Defence Research and Development Canada
Recherche et développement pour la défense Canada

Molecular Typing of *Bacillus anthracis* isolates using MLVA-15

Chad W. Stratilo and Doug Bader
DRDC Suffield

Defence R&D Canada – Suffield

Defence R&D Canada

Technical Memorandum

DRDC Suffield TM 2013-124

December 2013

Canada

Molecular Typing of *Bacillus anthracis* isolates using MLVA-15

Chad W. Stratilo and Doug Bader
DRDC Suffield

Defence R&D Canada – Suffield

Technical Memorandum
DRDC Suffield TM 2013-124
December 2013

Principal Author

Chad W. Stratilo
Microbiologist

Approved by

Dr. D. Pedersen
Head, Bio Threat Defence Section

Approved for release by

Dr. R. Clewley
Chairman, Document Review Panel

© Her Majesty the Queen in Right of Canada, as represented by the Minister of National Defence, 2013

© Sa Majesté la Reine (en droit du Canada), telle que représentée par le ministre de la Défense nationale, 2013

Abstract

Bacillus anthracis, the etiological agent of anthrax, lacks character homoplasy and is genetically very monomorphic. Variable number tandem repeats (VNTR) are rapidly evolving stretches of DNA that can serve as markers for strain differentiation (molecular typing). A collection of *B. anthracis* strains were previously characterized using multiple-locus VNTR analysis (MLVA-8). This work describes the characterization of 31 strains by MLVA-15. MLVA-15 allowed these strains to be placed within a phylogenetic framework as with MLVA-8. MLVA-15 did allow higher resolution and discrimination than MLVA-8. While MLVA-15 allowed many strains to be discriminated from each other, several Canadian isolates were indistinguishable. This work allowed verification of strain designation within the DRDC Suffield collection and allowed for the development of a high throughput genetic typing capability at DRDC Suffield.

Résumé

Bacillus anthracis, l'agent étiologique de l'anthrax, n'affiche pas d'homoplasie de caractère et est génétiquement très monomorphe. Les nombres variables de répétitions en tandem (VNTR) sont des séquences d'ADN en évolution rapide qui peuvent servir de marqueurs pour la différentiation des souches (typage moléculaire). Une collection de souches de B. anthracis a été caractérisée dans le passé par analyse multi-locus VNTR (MLVA-8). Les présents travaux décrivent la caractérisation de 31 souches par MLVA 15. Par l'analyse MLVA-15, on a pu placer ces souches dans un cadre phylogénique comme on l'avait fait lors de l'analyse MLVA 8, mais à une résolution et discrimination plus élevées que celles obtenues par MLVA-8. Toutefois, bien que l'analyse MLVA 15 ait permis de différencier de nombreuses souches les unes des autres, plusieurs isolats canadiens étaient indistinguables. Ces travaux ont permis de vérifier la désignation des souches figurant dans la collection de RDDC Suffield et de développer une haute capacité de production de typage génétique à RDDC Suffield.

Executive summary

Molecular Typing of *Bacillus anthracis* Isolates using MLVA-15

Chad W. Stratilo; Doug Bader; DRDC Suffield TM 2013-124; Defence R&D Canada – Suffield; December 2013.

Introduction or background: There is a need to characterize biothreat agents for biodefence and biosecurity purposes. Genetic typing techniques can be used to characterize biothreats and have the potential to determine the possible foci or origin of the agent. *B. anthracis*, a potential biothreat agent, appears to be among the most genetically and phenotypically homogenous pathogenic bacteria known. A CRTI project (CRTI 02-0069RD) was undertaken to establish molecular typing capabilities of biothreat agents using variable number tandem repeats (VNTR). DNA from a number of *B. anthracis* strains (31 strains) was analyzed using multiple-locus VNTR analysis (MLVA) at 15 loci.

Results: MLVA-15 analysis of DNA from 31 *B. anthracis* strains revealed a phylogenetic relationship between strains in which many strains were distinguished from each other, however, Canadian *B. anthracis* isolates displayed highly conserved MLVA-15 types.

Significance: DRDC Suffield has developed a molecular typing capability for *B. anthracis*, which has allowed for a phylogenetic framework that could be used to assess additional strains including those from a bioterrorism event. Typing technology and the methods developed in this project could be important tools for attempting to trace the possible foci or source of an outbreak or attack.

Future plans: MLVA typing schemes for other biothreat agents will be pursued. MLVA databases will be expanded as additional strains are acquired in order to increase the resolving power of this technique.

Sommaire

Typage moléculaire d'isolats de *Bacillus anthracis* par analyse MLVA 15

Chad W. Stratilo; Doug Bader; RDDC Suffield TM 2013-124; R et D pour la Défense Canada – Suffield; septembre 2013.

Introduction ou contexte : Il est nécessaire de caractériser les agents de bioterrorisme aux fins de défense biochimique et de sécurité biochimique. Les techniques de typage génétique peuvent être utilisées pour caractériser les agents de bioterrorisme et potentiellement déterminer les points de nécrose miliaires, ou origine, de l'agent. B. anthracis, un agent de bioterrorisme potentiel, semble figurer parmi les bactéries pathogènes connues les plus homogènes sur le plan génétique et phénotypique. Un projet de l'IRTC (no 02-0069RD) a été entrepris pour établir les capacités de typage moléculaire d'agents de bioterrorisme au moyen du nombre variable de répétitions en tandem (VNTR). L'ADN d'un certain nombre de souches de B. anthracis strains (31 souches) a été analysé au moyen d'une analyse multilocus VNTR (MLVA) à 15 loci.

Résultats : L'analyse MLVA-15 de l'ADN de 31 souches de B. anthracis a révélé une relation phylogénétique entre les souches, dans laquelle de nombreuses souches ont été différenciées l'une de l'autre. Toutefois, les isolats canadiens de B. anthracis affichaient des types MLVA-15 hautement conservés.

