Molecular characterisation of southern African *Bacillus anthracis* strains

By

Ayesha Hassim

Submitted in fulfilment of the requirements for the degree

Magister Scientiae (Veterinary Tropical Diseases)

Faculty of Veterinary Science, University of Pretoria

Promoter: Dr Henriette van Heerden

Co-promoter: Dr Jenny Rossouw

November 2011

Acknowledgements

All praises are due to the Almighty. I thank Him for leading me onto this path. It has been an incredibly rich and rewarding journey of learning, not in the academics alone, but in life experience as well.

I offer my sincere gratitude to my supervisor Dr. Henriette van Heerden for her gentle, but patient guidance and tutelage. Her integrity, ready humour and unwavering support continues to be a source of inspiration and strength to me.

This project would not have been possible without the funding provided by the National Research Foundation (NRF South Africa) and Institute of Tropical Medicine (ITM Belgium) grants and the facilities and expertise provided by the staff of the Department of Veterinary Tropical Diseases, University of Pretoria.

Special thanks to Mrs Lorraine Arntzen and my co-supervisor Dr. Jenny Rossouw of the National Institute of Communicable Diseases –Special Pathogens Unit, as well as, Dr. Awoke Kidanemariam, Analise van Zyl and the staff of the ARC-Onderstepoort Veterinary Institute for the training and assistance in providing isolates for this project.

I cannot sufficiently express my appreciation to Dr. Gilles Vergnaud, Dr. Christine Pourcel and Yolande Hauck of the Institute of Genetics and Microbiology, University Paris-South Orsay for their invaluable training in MLVA, but also for their warm hospitality and for leaving me with such pleasant memories of France.

Kind thanks to Prof. Allan Guthrie of the Equine Research Centre, University of Pretoria for the patience and time taken in setting up the analysis software and teaching me to be proficient in its use.

I would like to recognise the contributions of Dr. Jackie Picard for her tireless efforts, impeccable work ethic and generous spirit. Recognition also to Dr. Shuvai Musari and Dr. Calvin Gomo of Central Veterinary Laboratory (CVL Zimbabwe) for providing the blood smear samples.

To Melanie, Nedret, Judith and all the staff of SANparks, I thank you for your kind assistance in obtaining the data needed for the geographic modelling of the Kruger National Park isolates.

I offer my affection and appreciation to my friends Taryn, Fahmeeda, Mo'aaza, Raksha, Mamohale, Peninah, Ilse, Shepard, Habib, Melvyn and all the postgrads for the happy memories and in making every day one to look forward to.

I owe a debt of gratitude to my family. The unquestioning love and acceptance of my parents Mohammed and Fazila; the easy affection, gentle words of encouragement and unwavering support of Yusuf, Tahiyya, Jawid, Sakina, Imaan, Blessing, Zakiyya, Aneesa and Akinbowale. My family has ever been the greatest gift to me and my greatest strength. Special mention goes to both my brothers, but most especially Jawid, for the countless occasions they saved my computer from the brink of death. This dissertation would be even longer in coming if not for their mad resuscitation skills.

Table of Contents

CHAPTER 2: A comparative study between agarose and capillary gel electrophoresis using 31 multiple locus variable number tandem repeats analysis (MLVA) for *Bacillus anthracis sensu lato*

Index of Tables

Index of Figures

List of Abbreviations

AFLP	Amplified fragment length polymorphisms
cAMP	Cyclic andenosine monophosphate
canSNP	Canonical single nucleotide polymorphism
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonulceotide
EDTA	Ethylenediaminetetraacetic acid
EF	Oedema factor
KNP	Kruger National Park
LF	Lethal factor
MAPK	Mitogen activated protein kinase
$MgCl_2$	Magnesium chloride
MLVA	Multiple locus VNTR analysis
MLST	Multiple –locus sequence typing
MS	Molecular screening
MST	Minimum Spanning Tree
MSU	Minimum spanning unit
NCP	Northern Cape Province
OIE	Office International des Epizooties
ORF	Open reading frame
PCR	Polymerase chain reaction
PA	Protective antigen
rRNA	Ribosomal ribonucleic acid
SNP	Single Nucleotide Polymorphism
TBE	Tris-Borate EDTA
UPGMA	Unweighted pair group method using arithmetic averages
UV	Ultraviolet
VNTR	Variable number tandem repeats
WHO	World Health Organization

Chapter 1:

