for the strains shown in Fig. 3, and these diversity index values are reported in Table 2 (22). These values did not change significantly for calculations including only BoNT/A1 strains. The diversity index (D) for VNTR "*l*" was calculated as $D_l = 1 - \sum_{u} P_{lu}^2$ where P_{lu} is the frequency of the *u*th allele at the *l*th VNTR. Since the diversity index depends upon the number of different alleles and the

number of times that allele appears within a given VNTR, a higher value suggests that a given VNTR may provide greater discrimination among samples.

VNTR data were analyzed using PAUP 4.0 software (Sinauer Associates, Inc., Sunderland, MA), resulting in the unweighted-pair group method using average linkages dendrogram shown in Fig. 3. The fragment sizes generated by each VNTR locus for the strains are provided (see Table SA1 in the supplemental material). These different fragment sizes resulting from the number of tandem repeat units observed within each of the 10 VNTR regions were used to define a genotype.

AFLP analysis of samples. AFLP data were generated for these strains as previously described (7). Briefly, strain DNAs were digested with the restriction endonucleases EcoRI and MseI, and the resulting fragments were ligated to linkers. Dye-labeled primers, homologous to the linkers, were used to selectively amplify the fragments by PCR. The PCR products were then loaded in random fashion in triplicate plates and visualized on an ABI 3100 automated fluorescent sequencer. Forty DNA fragments present in the three runs were used as a "fingerprint" to represent each of the strains in the analysis.

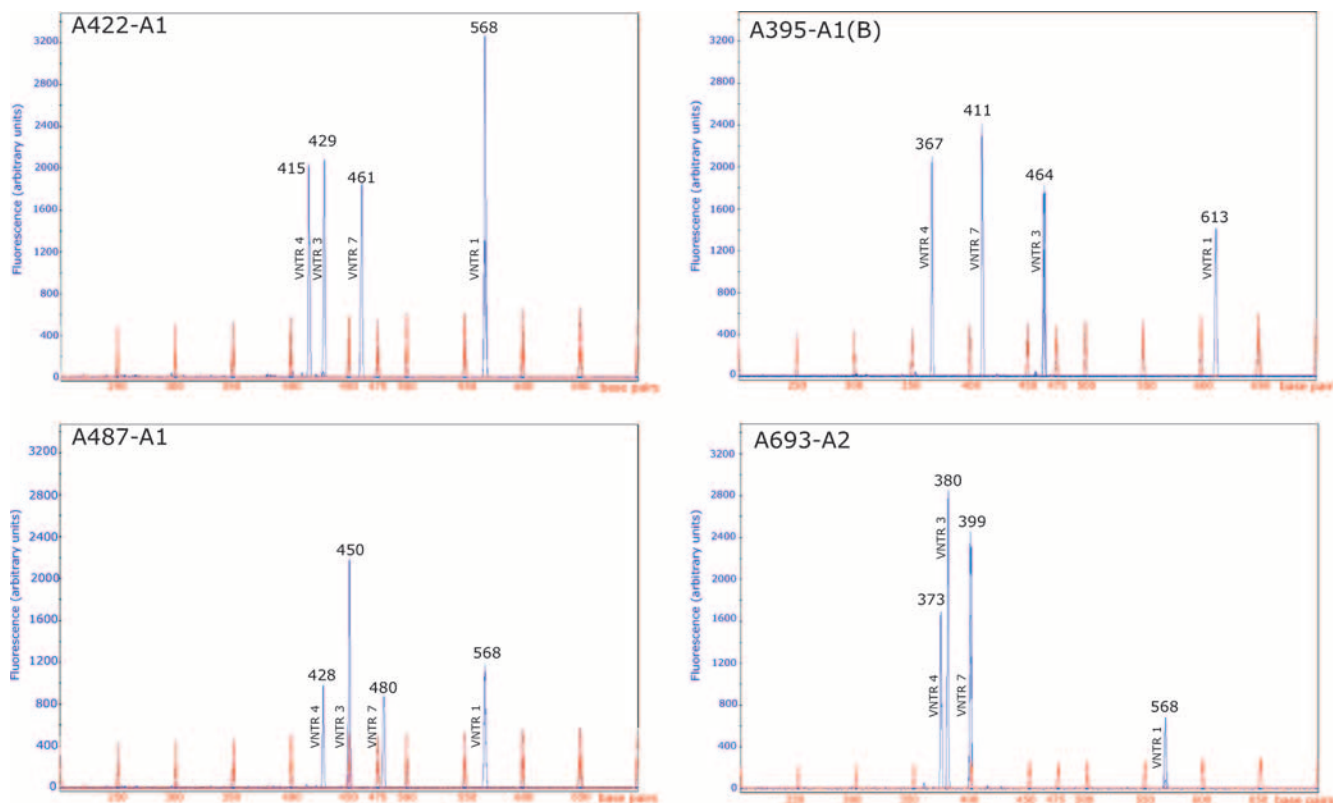The AFLP-based dendrogram in Fig. 1 was generated from the comparison of