Importance : RDDC Suffield a élaboré une capacité de typage moléculaire pour B. anthracis, qui a permis d'établir un cadre phylogénétique qui pourrait servir à évaluer d'autres souches, y compris celles provenant d'un événement de bioterrorisme. La technologie de typage et les méthodes développées lors de ce projet pourraient devenir des outils importants pour découvrir le point de nécrose miliaire ou la source possible d'une éclosion ou d'une attaque.

Perspectives : On continuera de rechercher des mécanismes de typage par MLVA pour d'autres agents de bioterrorisme. Les bases de données des MLVA seront élargies au fur et à mesure que des souches additionnelles seront acquises, afin d'augmenter la résolution.

Table of contents

Abstract …….. i
Résumé ….... i
Executive summary ii
Sommaire iii
Table of contents iv
List of figures v
List of tables v
Acknowledgements v
Introduction 1
Materials and Methods 3
    Bacterial Strains and DNA Isolation 3
    MLVA-15 3
Results 7
Discussion 11
References 15
Annex A - PCR Setup for MLVA-15 Multiplex Reactions 17
List of symbols/abbreviations/acronyms/initialisms 19

List of figures

Figure 1. MLVA-15 based dendogram of twenty-nine B. anthracis isolates. The 15 MLVA marker loci were used to calculate clustering analysis using a Categorical coefficient and UPGMA options. The genetic distance is presented as percent similarity. 13

List of tables

Table 1. *B. anthracis* Strains 4

Table 2. MLVA-15 Primers 5

Table 3. Characterization of *B. anthracis* VNTR used for MLVA 8

Table 4. Allele calls for *B. anthracis* MLVA analysis 9

Table 5. Amplicon size (bp) for *B. anthracis* MLVA Analysis 12

Acknowledgements

This project would not have been possible without the support and collaboration of M. Mulvey, L. Bryden and D. Johnstone at the Public Health Agency of Canada National Microbiology Laboratory. This project was supported by CRTI 02-0069RD. Thanks to L. Bryden for marker characterization data.

This page left intentionally blank

Introduction

Bacillus anthracis is a spore-forming bacterium endemic in soils throughout much of the world. Anthrax, caused by the *B. anthracis* bacterium, is a zoonotic disease affecting both animals and humans. Virulent strains of *B. anthracis* contain two virulence plasmids, pXO1 and pXO2. These plasmids contain genes that confer toxin producing ability and capsule synthesis activity, respectively, although there are chromosomally encoded factors that are crucial for the full virulence of *B. anthracis* [1]. Due to bioterrorism events involving anthrax, there has been an increased interest in *B. anthracis*, especially with respect to identification, detection, and molecular subtyping of this pathogen.

B. anthracis belongs to the *B. cereus* group, whose members include *B. cereus* and *B. thuringiensis*. As a group there is very little genetic diversity. Amplified fragment length polymorphism (AFLP) analysis of the *B. cereus* group revealed that *B. cereus* and *B. thuringiensis* are broadly interspersed across all branches of the AFLP phylogenetic tree, while *B. anthracis* shows very low genetic diversity and clusters to a sub-branch of the phylogenetic tree that is distinct from other members of the *B. cereus* group [2].

B. anthracis genetic homogencity makes reconstructing its evolutionary history problematic. *B. anthracis* is among the most homogeneous pathogenic bacteria described, lacking character homoplasy and containing few single nucleotide polymorphisms (SNPs) [3]. This lack of homoplasy may be due to its life history which includes spending long periods of time as dormant endospores. Most molecular techniques used to differentiate between strains in other species such as AFLP [2, 4], multi-locus sequence typing (MLST) [5], and pulsed field gel electrophoresis (PFGE) [6], fail to discriminate *B. anthracis* strains.

Several molecular typing methods, including whole genome SNP analysis and multi-locus variable-number tandem repeat analysis (MLVA), have been more successful in discriminating *B. anthracis* strains and have allowed the exploration of its phylogenetics [7].

A MLVA method exploiting the copy number differences of nucleotide repeat sequences at six chromosomal loci and one locus for each of the two plasmids has been developed [8, 9, 10]. MLVA analysis is useful due to increased mutation rates compared to SNPs, as well as a greater number of allelic states compared to biallelic SNPS [7]. MLVA-8 was used to analyze DNA from *B. anthracis* strains in the DRDC Suffield and Public Health Agency of Canada National Microbiology Laboratory (PHAC-NML) collections [11]. A more discriminatory MLVA typing scheme using an additional 7 loci has been developed [8]. This new technique (MLVA-15) was evaluated in this study for its ability to further discriminate *B. anthracis* strains previously tested using the MLVA-8 technique.

This page left intentionally blank

Materials and Methods

Bacterial Strains and DNA Isolation

Table 1 lists the 31 *B. anthracis* strains that were typed using MLVA-15. DNA from strains in the DRDC Suffield collection was prepared in BSL3 biocontainment as follows: *B. anthracis* strains were grown on sheep blood agar Petri dishes at 37 ˚C overnight. Strain DNA was isolated using MasterPure™ DNA & RNA Purification Kit (Epicentre Biotechnologies, Madison, WI, USA), Phase lock (Eppendorf, Westbury, NY, USA) or GNOME® DNA Isolation Kit (Q.BIOgene, Irvine, CA, USA). DNA preparations were checked for sterility by broth and plate culture prior to removal for MLVA analysis which was performed outside BSL3 biocontainment. DNA from strains in the PHAC-NML collection was provided by PHAC-NML.