Introduction and literature review

1.1 Introduction

Bacillus anthracis, responsible for the disease known as anthrax, has an age-old history going as far back as 1491 BC as one of the ten plagues in the Bible (Farrar, 1995). This microbe is aerobic or facultative anaerobic, Gram-positive, non-motile rods measuring 1.0 µm wide and 3.0-5.0 µm long capable of forming central or terminal spores (endospores) (Dixon *et al.,* 1999; Leppla *et al.,* 2002). Spores are highly resistant to a variety of environmental conditions namely heat, cold, ultraviolet (UV) radiation, ionisation radiation, pressure and chemical agents compared to the vegetative forms of *B. anthracis* (Turnbull, 1998; Vilas-Boas *et al.,* 2007). These highly resistant endospores are found in soil at sites where infected animals have previously died (De Vos, 1990; Leppla *et al.,* 2002). A host contracts anthrax through the uptake of spores or vegetative cells. Within the infected animal, the spores germinate to produce vegetative forms that multiply and which eventually kills the host (Leppla, 1982; Mock & Fouet, 2001). Endospores form during nutrient or oxygen starvation and persist for long periods in preparation for the next host, hence completing the life cycle of *B. anthracis* (Leppla, 1982; Farrar, 1994; Mock & Fouet, 2001).

Anthrax is a zoonotic condition and can be transmitted to humans through contact with infected animals or animal products. However, it is primarily a disease which affects wildlife and domesticated livestock. Despite the rare occurrence of human cases, anthrax remains a potential threat because of continued outbreaks of anthrax in animals and the importation of contaminated raw hair products (La Force, 1978; 1994). Anthrax is not transmitted directly from victim to victim, but instead is picked up by herbivores, such as wild ungulates, that spread anthrax with the aid of scavengers to waterholes. Once the infected animal dies, spores are shed and distributed through watersheds and the environment (Dixon *et al.,* 1999). Transmissions by stable flies, mosquitoes and blow flies have also been reported (Braack & De Vos, 1990; Turnbull, 2000). Since endospores form upon exposure to oxygen, the opening of a carcass is prohibited by law in any circumstance where anthrax is suspected and all anthrax cases should be reported to the OIE (Office International des Epizooties), where anthrax is a reportable disease (OIE, 2008).

Molecular typing to differentiate isolates helps in understanding the epidemiology of anthrax, like the relationship between the disease in wild and domestic animals as well as humans. The long resting stage of the spores greatly reduces the rate of evolutionary change (i.e. it is a slowly evolving organism) (Keim *et al.,* 2000). Therefore, developing

molecular typing for the highly monomorphic *B. anthracis* species is challenging. Low diversity among *B. anthracis* isolates has been reported with most molecular techniques worldwide (Keim *et al.*, 1999; Read *et al.,* 2002). Hypervariability among short tandem repeats (indicated as variable number tandem repeats, VNTR) has been useful for strain typing (Andersen *et al.*, 1996; Jackson *et al.*, 1997; Lindstedt, 2005; Hyytia-Trees *et al.*, 2010). VNTRs are short nucleotide sequences that are polymorphic in the number of tandem repeats (copy numbers). VNTRs that may vary in copy number are amplified by primers flanking the repeat. Different copy numbers of VNTR are detected by PCR amplicons of different sizes (Keim *et al.,* 2000). Tandem repeats are increasingly recognised as the markers of choice for genotyping many pathogens using multi-locus VNTR analysis (MLVA). MLVA has been developed by various laboratories using either agarose or capillary electrophoresis to determine the amplicon size (Le Fleche *et al.*, 2001; Lista *et al.,* 2006; Beyer & Turnbull, 2009).

1.2 Literature Review

1.2.1 Taxonomy

Bacillus anthracis is related to 3 genetically similar, but phenotypically different bacteria among the broader group known as the *Bacillus cereus* group. *Bacillus cereus*, *B. thuringiensis,* and *B. anthracis* are soil borne pathogens which demonstrate similarities in their ability to sporulate, but differ in aetiology and the presence of circular DNA fragments called plasmids. The other 3 members of the *B. cereus* group are: *B. mycoides, B. pseudomycoides* and *B. weihenstephanensis* (Daffonchio *et al.*, 2000; Radnedge *et al.*, 2003).

The *B. cereus* group are Gram positive, rod shaped bacteria, which means that their cell surfaces are positively charged and stain dark blue or violet when Gram stained. All Gram positive bacteria have a thick cell wall, which protects the cell contents from mechanical damage by external forces. *Bacillus anthracis* can be distinguished from other members of the *B. cereus* group by its characteristic lack of motility, its absence of haemolytic activity on blood agar and penicillin sensitivity. *Bacillus anthracis* colonies are mucoid in appearance, on bicarbonate medium, with optimal growth at 37°C and show a clear zone of lysis after inoculation with gamma phage (Parry *et al.*, 1983; Turnbull, 1998; Vilas-Boas *et al.*, 2007).

Anthrax spores are able to withstand harsh environmental conditions for prolonged periods (decades in length) due, in part, to its morphology. *Bacillus anthracis* spores are covered by three prominent protective layers: the cortex (composed of a thick peptidoglycan layer); the coat (covers the cortex in a large number of proteins); and the exosporium (a balloon-like layer) that serves as the primary contact surface between the spore and the environment (Ezzell & Abshire, 1988; 1992; Liu *et al.,* 2004; 2008).