FIG. 2. VNTR profiles within four *C. botulinum* strains. Four of the 10 VNTR profiles were overlaid to illustrate the different fragment sizes generated in four strains. VNTR 1, 3, 4, and 7 fragment sizes (in base pairs) are shown in blue for the BoNT/A1 strains A422 and A487, BoNT/A1(B) strain A395, and BoNT/A2 strain A693. The fragment sizes for the internal size standard are shown in red.

strain fingerprints/fragment sizes by use of the Jaccard similarity index. Within the dendrogram, the distance between two samples is the lowest point connecting the samples.

Toxin preparation. In addition to the DNA preparations, the supernatants of cultures containing unpurified toxin were tested. Crude toxin preparations were prepared by growing *C. botulinum* strains (A150, A254, and Ab149) anaerobically in 10 ml TPGY medium (50 g/liter trypticase peptone, 5 g/liter Bacto peptone, 4 g/liter D-glucose, 20 g/liter yeast extract, 1 g/liter cysteine-HCl, pH 7.4) at 35°C for 5 days. The cells were collected by centrifugation at 4,000 rpm and the supernatants filtered through a 0.22-μm filter. These supernatants were assayed by mouse neutralization assay and found to contain ≤10 μg/ml toxin. VNTR experiments using these crude toxin preparations were performed as described above for the DNA preparations, except that 1 μl of a 1:20 dilution of the crude toxin preparation was added as the template instead of purified DNA.

RESULTS

AFLP analyses of the 75 clostridial strains revealed their genetic diversity and differentiated the serotypes into several distinct branches (Fig. 1). The BoNT/A1 strains generally cluster within two branches: 36 are found within the A1 branch and 16 in the A1(B) branch. A384 is the only BoNT/A1 strain that is found within a cluster of bivalent strains distant from the BoNT/A1 and BoNT/A1(B) strains. The AFLP dendrogram also includes a cluster of BoNT/A2 strains, a cluster of four bivalent strains that includes the BoNT/A4 strain (Ba207), and a branch that contains the BoNT/A3 strain (A254). Strains of the other *C. botulinum* serotypes cluster according to their group I to IV designations. The two *C. sporogenes* strains

occupy a branch between the group I and group II *C. botulinum* strains. The AFLP-based dendrogram shows significant genetic diversity among the strains representing the seven serotypes and illustrates the genetic similarity among many of the BoNT/A1 and BoNT/A1(B) strains.

Because AFLP could not distinguish among the many BoNT/A1 strains, VNTR regions were identified and tested. The VNTR regions selected are dispersed throughout the *C. botulinum* ATCC 3502 genome (Table 2) and are distant (greater than 100 kb) from the location of the BoNT/A gene. These loci contained repeat units ranging from 6 to 27 nucleotides in length and provided differing degrees of discrimination within this strain collection (Table 2). An example of the fragment sizes generated for each of the VNTR regions within four strains is shown in Fig. 2. The fragment sizes generated are the result of the different numbers of repeat units within the four VNTR regions and illustrates how the BoNT/A1, BoNT/A1(B), and BoNT/A2 strains can be differentiated from one another.