MLVA-15

MLVA-15 of *B. anthracis* strain DNA was performed for loci as described by Keim *et al*. [8]. PCR reactions contained 1× AmpliTaq Gold PCR Buffer and 0.5 U of AmpliTaq Gold DNA polymerase (Applied Biosystems Inc., Foster City, CA, USA); 4 mM $MgCl_2$; dNTPs (0.2 mM each) and 0.2 μM of forward and reverse primers (Table 2). Approximately 2-20 ng of template DNA was used per 50 μL reaction. Phosphoramidite fluorescent dye (6-FAM, HEX, NED) covalently linked to the forward primer was used to detect the amplicons. Amplicon sequencing required the use of unlabeled forward and reverse primers. Thermocycling conditions were 95˚ C for 5 min followed by 35 cycles of 94 ˚C for 30s; 60 ˚C for 30s; 65˚ C for 30s and 1 cycle of 65 ˚C for 7 min. Multiplex reactions were carried out for fragment analysis. Detailed PCR setup is described in Annex A. HiDi formamide (8 μL) and 1 μL of the diluted PCR products were combined with 1 μL of size standard Rhodamine-X Mapmaker 70 to 400 bp and CST ROX 420-800 (BioVentures Inc., Murfreesboro, TN, USA). All 15 loci were amplified in three individual three individual PCR reactions. These products were analyzed on an ABI 3100 genetic analyzer and sized using GeneMapper (Applied Biosystems).

Some of the MLVA amplicons were sequenced to establish the size of the amplicon and the VNTR. These MLVA PCR products were purified using Montage PCR_{96} plates (Millipore, Nepean, ON, CA) and sequenced using the MLVA primers as sequencing primers. All sequencing reactions were carried out in 20 μL reactions with Big Dye 3.1 Terminator chemistry (Applied Biosystems) and analyzed on an ABI 3100 automated sequencer (Applied Biosystems). Contig assembly of the *B. anthracis* MLVA loci sequences was performed with Sequencher (Gene Codes Corp., AnnArbor, MI, USA). Clustering analysis of the MLVA was done using a Categorical coefficient and UPGMA options using Bionumerics software (Applied Maths, Inc.,Austin, TX, USA).

Table 1. B. anthracis Strains

Strain	DNA Source	Source / yr. of Isolation
9604	DRDC Suffield	AB, Canada, 1996
9609	DRDC Suffield	AB, Canada, 1996
9610	DRDC Suffield	AB, Canada, 1996
9614	DRDC Suffield	AB, Canada, 1996
94188C	DRDC Suffield	SK, Canada, 1994
93212C	DRDC Suffield	Canada, 1993
9619	DRDC Suffield	Canada, 1996
93-189C	DRDC Suffield	NWT, Canada, 1993
ATCC-4229	DRDC Suffield	Pasteur vaccine
RP42	DRDC Suffield	RP42 Δlef/Δcya
Vollum	DRDC Suffield	South Africa, 1948
Vollum 1B	DRDC Suffield	Derived from Vollum
SK61	DRDC Suffield	
ACB	DRDC Suffield	Ohio, USA, 1952
Ames	DRDC Suffield	Texas, USA, 1980
NH	DRDC Suffield	New Hampshire, USA, 1957
VH	DRDC Suffield	South Africa, 1952
Buffalo	DRDC Suffield	Iowa, USA, 1979
SK31	DRDC Suffield	South Africa, 1974
SK162	DRDC Suffield	Florida, USA, 1976
17T5	DRDC Suffield	South Afric, 1975
9807	NML-PHAC	MB, Canada, 1998
9937	NML-PHAC	AB, Canada, 1999
NML 03-0139	NML-PHAC	MB, Canada, 2003
NML 03-0191	NML-PHAC	MB, Canada, 2003
9911	NML-PHAC	AB, Canada, 1999
Sterne	NML-PHAC	Vaccine Strain
ATCC-6602	NML-PHAC	Pasteur vaccine
2000-77	NML-PHAC	MB, Canada, 2000
9946	NML-PHAC	AB, Canada, 1999
MM9497	NML-PHAC	Unknown

Table 2. MLVA-15 Primers

Marker[a]	Primers	Primer Sequences (5` | 3`)	Allele Size Range[a]
vrrA	vrrA-F1	CACAACTACCACCGATGGCACA	290-326
	vrrA-R1	GCGCGTTTCGTTTGATTCATAC	
vrrB1	vrrB1-F1	ATAGGTGGTTTTCCGCAAGTTATTC	224-251
	vrrB1-R1	GATGAGTTTGATAAAGAATAGCCTGTG	
vrrB2	vrrB2-F1	CACAGGCTATTCTTTATCAAACTCATC	135-171
	vrrB2-R1	CCCAAGGTGAAGATTGTTGTTGA	
vrrC1[b]	HCvrrC1F	CAGAAGAATCAGTGATTGTC	405-522
	HCvrrC1R	CACCATTTCTTGTTGCATTTCCTC	
vrrC2[b]	HCvrrC2F	TGGCTGATGAACAAACGAA	444-516
	HCvrrC2R	GCATGTCTTTCCATTAATCGCGCT	
CG3	CG3-F	TGTCGTTTTACTTCTCTCTCCAATAC	153-158
	GC3-R	AGTCATTGTTCTGTATAAAGGGCAT	
pXO1-aat	pXO1-AAT-F3	CAATTTATTAACGATCAGATTAAGTTCA	123-144
	pXO1-AAT-R3	TCTAGAATTAGTTGCTTCATAATGGC	
pXO2-at	pXO2-AT-F1	TCATCCTCTTTTAAGTCTTGGGT	135-145
	pXO2-AT-R1	GTGTGATGAACTCCGACGACA	
BaVNTR12	BaVNTR12-F	CGTACGAAGTAGAAGTCATTAA	112-116
	BaVNTR12-R	GCATATAATTGCACCTCATCTAG	
BaVNTR16	BaVNTR16-F	CTCTTGAAAATATAAAACGCA	273-305
	BaVNTR-R	GAATAATAAGGGTTCTCATGGTAT	
BaVNTR17	BaVNTR17-F	TAGGTAAACAAATTTTCGTAATC	378-452
	BaVNTR17-R	GATCGTACAACAGCAATTATCAT	
BaVNTR19-2	BaVNTR19-2-F	GTGATGAAATCGGACAAGTTAGGAG	91-94
	BaVNTR19-2-R	GAAATATTTTATTAAACATGCTTTCCATCC	
BaVNTR23	BaVNTR23-F	TTTAGAAACGTTATCACGCTTA	170-206
	BaVNTR23-R	GTAATACGTATGGTTCATTCCC	
BaVNTR32	BaVNTR32-F	AACTGGATCCAGGAGATTATA	409-577
	BaVNTR32-R	GAAACAAGAGCAAACCCAAT	
BaVNTR35	BaVNTR35-F	AAATAATATGTTCCTTTTGCTG	102-114
	BaVNTR35-R	GTCCTGAAATAAATGCTGAAT	