1.2.2 Ecology

The ability of anthrax spores to persist in the earth for extended periods is due to specific soil and climatic conditions, which provide an environment conducive to its survival (De Vos, 1990; Smith *et al.*, 2000). Soils rich in organic matter, ample in calcium, prone to alkalinity and with an ambient temperature above 15.5 °C are considered optimal for persistence of spores. The exosporium of the spore is negatively charged (this charge and its strength are pH dependent) whereas humus particles are positively charged and are therefore able to chelate (i.e. collect and hold bacteria), which then provides environmental stability for the spores (Hugh-Jones & Blackburn, 2009). This ensures that high levels of spores are available at sites where carcasses reside, levels which could be potentially lethal to the next grazer (Coker *et al.*, 2002). These humus-spore clumps have the added advantage of being buoyant. During flooding, these clumps can be deposited or concentrated at water collection sites once the water evaporates, as the spores remain unaffected by UV radiation (Hugh-Jones and de Vos, 2002; Vilas-Boas *et al.*, 2007). It is possible that the negatively charged anthrax spores would attract divalent cations, such as calcium, which would then aid in maintaining the spore core matrix, spore viability and germinative ability (Smith *et al.,* 1955*;* 2000; Hugh-Jones & De Vos, 2002; Hugh-Jones & Blackburn, 2009).

1.2.3 Routes of Infection

There are three routes to infection: cutaneous, gastrointestinal and inhalation (Turnbull 1998; Vilas-Boas *et al.,* 2007). If left untreated, all three can progress to fatal systemic anthrax (Dixon *et al.,* 1999), although the inhalational form is considered the most lethal of the three (Turnbull *et al.,* 1988). Incubation periods before death are 2-7 days for cutaneous infection, 3 hours up to 3 days for inhalational infection and generalized toxaemia during gastrointestinal infection resulting from 1 to 7 day incubation period (Hambleton & Turnbull, 1990).

1.2.4 Mode of Action

Bacillus anthracis strains contain two plasmids, pXO1 (181 kb) and pXO2 (95 kb), which are required for virulence. They are responsible for toxin production and encapsulation, respectively (Green *et al.,* 1985; van Schaik *et al.,* 2007). These two main virulence factors, pXO1 and pXO2, are dependent on the action of antiphagocytic poly-D-glutamic acid capsule and the tripartite protein (Dixon *et al.,* 1999). Anthrax exotoxin is composed of the cell-binding protein, protective antigen (PA, *pag* gene, 83 kDa), which is cleaved after binding to a receptor on the host cell, creating a binding site for either lethal factor (LF, *lef* gene, 90 kDa) or oedema factor (EF, *cya* gene, 89 kDa) (Petosa *et al.*, 1997; Pannifer *et al.*, 2001).

The two exotoxins are therefore produced through an interaction of PA and LF (Figure 1.1) forming the lethal toxin and PA and EF forming the oedema toxin (Singh *et al.,* 1994). PA is cleaved and activated by a host protease to form PA_{63}, which forms a heptamer with an exposed binding site for either of the two toxic enzymes, LF or EF (Leppla, 1991; Redmond *et al.,* 2004). The resulting toxin complex is endocytosed into intracellular compartments called endosomes. A conformational change in PA_{63} is triggered by the natural acidification of the endosome, which then confers the toxic enzyme into the host cell interior. Once inside, the toxin is able to damage the host cell defence system (O'Brien *et al.*, 1985; Duesbury *et al.*, 1998).