The 10 VNTR markers amplified targets within all of the serotype A strains, including strains representing the BoNT/A1-A4 subtypes and the bivalent B strains (Ba207, Bf258, Bf698). VNTR analysis using all 10 markers differentiated 61 strains into 38 genotypes (Fig. 3). The in silico VNTR analysis results for three fully sequenced BoNT/A1 strain genomes from the Sanger Institute (ATCC 3502; http://www.sanger.ac.uk/Projects/Microbes/), A150 Hall (GenBank CP000726), and
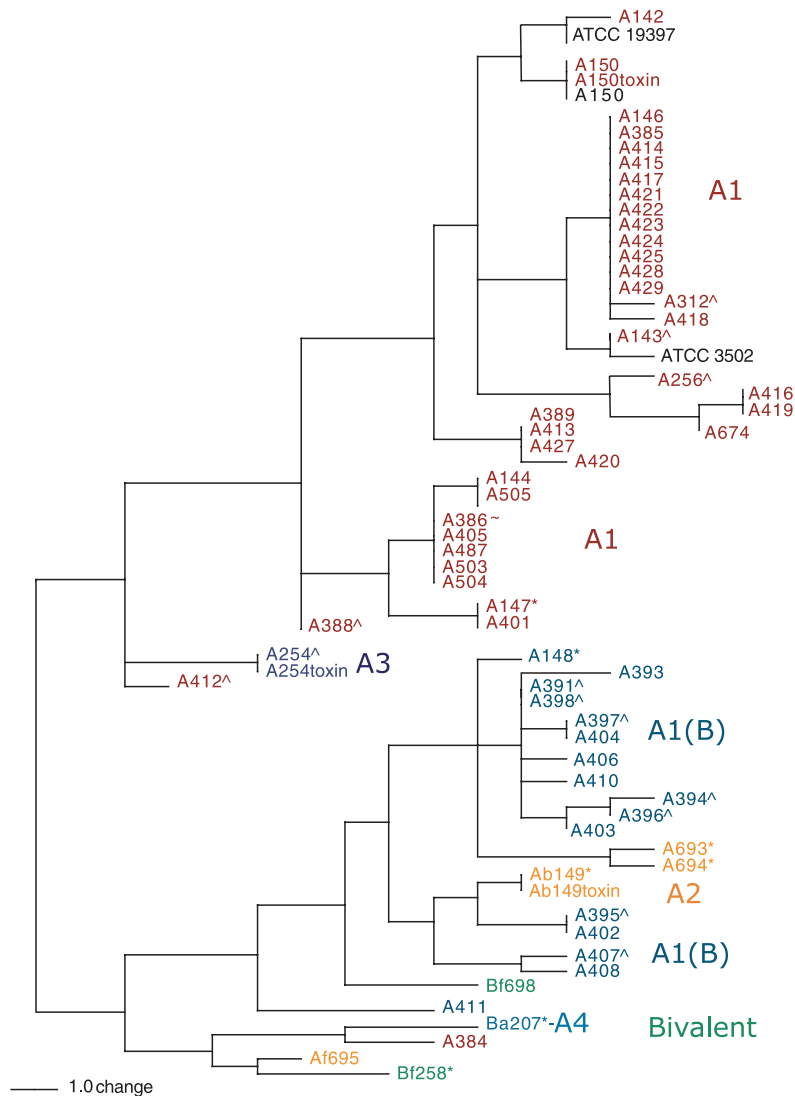
FIG. 3. MLVA of 61 *C. botulinum* strains. Fifty-nine BoNT/A1-A4-producing strains and three bivalent B strains were differentiated into 38 genotypes by use of 10 VNTR loci. The completed genomic sequences of BoNT/A1 strains ATCC 3502, ATCC 19397, and A150 were queried in silico and are shown in black. Three crude toxin preparations from strains A150, Ab149, and A254 were also tested and are designated as "toxin." Symbols: ^, food-borne sources; *, infant cases; ~, soil or marine sediments. The scale indicates the number of alleles that are different between samples.