[a] Based on allele name not allele size (bp) therefore may differ from previous allele naming conventions. For actual size range refer to Table 4 "Location"

[b] Primers for markers vrrC1 and vrrC2 were modified to accommodate a sizing algorithm of less than 550 bp and designated with HC-prefix

This page left intentionally blank

Results

The nature and extent of known variability of the various VNTR loci are described in Table 3. Direct sequencing of the 15 loci allowed for sizing of the alleles present at those loci. Allele names are based on sizing established by direct sequencing and the integrating of sizings established in the literature (Table 4) [8, 9, 10]. The sizes of the amplicons during fragment analysis are also listed (Table 4). The differences between amplicon length established by fluorescent primer fragment length analysis and direct sequencing can be ascertained by comparison between the allele name established by direct sequencing, and the location of the allele established by fragment length analysis (Table 4). The difference between length established by sequencing and length determined by fragment analysis is +/- 4 bp although there may be no difference at a particular locus or allele. This discrepancy may be attributable to the type of phosphoramidite fluorescent dye used in 5’ labelling for fragment analysis. The name of the allele for a given locus is based on the size of the VNTR amplicon as determined by sequence analysis as opposed to fragment size analysis. The allele names for several loci have been revised from previous work to more closely integrate with allele names used in the literature.

The results of the MLVA-15 analysis are listed in Table 5. The phylogenetic relationship of the closest related 29 strains analyzed by MLVA-15 are shown in Figure 1. The vaccine strain (Sterne) is missing locus pXO2 since it lacks the pXO2 plasmid. Strain 4229 lacks the pXO1 plasmid hence it lacks an allele for the pXO1 locus. Strain RP42 also lacks the pXO2 plasmid. Strain 6602 lacks both plasmids and does not have alleles for pXO1 or pXO2 loci.

The primers for vrrC1 and vrrC2 were redesigned in this study (HCvrrC1 and HCvrrC2) to produce smaller amplicons than those used by others [8, 9, 10]. Amplicons that are smaller than 600 bp in length allowed for a standard sizing algorithm to be used for analysis by the ABI Genemapper software. HCvrrC1 and HCvrrC2 were used for fragment analysis but the traditional primers for those loci were used for sequencing.

Relationships between classical *B. anthracis* strains using cluster analysis were found to be as expected based on previous reports [8, 11, 12]. While MLVA-15 resolved additional isolates compared to MLVA-8, a large number of isolates from North Western Canada were still not resolved using this technique.

Table 3. Characterization of B. anthracis VNTR used for MLVA

Marker Locus[a]	**Consensus Repeat Motif**	**Repeat Size**	**Array Size**	**# Of Alleles**
vrrA1	CAATATCAACAA	12	2 - 6	5
vrrB1[b]	CA(T/C)CA(T/C/A)(G/C)(G/A)(C/A/T/G)	9	11 - 23	5
vrrB2[b]	CATCAAGGT or CATCACCAC	9	2 - 4	4
vrrC1	AA(C/T)TGC(A/G)GAA(A/G)CAGAAGAGT(T/C)AGAAGAAGT(G/A)GAAGT	36 (9)[c]	4 - 12	6
vrrC2	G(C/T)AGA(A/G)GAA(G/C)(A/C)A(C/T)C(G/A)(G/A)(T/C)(T/C)	18	16 - 19	3
CG3	ATATT	5	1 - 2	2
pXO1-aat	TAA	3	4 - 13	8
pXO2-at	AT	2	6 - 15	9
BAVNTR12	AT	2	5 - 7	3
BAVNTR16[d]	TTTACACA	8	2 - 6	4
BAVNTR17	AATGTGTA	8	2 - 4	5
BAVNTR19-2	ATA	3	5 - 6	2
BAVNTR23	TTTCTTTTTCAT	12	1 - 5	4
BAVNTR32	AGG(C/T)GAAGAAACAGAAAAGCC	21	2 - 8	4
BAVNTR35	CAATCA	6	2 - 4	3

a. VNTR markers found in ORFs are shown in italics.
b. *vrrB* repeats are not all identical, as some contain multiple nucleotide differences.
c. V*rrC1* marker contains a degenerate 9-nucleotide sub-repeat structure that results in fractional repeat sizes in some alleles (the most frequent of the repeats are given in the above table)
d. BaVNTR16 contains two repeat areas