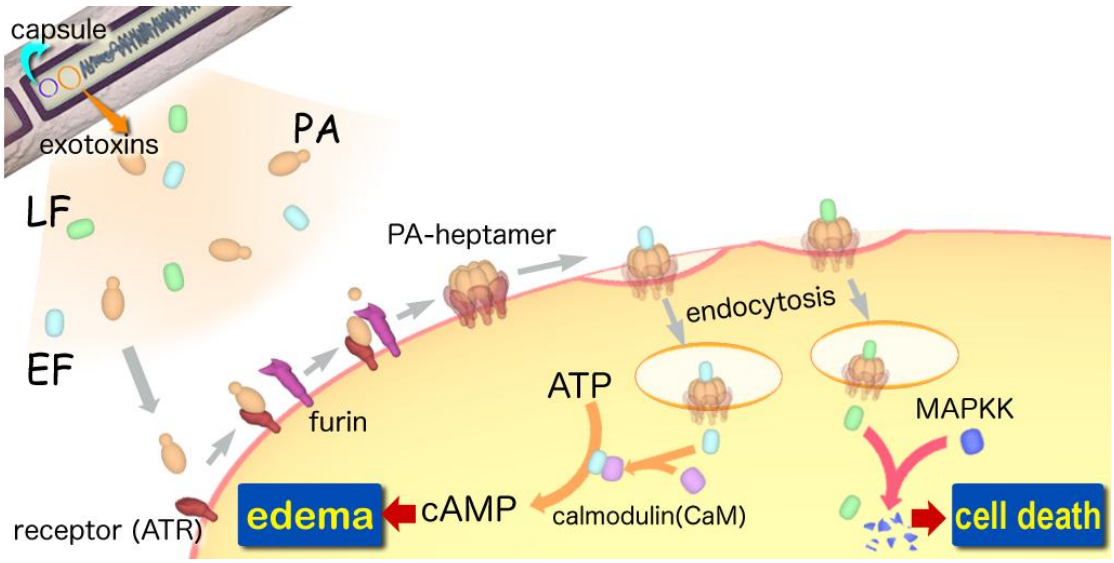


Figure 1.1: Mode of action of anthrax exotoxins leading to oedema and cell death (GNU Free documentation agreement) http://en.wikipedia.org/wiki/Image:Anthraxtoxins_diagram_en.png

The PA and two enzyme components, EF and LF are expressed by pXO2. The number of pXO2 plasmids in each bacterial cell, contributes to the level of virulence associated with that isolate (Pannifer *et al.*, 2001; Coker *et al.,* 2003). However, a tremendous variation in pXO1 and pXO2 copy numbers have been observed in a genetically diverse collection of isolates using a semi-quantitative technique (Coker *et al.,* 2003; Hoffmaster *et al.*, 2004; Rasko *et al.,* 2007). It is during the vegetative form of *B. anthracis* that all known anthrax virulence genes are expressed. This results from the germination of spores within the body (Dixon *et al.,* 2000).

EF is a calmodulin-activated adenylate cyclase that generates cyclic adenosine monophosphate (cAMP) to abnormal levels, which attracts fluid within and between cells, a process known as oedema. The disruption of normal signalling pathways results in cell lysis, the sudden release of messenger molecules and toxic shock (Leppla, 1982; Pannifer *et al.,* 2001). LF helps the pathogen to evade the immune system through killing macrophages. In macrophages LF acts as an endoprotease that removes the N-terminus of the mitogen activated protein kinase kinase (MAPKK) (Duesbury *et al.*, 1998). LF achieves this by removing the specific fragment from individual MAPKs that is crucial for immediate interaction with other signalling proteins. This action rapidly blocks the signals that would normally recruit other immune cells to fight the infection that results in apoptosis/ programmed cell death (O'Brien *et al.*, 1985; Duesbury *et al.*, 1998).

1.2.5 Symptoms and Pathology

Clinical symptoms in animals include fever, shock, multiple organ failure (Leppla *et al.,* 2002), discharge/bleeding from various orifices, lesions which resemble black eschars/ carbuncles (Vilas-Boas *et al.,* 2007), cardiac distress, respiratory distress and oedematous swelling of the tongue, face and throat. Often, death occurs shortly after the appearance of the first symptoms (Hambleton *et al.,* 1984). Usually, animals that die from anthrax do not display signs of *rigor mortis.* As the names 'splenic disease' / "miltsiekte" suggest, the most obvious sign during necropsy is severe splenomegaly (Hugh-Jones & De Vos, 2002).

1.2.6 Diagnostics

A Gram stain or Giemsa stain of the blood smear is the first step taken in confirming the presence of *B. anthracis* in any case where anthrax is suspected. Furthermore, M'Faydean reaction (polychrome methylene blue stain) can be included for confirmation of the presence of capsules. Once the presence of Bacilli is confirmed microscopically, a sample (blood/tissue/bone) is sent to the reference laboratory for confirmation (Turnbull, 1998). In human cases, a serum sample should also be included for serological assays, which involve testing for toxin antigen (Henry, 2001). At the reference laboratory, the Bacilli are isolated and thereafter cultured on blood agar and examined for colony morphology, lack of haemolytic activity, penicillin sensitivity and bacteriophage sensitivity in order to be identified as *B. anthracis* (Turnbull, 1998). A small but significant percentage of anthrax isolates demonstrate either bacteriophage or penicillin resistance. Therefore, none of these tests should be performed in isolation for the definitive identification of *B. anthracis* (Coker *et al.,* 2002; Fulmer *et al.,* 2003). Further testing by PCR for the *cap* and *pag* genes which code for the plasmid encoded toxins is commonly practiced (Turnbull, 1998; OIE, 2008).

1.2.7 Vaccine and Control

Prompt and timely antibiotic treatment generally results in the recovery of the animal or human infected with anthrax. Administration of antibiotics does not guarantee patient survival once anthrax toxins are present systemically in lethal quantities, especially in the case of inhalational anthrax (Turnbull, 1991; 1998; Henry, 2001). *Bacillus anthracis* is penicillin sensitive and can also be treated with tetracycline, chloramphenicol and streptomycin quite effectively (Odendaal *et al.,* 1991). Even if animals die after antibiotic treatment, the infective load of *B. anthracis* will be greatly reduced (Turnbull, 1998).