ATCC 19397 (GenBank CP000727) were included in the dendrogram for comparison.

The VNTR-based dendrogram shows that most of the strains are divided into two groups that generally contain either BoNT/A1 or BoNT/A1(B) and bivalent and BoNT/A2 strains. Although there is little or no detailed information available for many of these strains, sources include infant botulism cases (A147, A148, A149, Ba207, Bf258, A693, A694), food-borne cases (A143, A256, A254, A312, A388, A391, A412, A394 to A398, B152, B512, E213 and F188), and soil or marine sediments (A386, F550). The strains originated from different geographic regions (United States, United Kingdom, France, Sweden) and date from as early as 1922.

The VNTR-based analyses resulted in discrimination among strains beyond the subtype level. The 53 BoNT/A1 and BoNT/ A1(B) strains are separated into 30 distinct genotypes, defined by the number of repeat units observed at each of the 10 VNTR loci. Thirty-six of these strains contain an identical BoNT/A1 gene sequence that would prevent differentiation by antisera or toxin gene analysis. The A144 and A505 strains, which show identical MLVA patterns, are also identical to strain ATCC 17862, known as McClung 457. These two strains were received from different collections and are apparently the same strain. Seventeen of the 53 BoNT/A1 strains are from the Institut Pasteur and are presumed to originate mainly from European sources. Eleven of these cluster with A146 (ATCC 25763, the type strain) and are indistinguishable by this method. The histories of several of these strains indicate that they were obtained from liver samples or from bovine botulism cases. Several were obtained in 1953 and 1954 and may have

been associated with a single outbreak or a series of related outbreaks. The A146 (ATCC 25763) strain used in these experiments was obtained from a different source and clusters with these strains; its history can also be traced to the Institut Pasteur (3).

The 16 BoNT/A1(B) strains were separated by VNTR analysis into 13 genotypes. While the BoNT/A1 strains in the collection are from several continents, all of the BoNT/A1(B) strains of known origin were isolated in the United States. *C. botulinum* A148 differs in origin from other BoNT/A1(B) strains and was collected from an infant case in Pennsylvania, while most other BoNT/A1(B) strains were collected from contaminated foods. Several strains were isolated from food-borne incidents in the relatively small geographic region encompassed by eastern Colorado, New Mexico, and western Nebraska between 1922 and 1935. Two BoNT/A1(B) strains, A396 and A398, represent food-borne botulism cases occurring in the same town 4 years apart, yet they are genetically distinct based on VNTR analysis (Fig. 3). Unfortunately, there is no source information available for the other two genotypes that contain two members each and cannot be differentiated (A397/A404 and A395/A402).

Strain A143, identified as ATCC 3502 in this collection, did not generate the expected VNTR results, which were based on the genomic sequence data of ATCC 3502 (http://www.sanger .ac.uk/Projects/Microbes/). The data differed from the expected results in only one 18-base pair repeat unit of VNTR 9. All of the other nine VNTR regions were identical to the Sanger ATCC 3502 sequence. This A143 strain is very closely related to the Sanger ATCC 3502 strain but is not the exact isolate.

The VNTR markers were also used to distinguish among the bivalent strains. Four of the five bivalent strains are from infant cases: Ba207 (CDC 657, 1988), Bf698 (New Mexico, 1980), Ab149 (Utah, 1977), and Bf258 (Sweden). These strains have similar genetic backgrounds as determined by AFLP analyses (Fig. 1) but contain different combinations of BoNT genes. Each of these strains can be differentiated by sequencing the BoNT/A, BoNT/B, or BoNT/F genes; however, VNTR analysis can rapidly distinguish these strains.