Table 4. Allele calls for B. anthracis MLVA analysis

Marker Name	Bin Parameters	Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6	Allele 7
vrrA1	Sequence (bp)	290	302	314	326			
	Location	285.8	297.4	309.7	321.8			
	Offset (L/R)	0.5/0.5	0.5/0.5	0.75/0.75	0.5/0.5			
	Range	0.42	0.12	1.02	0.51			
vrrB1	Sequence (bp)	224	251					
	Location	225.64	252.21					
	Offset (L/R)	0.5/0.5	0.5/0.5					
	Range	0.31	0.13					
vrrB2	Sequence (bp)	135	153	162	171			
	Location	128.98	147.42	156.65	165.54			
	Offset (L/R)	0.5/0.5	0.5/0.5	0.75/0.75	0.5/0.5			
	Range	0	0.14	1.17	0.07			
HCvrrC1	Sequence (bp)	405	450	486	522			
	Location	398.51	443.05	480.74	516.44			
	Offset (L/R)	0.5/0.5	0.5/0.5	0.75/0.75	0.5/0.5			
	Range	0.07	0.72	0.67	0.05			
HCvrrC2	Sequence (bp)	444	516					
	Location	436.83	510.28					
	Offset (L/R)	0.5/0.5	0.5/0.5					
	Range	0.14	0.31					
CG3	Sequence (bp)	153	158					
	Location	149.41	154.31					
	Offset (L/R)	1/1	0.75/0.75					
	Range	1.18	0.68					
pX01-AAT	Sequence (bp)	123	126	129	132	135	138	144
	Location	118.79	121.67	124.72	127	130.74	133.64	139.32
	Offset (L/R)	1/1	1/1	1/1	1/1	1/1	1/1	1/1
	Range	0.88	1.11	1.22	0.0	1.2	0.53	0.0
pX02-AT	Sequence (bp)	135	137	139	141	143	145	
	Location	132.05	133.78	135.63	137.7	139.78	141.7	
	Offset (L/R)	0.5/0.5	0.5/0.5	0.5/0.5	0.5/0.5	0.5/0.5	0.5/0.5	
	Range	0.04	0.37	0.45	0.25	0.16	0	
BAVNTR12	Sequence (bp)	112	114	116				
	Location	108.44	110.47	112.55				
	Offset (L/R)	0.5/0.5	0.5/0.5	0.5/0.5				
	Range	0.07	0.18	0				
BAVNTR16	Sequence (bp)	273	281	305				
	Location	268.11	276.15	300.01				
	Offset (L/R)	0.5/0.5	0.5/0.5	0.5/0.5				
	Range	0.23	0.15	0.15				
BAVNTR17	Sequence (bp)	378	386	394	452			
	Location	376.59	384.72	392.88	451.62			
	Offset (L/R)	0.5/0.5	0.5/0.5	0.5/0.5	0.5/0.5			
	Range	0.15	0.2	0.13	0.11			
BAVNTR19-2	Sequence (bp)	91	94					
	Location	90.83	93.39					
	Offset (L/R)	0.5/0.5	1/1					
	Range	0.07	0.13					
BAVNTR23	Sequence (bp)	170	182	194	206			
	Location	169.55	181.59	193.56	205.49			
	Offset (L/R)	0.5/0.5	0.75/0.75	0.75/0.75	0.75/0.75			
	Range	0.05	0.03	0.1	0			
BAVNTR32	Sequence (bp)	409	472	514	577			
	Location	405.68	469.46	511.35	574.09			
	Offset (L/R)	0.5/0.5	0.75/0.75	0.75/0.75	0.75/0.75			
	Range	0	0.14	0.27	0.27			
BAVNTR35	Sequence (bp)	102	108	114				
	Location	99.15	105.66	112.12				
	Offset (L/R)	0.5/0.5	0.5/0.5	0.5/0.5				
	Range	0.03	0.23	0.02				

This page left intentionally blank

Discussion

The MLVA-15 typing method presented in this report is a robust and well accepted approach for characterizing *B. anthracis* isolates. The use of fragment analysis allows for quick and reproducible results. Analysis of strains between laboratories world wide is convenient and easily converted into simple numerical data unlike AFLP or PFGE. VNTR analysis is required for *B. anthracis* typing due to its extremely low mutation rates, likely due to its life history.

Increasing the number of MLVA loci from eight to fifteen increased the resolving power of this typing method. As an example, MLVA-8 analysis of strain SK162 and strain Buffalo showed identical MLVA-8 types [11]. Increasing the number of MLVA loci to 15 allowed these two strains to have distinct MLVA-15 types. These differences were due to BaVNTR16 and BaVNTR17. It is interesting to note that both of these markers are on the pXO2 plasmid. Unfortunately MLVA-15 did not improve the resolution of most of the Canadian isolates from each other although other studies have found some heterogeneity within their library of Canadian isolates. The vast majority of Canadian isolates still however fall into the Western North American genotype as described by Van Ert et al. [10]. It should be noted these isolates may in fact be identical even coming from the same outbreak. The provenance of some strains / isolates is very limited. Other typing methods such as SNR, SNP or whole genome sequencing might differentiate these isolates to a greater degree [9, 13]. With current technology whole genome sequencing would be used to analyze an individual unknown sample or a few unknown samples, although MLVA and SNP typing may be used to screen a large collection of subsequent samples to ensure they are from the same source. Over 5,730 environmental samples from 60 site locations were collected during the investigation of the 2001 anthrax attacks in the United States [14]. Even with current technology this number of samples would be daunting to analyze via whole genome sequencing.

Discrepancies between the allele calls used in this paper and those found in other publications [8] are due to designations adopted based on fragment sizing analysis as opposed to sequence data. For example, in Table 4, *vvr*B2 allele name is 135 whereas its sequence location is 128.98 ± 0.5 bp. It should be noted the choice of allele name also takes into account the repeat size. It may be beneficial in the future to move away from allele names that are based on amplicon or sequence size but rather adopt a scheme that uses a letter or number designation based on the repeat structure and length [10].

B. anthracis has been used as a biological threat agent and could be used against the CF as a biowarfare agent or as a bioterrorist threat against Canadians. This project has allowed DRDC Suffield to establish the capability to perform high resolution subtyping of *B. anthracis*. The methods and cluster database that was developed in this project can be used to type and characterize genetic relatedness of additional *B. anthracis* strains in the future. High resolution subtyping could be used for forensic investigations where intentional use is suspected. Our present understanding of Canadian *B. anthracis* isolates is that they are highly monomorphic represented by two clades within the clonal lineage of *B. anthracis* [10]. This may assist in including or excluding Canada as the place of origin of an unknown isolate. Although it should be noted the true natural diversity of *B. anthracis* in Canada may not accurately be reflected by the current isolate collections tested. The larger and more comprehensive an isolate collection is, the better the understanding of the true genetic diversity of *B. anthracis* in natural populations in Canada and worldwide would be. This would better enable the scientific community to distinguish a natural outbreak from an intentional use situation.