Robert Koch established the bacterial aetiology of the disease in the 1850's *(*Hambleton *et al.,* 1984). *Bacillus anthracis* was used as the first bacterial vaccine by Pasteur in 1881 (Turnbull, 1991; Leppla *et al.,* 2002). The heat attenuated *B. anthracis* culture was capable of forming capsules, but not toxins (Sterne, 1937; Farrar, 1994). The Pasteur vaccine provides a lower level of protective immunity than toxigenic vaccine strains, like the Sterne

strain $34F_2$ (Sterne, 1937). Widespread vaccination began in the 1930's, using the Sterne attenuated strain, which virtually abolished anthrax in industrialized countries (Leppla *et al.* 2002). By 1939 the Sterne vaccine had proven itself to be safe and effective for use in domesticated animals. It seemed to completely protect them against challenge with highly virulent strains of *B. anthracis*. Live attenuated strains are still undesirable for human vaccine use (Hambleton *et al.,* 1984; Hambleton & Turnbull, 1990). While vaccination programs and good farming practices have gained the upper hand in control of anthrax in industrialised countries, it still remains a serious problem in many developing countries (Farrar, 1994).

1.2.8 Distribution and Spread

Anthrax is a disease which occurs in many parts of the world, including Asia, central and South America, southern Europe, sub-Saharan Africa and small pockets of Australia, northwest Canada and Haiti (Hugh-Jones *et al.*, 2000; Keim *et al.*, 2000; Smith *et al.*, 2000). Sporadic cases may occur in other countries, usually traceable to imports of contaminated animal material (Bales *et al.*, 2002). Southern Africa is among the regions where anthrax regularly occurs endemically or epidemically (OIE handistatus II).

In South Africa there is currently two endemic regions namely the Kruger National Park (KNP) and Ghaap region in Northern Cape Province (NCP). Annual reports to OIE show the occurrence of outbreaks in domestic and wild animals every year in Namibia, South Africa and Zimbabwe and, at least in wild animals, most years in Botswana (OIE handistatus II). It is suspected that under-reporting occurs in this region especially in wildlife because of surveillance inadequacies and difficulties monitoring range wildlife. The official reports therefore probably fall short of the true incidence in many instances (OIE handistatus II).

The introduction of the Sterne spore vaccine and other legislative processes influenced outbreak trends in South Africa (Figure 1.2). At the peak of the outbreaks in 1920s, anthrax killed between 30000-60000 animals in one year (1923) in South Africa (Sterne, 1967; Fasanella *et al.,* 2010). During the time the vaccine was successful in reducing the incidence of outbreaks, especially once it was freely dispensed and its use became common practice. The outbreak events steadily dropped over the next 4 decades. The decrease is quite marked after 1937 when the Sterne vaccine was introduced as a compulsory annual vaccine

distributed freely by the government amongst farmers and African-owned stock (Gilfoyle, 2006). Between 1968 and the mid 1980's, it became mandatory to inoculate livestock in endemic areas and the incidence of outbreaks dwindled to virtually nothing except for isolated cases (De Vos & Turnbull, 1994).