VNTR analysis of 12 *C. botulinum* strains representing serotypes B to G and two *C. sporogenes* strains provided various degrees of resolution. Only 4 or 5 of the 10 VNTR loci were amplified from B512 and B152, respectively. The VNTR primers amplified fewer loci from other serotypes or amplified fragments that were not of the expected size. Seven of the 10 VNTR loci failed to amplify from serotype C, E, and G strains. The lack of amplification within these serotypes does not indicate that these VNTR regions are not present within the tested strains but only that the primers failed to find homology within the sequences. VNTR loci 1, 2, 4, 6, 7, and 10 were amplified within both *C. sporogenes* strains. This result indicates that these VNTR assays are not specific for BoNT-producing clostridia and confirms the need for additional assays to identify the BoNT.

The VNTR markers were also tested to determine whether target sequences could be amplified from crude botulinum toxin preparations. Supernatants from 5-day-old cultures of three *C. botulinum* BoNT/A1 (A150), BoNT/A2 (Ab149), and BoNT/A3 (A254) strains containing ≤10 μg/ml of toxin and unknown amounts of contaminating DNA were tested. The crude toxin preparations were found to support amplification for all 10 VNTR markers and gave identical results to those obtained using purified DNA preparations of the same strain (designated as "toxin" in Fig. 3).

DISCUSSION

Markers that can differentiate among BoNT/A-producing strains can be used by public health agencies to identify and track botulism outbreaks, by forensic laboratories for analysis of toxin samples, and by institutions to further differentiate the strains residing in their collections. Previous analysis of *C. botulinum* strains has focused primarily on neurotoxin serotyping, which is sufficient to prescribe proper medical treatment for a patient. The identification of subtypes within serotypes is possible with the use of specific monoclonal antibodies in enzyme-linked immunosorbent assays or sequencing of toxin genes. MLVA can now provide an additional tool to further distinguish among strains within the BoNT/A1 subtype, providing valuable information for forensics or epidemiological studies.

Within this collection of 59 BoNT/A-producing and 3 bivalent B strains, 10 different VNTR markers were able to differentiate 38 genotypes; 30 genotypes were identified within the 53 strains of the BoNT/A1 subtype. Many genotypes contained only one or two members, also suggesting significant diversity within this strain collection. The only two genotypes that contained more members (5 or 12) contain isolates for which the available historical information is insufficient to provide any understanding of their relationship, if any, to one another. Two isolates with identical VNTR profiles are the same strain obtained from two different collections (A144, A505). The VNTR profile for A146, also known as ATCC 25763, CDC KA38, and 104310T CIP (Institut Pasteur collection), clusters with 11 other isolates and appears to have a French origin, suggesting that this strain was later transported to the United States.

Two isolates expected to be the same based on their origin and labeling were different. ATCC 3502 and A143 differ by a single repeat unit (18 nucleotides) within VNTR 9. The frequency of VNTR mutations observed for 96,000 generations of *Y. pestis* (21) suggests that a significant number of generations separate these two isolates. Therefore, it is likely A143 and the ATCC 3502 strain sequenced by the Sanger Institute share a fairly recent common ancestor and may not represent differences from multiple passages within the environment or the laboratory.

These MLVA profiles suggest a phylogeny similar to that obtained by AFLP analysis. In both methods, the BoNT/A1 and BoNT/A1(B) strains appear to be distinct lineages and located within separate branches of either dendrogram. The BoNT/A3, or Loch Maree, strain is also unique by both methods but clusters within the BoNT/A1 strains by VNTR analysis. According to either method, the bivalent strains generally cluster; AFLP analyses show that four of these five strains cannot be distinguished from one another, and within the VNTR-based dendrogram the five bivalent strains cluster among the BoNT/A1(B) strains. The bivalent Af695 strain contains a BoNT/A2 gene and is located among the BoNT/A1(B), bivalent, and other BoNT/A2 strains. In both analyses, the A384 strain is quite different from the other BoNT/A1 strains. It has

a unique BoNT/A1 gene sequence (differing by only a few nucleotides) but is not a different subtype. Its location in both the VNTR- and AFLP-based dendrograms confirms its unusual genetic background. Another strain, A411, appears to be different from the other BoNT/A1(B) strains based on AFLP and VNTR analyses.