Table 5. Amplicon size (bp) for B. anthracis MLVA Analysis

Strain	vrrA	vrrB1	vrrB2	HCvrrC1	HCvrrC2	HCCG3	pXO1	pXO2	BAVNTR 12	BAVNTR 16	BAVNTR 17	BAVNTR 19	BAVNTR 23	BAVNTR 32	BAVNTR 35
ATCC4229	314	224	162	486	516	153	NEG	137	114	273	386	94	194	514	108
ATCC6602	326	224	135	486	516	158	NEG	NEG	114	NEG	NEG	94	206	472	108
9604	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
9609	314	224	162	486	516	153	129	135	114	281	386	94	194	514	108
9610	314	224	162	450	444	158	129	141	114	281	386	94	194	514	108
9614	314	224	162	486	516	153	129	135	114	281	386	94	194	514	108
9619	314	224	162	486	516	153	126	135	114	281	386	94	194	514	108
9807	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
9911	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
9937	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
9946	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
200077	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
NML030139	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
NM 030191	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
17T5	302	251	171	450	444	158	123	143	112	305	378	94	182	472	102
93-189C	314	224	162	486	516	153	129	135	114	281	386	94	194	514	108
93212C	314	224	162	486	516	158	135	135	114	281	386	94	194	514	108
94188c	314	224	162	486	516	153	129	137	114	281	386	94	194	514	108
ACB	302	224	153	486	444	158	129	141	114	273	378	91	194	577	108
AMES	314	224	153	450	444	158	126	141	114	273	386	91	194	577	114
Buffalo	314	224	162	486	516	153	129	139	114	281	386	94	194	514	108
MM59497	314	224	162	486	516	153	129	139	114	281	386	94	194	514	108
NH	314	224	162	405	516	158	138	139	114	273	452	94	194	577	108
Rp42	314	224	162	450	444	158	132	NEG	114	NEG	NEG	91	182	577	114
SK162	314	224	162	486	516	153	129	139	114	273	394	94	194	514	108
SK31	302	251	171	450	444	158	126	143	112	305	378	94	182	472	102
SK61	314	224	162	486	444	158	135	145	114	273	378	91	194	409	108
Sterne	314	224	162	450	444	158	132	NEG	114	NEG	NEG	91	194	577	114
VH	326	224	162	405	516	158	144	139	116	273	386	94	194	577	108
Vollum	290	224	153	405	516	158	135	139	114	273	452	94	170	514	108
Vollum1B	290	224	153	405	516	158	135	139	114	273	452	94	170	514	108

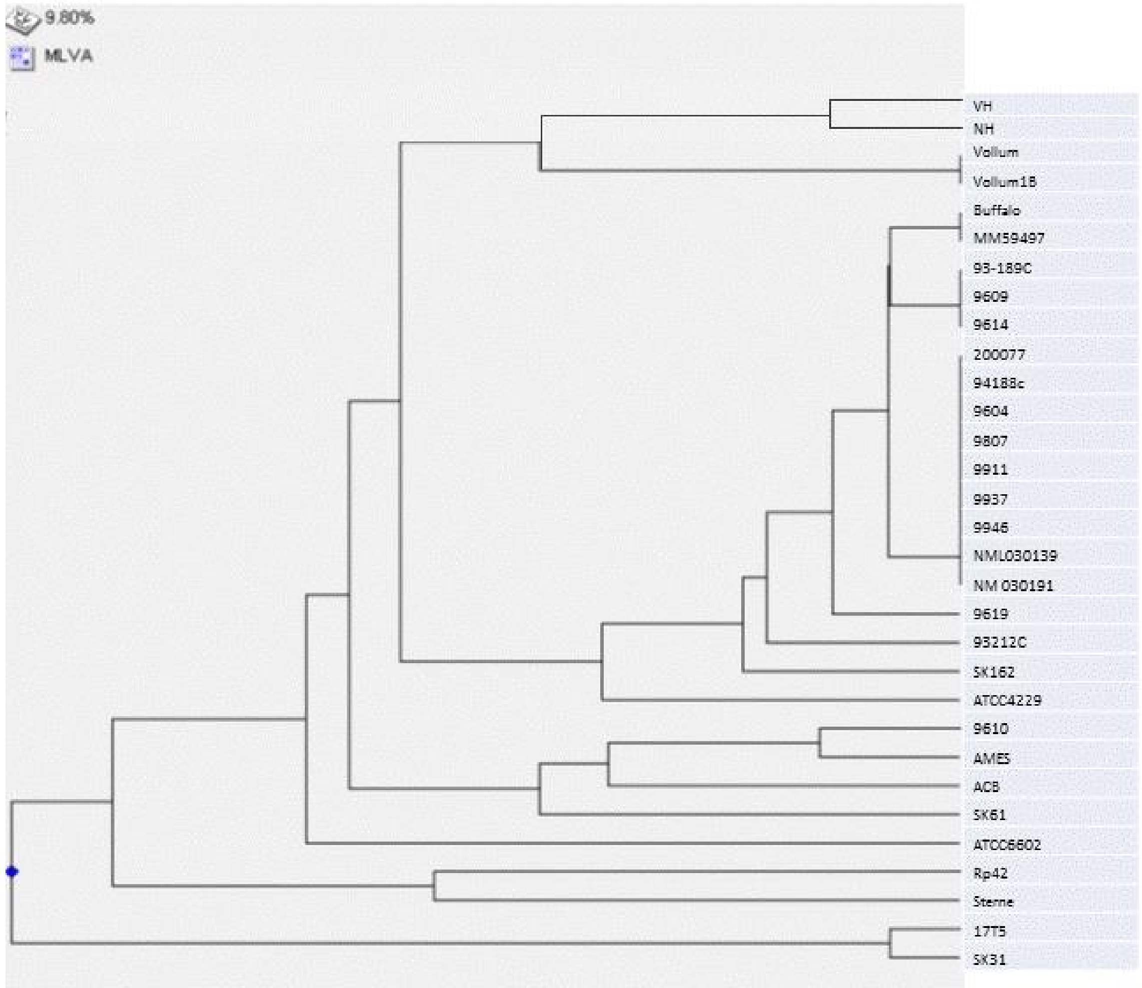

Figure 1. MLVA-15 based dendogram of thirty- one B. anthracis isolates. The 15 MLVA marker loci were used to calculate clustering analysis using a Categorical coefficient and UPGMA options. The genetic distance is presented as percent similarity.