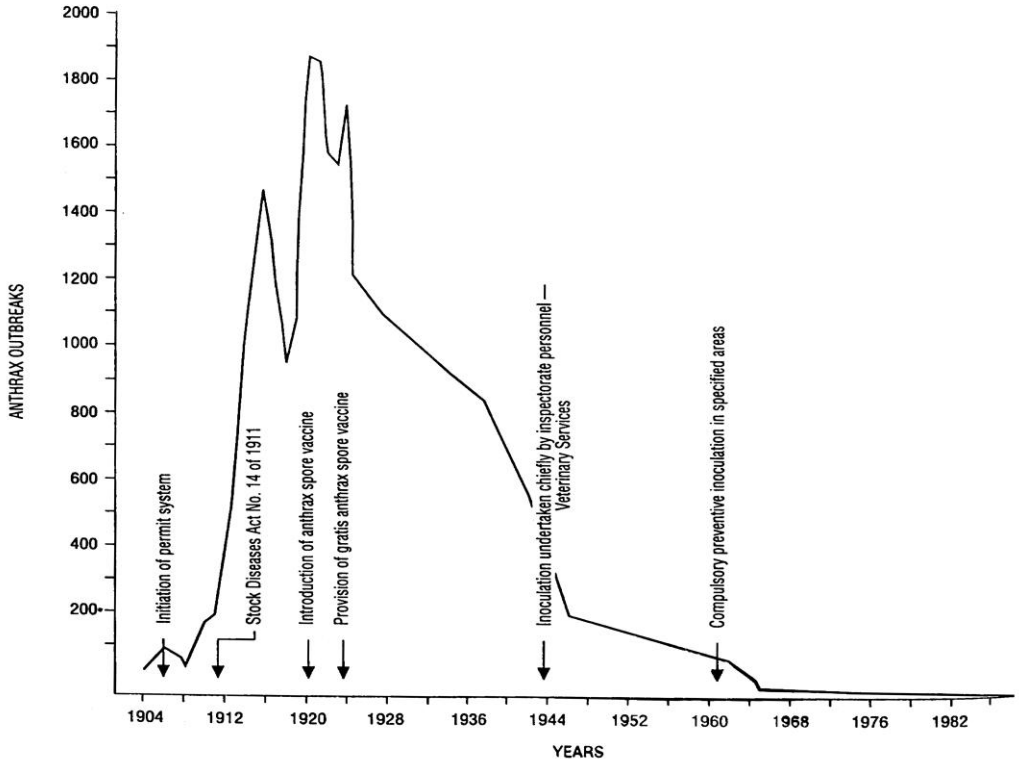


Figure 1.2: Distribution and spread of anthrax in South Africa demonstrating peak periods of the disease. Taken from De Vos & Turnbull (1994) in Infectious Diseases of Livestock, Volume 3, Chapter 176 (editors Coetzer, J.A. and Tustin, R.C.).

1.2.9 Southern African Outbreaks

1.2.9.1 Outbreaks in African Game Parks

Keim *et al.* (2000) believe sub-Saharan Africa to be the geographic origin of anthrax. Many of the pristine natural reserves experience cyclical/periodic epidemics. The following conservation parks experience outbreaks as part of the ecological cycle (topography and climatic conditions): the Queen Elizabeth National Park in Uganda; the Omo-Mago National Park in Ethiopia; the Selous Nature Reserve in Tanzania; the Luangwa Valley in Zambia; the Etosha National Park in northern Namibia; the Kgalagadi Transfrontier Park bordering South Africa and Botswana; Hwange as well as Malilangwe Wildlife Reserve in Zimbabwe and KNP in South Africa (Hugh-Jones & De Vos, 2002).

In southern Africa, anthrax does not affect all herbivorous animal species equally and there is an apparent preference by the disease for a particular species in any one region. Zebras (*Equus burchelli*), for example, are the most commonly affected species (45 % of all recorded anthrax deaths) in the Etosha National Park with kudu (*Tragelaphus strepsiceros*) only occasionally affected (0.8 % of anthrax deaths) (Lindeque & Turnbull, 1994). In the KNP, the kudu is the principal host accounting for >50% of all recorded anthrax cases with zebra falling into a relatively small group of "other affected species" (De Vos, 1990). In Malilangwe, the kudu once again seems to be the predominant host followed by the hippopotamus (*Hipopotamus amphibious*) (Clegg *et al.,* 2007). In the Luangwa Valley hippopotamus and elephant (*Loxodonta Africana)* have seen the highest mortalities (Turnbull *et al.*, 1991) while in the Kgalagadi, the nyala (*Tragelaphus Agassi*), buffalo (*Syncerus caffer*), kudu and giraffe are equally affected (Hugh-Jones & De Vos, 2002). Anthrax is a multispecies disease that can infect mammals. Ruminants and hindgut digesters are most susceptible, whereas carnivores and primates (including humans) are more resistant. This is supported in serological surveys in the anthrax endemic Etosha National Park where it was discovered that naturally acquired anthrax antibodies were rare in herbivores, but common amongst carnivores (Turnbull *et al.*, 1992).

During outbreaks in the KNP in previous years, it was noted that a high stocking rate and large percentage of susceptible host species needed to be present in order for an epidemic to occur. Hugh-Jones & De Vos (2002) reported that anthrax failed to kill up to 20% of its hosts and left a great proportion of young animals to be counted among the living, which makes it an ideal means of natural culling. The authors concluded that in a natural setting, an anthrax outbreak is dependent on the stocking density and is self-limiting (Hugh-Jones & De Vos, 2002).

1.2.9.2 Vectors

Circumstantial evidence points to certain ecological and behavioural factors as being responsible for the differences in preferential species infection and dissemination during an outbreak (Figure 1.3). It has been observed that blowflies (*Chrysomyia albiceps and C.*

marginalis) in the KNP, after feeding on anthrax carcasses, fly to nearby shrubbery and deposit infected blood on the leaves that are 1 – 3 m in height, which then become the source of infection for the browsing kudus (Hugh-Jones & De Vos, 2002). It has been speculated that insects spread anthrax to nearby and more distant shrubbery as *C. albiceps* has been known to circulate up to a 32.5 km radius, while *C. marginalis* has been noted 25 km from a source (Braack & De Vos, 1990; Turnbull, 2000).