These VNTR loci are not associated with the BoNT gene, but the method shows that the different BoNT/A1-A4 genes are located within very different genetic backgrounds, as supported by the AFLP analyses. The recent finding that the BoNT/A3 and BoNT/A4 genes are located within plasmids suggests a mechanism for their movement into different genetic backgrounds (16).

The VNTR markers designed from the complete genomic sequence of ATCC 3502 were effective in discriminating among the serotype A and bivalent strains in this collection. Other serotypes were problematic because a significant number of VNTR primers did not amplify target sequences. This is not unexpected based on both the AFLP and previous 16S rRNA gene sequence analyses that suggested extensive genetic diversity within this species (2). Six of the 10 VNTR markers successfully amplified products of the expected sizes when DNA from the two *C. sporogenes* strains was used as the template. This species appears to be more closely related to group I *C. botulinum* based on AFLP and 16S rRNA gene analyses (2). This illustrates the complexity of developing assays that unambiguously identify *C. botulinum* isolates without targeting closely related microbes. These VNTR markers successfully amplified targets from trace amounts of contaminating DNA in crude toxin preparations from BoNT/A1-, BoNT/A2-, or BoNT/A3-producing strains. These results indicate that VNTR markers may have considerable value in providing molecular signatures for different toxin preparations that may provide insight into the toxin source. However, it is unknown whether more-purified toxin or samples from the environment, food, human tissues, or stool may yield similar results.

These VNTR markers provide a tool not previously available to interrogate *C. botulinum* samples. MLVA of other bacterial pathogens has been valuable in forensic and epidemiological studies, such as the typing of a *Y. pestis* isolate infecting a human and tracing it to an isolate found within a prairie dog colony nearby (15). It is hoped that these VNTR markers in *C. botulinum* can be used to further epidemiological studies of botulism outbreaks, characterize archival strain collections, and assist in investigations of *C. botulinum* infection and spread in humans and wildlife.