This page left intentionally blank

References

[1] Koehler, T. M. (2002) *Bacillus anthracis* genetics and virulence gene regulation. *Curr. Top. Microbiol. Immunol.,* 71,143-164.

[2] Hill, K. K., Ticknor, L. O., Okinaka, R. T., Asay, M., Blair, H., Bliss, K. A., Laker, M., Pardington, P. E., Richardson, A. P., Tonks, M., Beecher, D. J., Kemp, J. D., Kolsto, A. B., Wong, A. C., Keim, P., and Jackson, P. J. (2004), Fluorescent amplified fragment length polymorphism analysis of *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis* isolates, *Appl. Environ. Microbiol.*, 70 (2), 1068-80.

[3] Pearson, T., Busch, J. D., Ravel, J., Read, T. D., Rhoton, S. D., U'Ren, J. M., Simonson, T. S., Kachur, S. M., Leadem, R. R., Cardon, M. L., Van Ert, M. N., Huynh, L. Y., Fraser, C. M., and Keim, P. (2004), Phylogenetic discovery bias in *Bacillus anthracis* using single-nucleotide polymorphisms from whole-genome sequencing, *Proc. Natl. Acad. Sci .,USA*, 101 (37), 13536-41.

[4] Keim, P., Kalif, A., Schupp, J., Hill, K., Travis, S. E., Richmond, K., Adair, D. M., Hugh-Jones, M., Kuske, C. R., and Jackson, P. (1997), Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers, *J. Bacteriol.*, 179 (3), 818-24.

[5] Okinaka, R., Cloud, K., Hampton, O., Hoffmaster, A., Hill, K., Keim, P., Koehler, T., Lamke, G., Kumano, S., Manter, D., Martinez, Y., Ricke, D., Svensson, R., and Jackson, P. (1999), Sequence, assembly and analysis of pX01 and pX02, *J. Appl. Microbiol.*, 87 (2), 261-2.

[6] Harrell, L. J., Andersen, G. L., and Wilson, K. H. (1995) Genetic variability of *Bacillus anthracis* and related species. *J. Clin. Microbiol*., 33,1847-1850.

[7] Keim, P., Van Ert, M. N., Pearson, T., Vogler, A. J., Huynh, L. Y., and Wagner, D. M. (2004), Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales, *Infect. Genet. Evol.*, 4 (3), 205-13.

[8] Keim, P., Price, L. B., Klevytska, A. M., Smith, K. L., Schupp, J. M., Okinaka, R., Jackson, P. J., and Hugh-Jones, M. E. (2000), Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*, *J. Bacteriol.*, 182 (10), 2928-36.

[9] Kenefic, L. J., Pearson, T., Okinaka, R. T., Chung, W. K., Max, T., Trim, C. P., Beaudry, J. A., Schupp, J. M., Van Ert, M. N., Marston, C. K., Gutierrez, K., Swinford, A. K., Hoffmaster, A. R., and Keim, P. (2008), Texas isolates closely related to *Bacillus anthracis* Ames, *Emerg. Infect. Dis.*, 14 (9), 1494-6.

[10] Van Ert, M. N., W. R. Easterday, L. Y. Huynh, R. T. Okinaka, M. E. Hugh-Jones, J. Ravel, S. R. Zanecki, T. Pearson, T. S. Simonson, J. M. U'Ren, S. M. Kachur, R. R. Leadem-Dougherty, S. D. Rhoton, G. Zinser, J. Farlow, P. R. Coker, K. L. Smith, B. Wang, L. J. Kenefic, C. M. Fraser-Liggett, D. M. Wagner, and Keim, P. (2007) Global genetic population structure of *Bacillus anthracis*. PLoS ONE, **2**e461.

[11] Stratilo, C. W. and Bader, D. (2007). Molecular Typing of *Bacillus anthracis* using Multiple-Locus Variable Number Tandem Repeat Analysis (MLVA-8). (DRDC Suffield TM 2007-182). Defence R&D Canada – Suffield.

[12] Stratilo, C. W., Lewis, C. T., Bryden, L., Mulvey, M. R., and Bader, D. (2006), Single-nucleotide repeat analysis for subtyping *Bacillus anthracis* isolates, *J. Clin. Microbiol.*, 44 (3), 777-82.

[13] Van Ert, M. N., Easterday, W. R., Simonson, T. S., U'Ren, J. M., Pearson, T., Kenefic, L. J., Busch, J. D., Huynh, L. Y., Dukerich, M., Trim, C. B., Beaudry, J., Welty-Bernard, A., Read, T., Fraser, C. M., Ravel, J., and Keim, P. (2007), Strain-specific single-nucleotide polymorphism assays for the Bacillus anthracis Ames strain, *J. Clin. Microbiol.*, 45 (1), 47-53.

[14] Trends in Science and Technology Relevant to the Biological and Toxin Weapons Convention: Summary of an International Workshop 2011: Washington (DC).

Annex A - PCR Setup for MLVA-15 Multiplex Reactions

Multiplex 1

Stock Concentration	Volume (µL) Required for One Reaction in a 25 µL total volume
10 × PCR Buffer	2.5
2 mM dNTPs	2.5
50 mM $MgCl_2$	2.0
Sterile, 18 MS dH_2O	5.95
vrrA1-f1-HEX (10 µM) vrrA-r1 (10 µM)	0.125 0.125
vrrB2-f1-FAM (10 µM) vrrB2-r1 (10 µM)	0.25 0.25
CG3-F1-NED (10 µM) CG3-R1(10 µM)	0.5 0.5
pXO1-AAT-F3-NED (10 µM) pXO1-AAT-R3 (10 µM)	1.0 1.0
pXO2-AT-F1-HEX (10 µM) pXO2-AT-R1 (10 µM)	0.25 0.25
BAVNTR17-6FAM (10 µM) BAVNTR17-REV (10 µM)	1.25 1.25
BAVNTR32-HEX (10 µM) BAVNTR32-REV (10 µM)	1.25 1.25
Platinum *Taq* DNA Polymerase (5U/µL)	0.8
Template DNA (25ng/µL)	2
Total	**25**