Besides predators and scavengers like the spotted hyena (*Crocuta crocuta*), other carrion feeders of interest are vultures including the white backed vulture (*Gyps africanus*), Cape vulture (*G. coprotheres*) and Lap pet-faced vultures (*Aegypius tracheliotus*), which are believed to open the carcasses and spread the pathogen from anthrax infected carcasses to the environment and watering holes (Turnbull *et al.*, 2008). It is believed that vultures are able to consume the bacilli laden soft-tissue of an anthrax carcass before the bacterium is exposed to air and is able to sporulate and spread spores whereas the hardy digestive system of this scavenger easily kills off the bacteria in its vegetative form (De Vos, 1990; Hugh-Jones & De Vos, 2002; Saggese *et al.*, 2007; Turnbull *et al.*, 2008). Cases of vulture deaths from anthrax are extremely rare. Vultures from the Magaliesburg Mountains in South Africa have a documented flight radius of 300 – 500 km (Saggese *et al.*, 2007; Turnbull *et al.*, 2008) and have been linked to outbreaks in neighbouring countries, like Botswana, that fall within this radius (Hugh-Jones & De Vos, 2002).

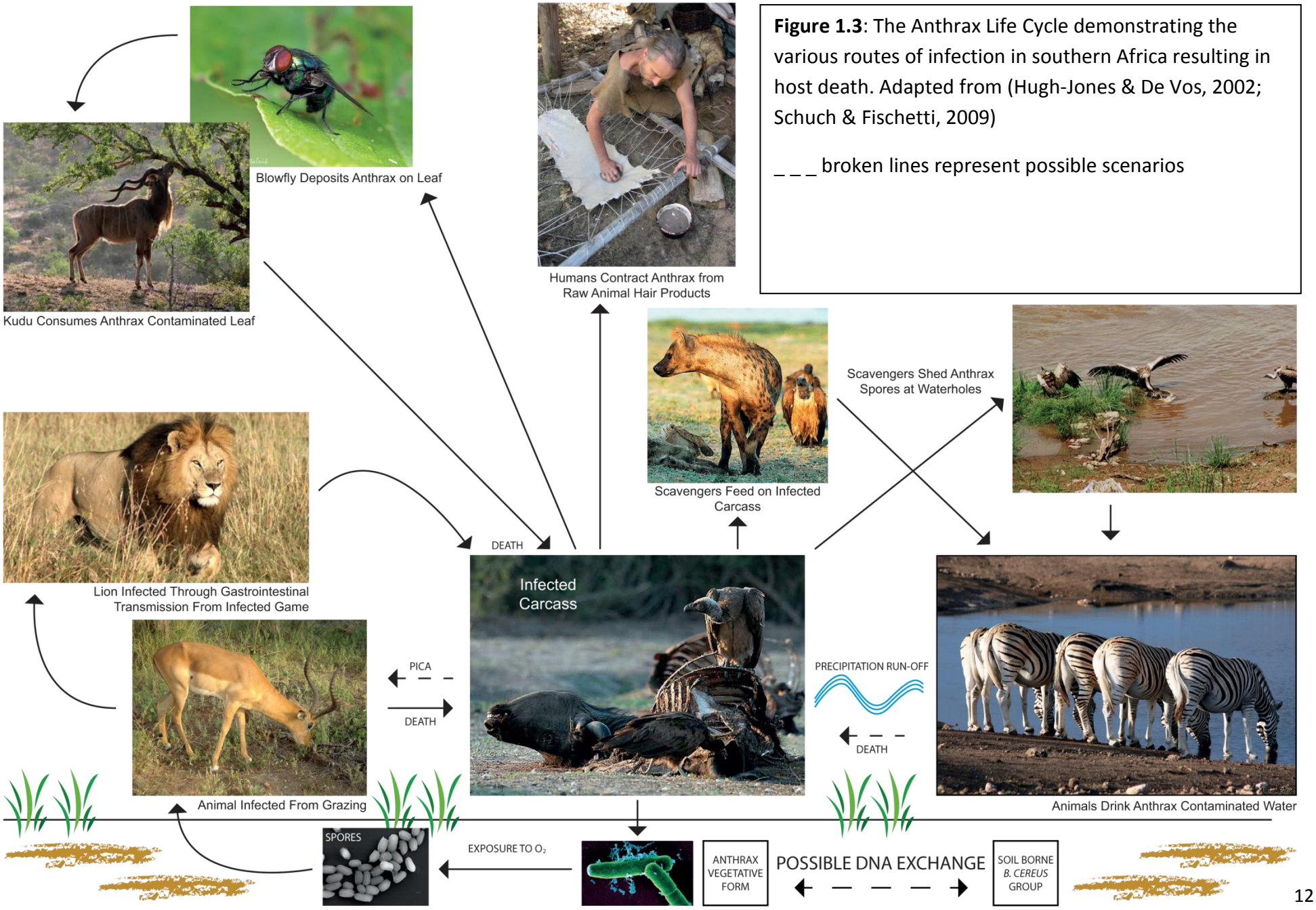


Figure 1.3: The Anthrax Life Cycle demonstrating the various routes of infection in southern Africa resulting in host death. Adapted from (Hugh-Jones & De Vos, 2002; Schuch & Fischetti, 2009)