REFERENCES

1. **Benson, G.** 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. **27:**573–580.
2. **Collins, M. D., and A. K. East.** 1998. Phylogeny and taxonomy of the foodborne pathogen *Clostridium botulinum* and its neurotoxins. J. Appl. Microbiol. **84:**5–17.
3. **Fach, P., S. Perelle, F. Dilasser, J. Grout, C. Dargaignaratz, L. Botella, J. Gourreau, F. Carlin, M. R. Popoff, and V. Broussolle.** 2002. Detection by PCR-enzyme-linked immunosorbent assay of *Clostridium botulinum* in fish and environmental samples from a coastal area in northern France. Appl. Environ. Microbiol. **68:**5870–5876.
4. **Farlow, J., K. L. Smith, J. Wong, M. Abrams, M. Lytle, and P. Keim.** 2001. *Francisella tularensis* strain typing using multiple-locus, variable-number tandem repeat analysis. J. Clin. Microbiol. **39:**3186–3192.
5. **Hatheway, C. L.** 1988. Botulism, p. 112–133. *In* A. Balows, W. H. Hausler, Jr., M. Ohashi, and A. Turano (ed), Laboratory diagnosis of infectious diseases: principles and practice. Springer-Verlag, New York, NY.
6. **Hatheway, C. L.** 1990. Toxigenic clostridia. Clin. Microbiol. Rev. **3:**66–98.
7. **Hill, K. K., T. J. Smith, C. H. Helma, L. O. Ticknor, B. T. Foley, R. T. Svensson, J. L. Brown, E. A. Johnson, L. A. Smith, R. T. Okinaka, P. J. Jackson, and J. D. Marks.** 2007. Genetic diversity among botulinum neurotoxin-producing clostridial strains. J. Bacteriol. **189:**818–832.
8. **Hyytia, E., J. Bjorkroth, S. Heilm, and H. Korkeala.** 1999. Characterization of *Clostridium botulinum* groups I and II by randomly amplified polymorphic DNA analysis and repetitive element sequence-based PCR. Int. J. Food Microbiol. **48:**179–189.
9. **Jackson, P. J., M. Hugh-Jones, D. M. Adair, G. Green, K. K. Hill, C. R. Kuske, L. V. Grinberg, F. A. Abramova, and P. Keim.** 1998. PCR analysis of tissue samples from the 1979 Sverdlosk anthrax victims: the presence of multiple *B. anthracis* strains in different victims. Proc. Natl. Acad. Sci. USA **95:**1224–1229.
10. **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 tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. J. Bacteriol. **182:**2928–2936.
11. **Keto-Timonen, R., M. Nevas, and H. Korkeala.** 2005. Efficient DNA fingerprinting of *Clostridium botulinum* types A, B, E, and F by amplified fragment length polymorphism analysis. Appl. Environ. Microbiol. **71:**1148–1154.
12. **Klevytska, A. M., L. B. Price, J. M. Schupp, P. L. Worsham, J. Wong, and P. Keim.** 2001. Identification and characterization of variable-number tandem repeats in the *Yersinia pestis* genome. J. Clin. Microbiol. **39:**3179–3185.
13. **Lin, W. J., and E. Johnson.** 1995. Genome analysis of *Clostridium botulinum* type A by pulsed-field gel electrophoresis. Appl. Environ. Microbiol. **61:**4441–4447.
14. **Lindstrom, M., and H. Korkeala.** 2006. Laboratory diagnostics of botulism. Clin. Microbiol. Rev. **19:**298–314.
15. **Lowell, J. L., D. M. Wagner, B. Atshabar, M. F. Antolin, A. J. Vogler, P. Keim, M. C. Chu, and K. L. Gage.** 2005. Identifying sources of human exposure to plague. J. Clin. Microbiol. **43:**650–656.
16. **Marshall, K. M., M. Bradshaw, S. Pellett, and E. A. Johnson.** 2007. Plasmid encoded neurotoxin genes in Clostridium botulinum serotype A subtypes. Biochem. Biophys. Res. Commun. **361:**49–54.
17. **Nevas, M., M. Lindstrom, S. Hielm, K. J. Bjorkroth, M. W. Peck, and H. Korkeala.** 2005. Diversity of proteolytic *Clostridium botulinum* strains, determined by a pulsed-field gel electrophoresis approach. Appl. Environ. Microbiol. **71:**1311–1317.
18. **Pourshafie, P., P. Vahdani, and M. Popoff.** 2005. Genotyping *Clostridium botulinum* toxin type A isolates from patients using amplified rDNA restriction analysis. J. Med. Microbiol. **54:**933–936.
19. **Skinner, G. E., S. M. Gendel, G. A. Fingerhut, H. A. Solomon, and J. Ulaszek.** 2000. Differentiation between types and strains of *Clostridium botulinum* by ribotyping. J. Food Prot. **63:**1347–1352.
20. **Smith, L. D. S.** 1977. Botulinum, the organism, its toxins, the disease. Charles C Thomas, Springfield, IL.
21. **Vogler, A. J., C. E. Keys, C. Allender, I. Bailey, J. Girard, T. Pearson, K. L. Smith, D. M. Wagner, and P. Keim.** 2007. Mutations, mutation rates, and evolution at the hypervariable VNTR loci of *Yersinia pestis*. Mutat. Res. **616:**145–158.
22. **Weir, B. S.** 1996. Genetic data analysis II. Sinauer Associates, Inc., Sunderland, MA.
23. **Whatmore, A. M., S. J. Shankster, L. L. Perrett, T. J. Murphy, S. D. Brew, R. E. Thirlwall, S. J. Cutler, and A. P. MacMillan.** 2006. Identification and characterization of variable-number tandem-repeat markers for typing of *Brucella* spp. J. Clin. Microbiol. **44:**1982–1993.