Multiplex 2

Stock Concentration	Volume (µL) Required for One Reaction in a 25 µL total volume
10 × PCR Buffer	2.5
2 mM dNTPs	2.5
50 mM $MgCl_2$	2.0
Sterile, 18 MS dH_2O	4.3
HCvrrC1F-6-FAM (10 µM) HCvrrC1R (10 µM)	1.5 1.5
HCvrrC2F-HEX (10 µM) HCvrrC2R (10 µM)	1.5 1.5
vrrB1-f1-HEX (10 µM) vrrB1-r1 (10 µM)	1.0 1.0
BAVNTR19-2-6-FAM (10 µM) BAVNTR19-2-REV (10 µM)	0.5 0.5
BAVNTR35-HEX (10 µM) BAVNTR35-REV (10 µM)	0.5 0.5
BAVNTR23-6-FAM (10 µM) BAVNTR23-REV (10 µM)	0.5 0.5
Platinum *Taq* DNA Polymerase (5 U/µL)	0.7
Template DNA (25 ng/µL -50 ng/µL)	2
Total	**25**

Multiplex 3

Stock Concentration	Volume (µL) Required for One Reaction in a 25 µL total volume
10 × PCR Buffer	2.5
2 mM dNTPs	2.5
50 mM $MgCl_2$	2.0
Sterile, 18 MS dH_2O	9.75
BAVNTR12-6FAM (10 µM) BAVNTR12-REV (10 µM)	1.5 1.5
BAVNTR16-6FAM (10 µM) BAVNTR16-REV (10 µM)	1.5 1.5
Platinum *Taq* DNA Polymerase (5U/µL)	0.25
Template DNA (25ng/µL)	2
Total	**25**

List of symbols/abbreviations/acronyms/initialisms

AFLP	Amplified Fragment Length Polymorphism
CRTI	CBRN Research and Technology Initiative
DND	Department of National Defence
DRDC	Defence Research & Development Canada
DRDKIM	Director Research and Development Knowledge and Information Management
MLST	Multi-Locus Sequence Typing
MLVA	Multi-locus variable-number tandem repeat analysis
PCR	Polymerase Chain Reaction
PFGE	Pulse Field Gel Electrophoresis
PHAC-NML	Public Health Agency of Canada National Microbiology Laboratory
R&D	Research & Development
SNP	Single Nucleotide Polymorphism
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
VNTR	Variable Number Tandem Repeats

This page left intentionally blank

DOCUMENT CONTROL DATA

(Security markings for the title, abstract and indexing annotation must be entered when the document is Classified or Designated)

1. ORIGINATOR (The name and address of the organization preparing the document. Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's report, or tasking agency, are entered in section 8.) Defence Research and Development Canada – Suffield P.O. Box 4000, Station Main Medicine Hat, Alberta T1A 8K6	2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.) UNCLASSIFIED 2b. CONTROLLED GOODS (NON-CONTROLLED GOODS) DMC A REVIEW: GCEC APRIL 2011

3. TITLE (The complete document title as indicated on the title page. Its classification should be indicated by the appropriate abbreviation (S, C or U) in parentheses after the title.)

Molecular typing of Bacillus anthracis isolates using MLVA-15

4. AUTHORS (last name, followed by initials – ranks, titles, etc. not to be used)

Stratilo, C.W; Bader, D.

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a. NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
December 2013	31	14

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of report, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Technical Memorandum

8. SPONSORING ACTIVITY (The name of the department project office or laboratory sponsoring the research and development – include address.)

Defence Research and Development Canada – Suffield
P.O. Box 4000, Station Main
Medicine Hat, Alberta T1A 8K6

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.) DRDC Suffield TM 2013-124	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)

11. DOCUMENT AVAILABILITY (Any limitations on further dissemination of the document, other than those imposed by security classification.)

Unlimited

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11). However, where further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

Bacillus anthracis, the etiological agent of anthrax, lacks character homoplasy and is genetically very monomorphic. Variable number tandem repeats (VNTR) are rapidly evolving stretches of DNA that can serve as markers for strain differentiation (molecular typing). A collection of *B. anthracis* strains were previously characterized using multiple-locus VNTR analysis (MLVA- 8). This work describes the characterization of 31 strains by MLVA-15. MLVA-15 allowed these strains to be placed within a phylogenetic framework as with MLVA-8. MLVA-15 did allow higher resolution and discrimination than MLVA-8. While MLVA-15 allowed many strains to be discriminated from each other, several Canadian isolates were indistinguishable. This work allowed verification of strain designation within the DRDC Suffield collection and allowed for the development of a high throughput genetic typing capability at DRDC Suffield.

Bacillus anthracis, l'agent étiologique de l'anthrax, n'affiche pas d'homoplasie de caractère et est génétiquement très monomorphe. Les nombres variables de répétitions en tandem (VNTR) sont des séquences d'ADN en évolution rapide qui peuvent servir de marqueurs pour la différentiation des souches (typage moléculaire). Une collection de souches de B. anthracis a été caractérisée dans le passé par analyse multi-locus VNTR (MLVA-8). Les présents travaux décrivent la caractérisation de 31 souches par MLVA 15. Par l'analyse MLVA-15, on a pu placer ces souches dans un cadre phylogénique comme on l'avait fait lors de l'analyse MLVA 8, mais à une résolution et discrimination plus élevées que celles obtenues par MLVA-8. Toutefois, bien que l'analyse MLVA 15 ait permis de différencier de nombreuses souches les unes des autres, plusieurs isolats canadiens étaient indistinguables. Ces travaux ont permis de vérifier la désignation des souches figurant dans la collection de RDDC Suffield et de développer une haute capacité de production de typage génétique à RDDC Suffield.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

Bacillus anthracis; Genetic Typing; Identification of bacteria