_ _ _ broken lines represent possible scenarios

1.2.10 Molecular Techniques

Molecular techniques have provided valuable insight into discrimination of bacterial species. The ribosomal genes are normally worthwhile targets in differentiating bacterial strains (Woese, 1987). Ash *et al.* (1991); Ash and Collins (1992) have determined that amongst the *B. cereus* group, which includes *B. anthracis*, the 16S rRNA is identical and the 23S rRNA sequences differ at only 2 nucleotides amongst the *B. cereus* group. Multi-locus sequence typing (MLST) is another method which is used to differentiate other pathogens, but which fails to adequately discriminate *B. anthracis* strains, due to the rarity of its sequence variation (Daffonchio *et al.*, 2006; Hoffmaster *et al.*, 2008). Tandem repeats on the other hand, have proven to offer informative markers for strain genotyping (Denoeud & Vergnaud, 2004). Tandem repeats are characterized as minisatellites (repeat unit consisting of 6 – 100 base pairs (bp) and microsatellites (repeat unit consisting of 1 – 5 bp) and has been used for the identification of pathogenic bacteria (Le Fleche *et al.*, 2001). Slip strand repair mutations caused by DNA polymerase is known to occur more frequently on the short repeats and can therefore be used for strain typing in the monomorphic *B. anthracis* species (Keim *et al.,* 1999; Achtman *et al.,* 2004) as multiple alleles can exist for each tandem repeat locus (Lista *et al.,* 2006). Tandem repeats therefore undergo rapid evolution which then augments the phenotypic variability of the pathogen i.e. tandem repeats play an important role in the adaptation of the pathogen to the host (Le Fleche *et al.*, 2001).

Tandem repeats contribute phenotypic variation in 2 ways: (1) Tandem repeats in the regulatory region of a gene can constitute an on/ off switch of gene expression at a transcriptional level; and (2) Tandem repeats within coding regions, with repeat unit lengths, which are not a multiple of three can induce a premature end of translation when a mutation changes the number of repeats. Loci with tandem repeats in the microsatellite class are called simple sequence contingency loci, which are believed to have limited value to epidemiological studies but are of interest for pathogenicity studies (Le Fleche *et al.,* 2001; Fouet *et al.*, 2002). The variability in the number of tandem repeats (VNTRs) of a specific VNTR locus only provides a small amount of information. Multiple loci are examined concurrently to further increase its discriminatory power, hence a MLVA system (Le Fleche *et al.*, 2001).

The first panel of markers MLVA8 markers (Keim *et al.,* 2000) were identified using the VNTR loci found in the *Vrr*A gene (Andersen *et al.,* 1996) as well as sequenced amplified fragment length polymorphisms (AFLP) markers of *B. anthracis* isolates *(*Keim *et al.,* 1999; 2000), and VNTR loci found in the plasmids (pXO1 and pXO2) (Keim *et al.,* 1999). Keim *et al.* (2000) used an automated fluorescent DNA sequencer (automated capillary electrophoresis) to determine the size of the PCR fragments. Cluster analysis with the MLVA8 loci identified two most dissimilar cluster groups (clade A and B) (Keim *et al.,* 2000). Isolates in clade A are found internationally and are therefore responsible for most epidemics and outbreaks globally. Clade B strains in contrast are almost exclusively restricted to southern Africa. However, the resolution achieved with the MLVA8 was found to be limiting since it did not have enough resolving power to discriminate between two geographically close isolates in natural outbreaks (Keim *et al.,* 2004; Lista *et al.,* 2006).

Additional markers were developed and the MLVA15 assay was used in combination with canonical single nucleotide polymorphisms (canSNPs) (Keim *et al.,* 2004). The canSNPs on their own have limited resolving power when compared to MLVA. However, canSNPs can be used to define the main clades and key phylogenetic positions in the clades (Keim *et al.*, 2004). The MLVA15 divided the *B. anthracis* isolates into 3 major clades, namely clade A, B and C. The Branch C isolates were found in Asia (Van Ert *et al.*, 2007).

The European group extended the MLVA8 assay to MLVA25 (Le Fleche *et al.,* 2001; Ciammaruconi *et al.,* 2008). The PCR fragment sizes were determined using standard agarose gel, making this technique more accessible by using basic equipment. The MLVA25 loci enabled the description of two new clades (D and E). The D cluster was composed of an isolate of unknown origin and another from Italy, while the E cluster was composed of isolates from Cameroon (Lista *et al.,* 2006). The typing of strains using the MLVA markers enables distinction of outbreaks caused by different strains from those caused by the spread of a single strain, to trace an outbreak strain back to its possible origin and to track the routes of transmission of an outbreak strain within and between animal populations. It is becoming possible to study genotypic diversity in relation to the spatial and temporal

dynamics behind the spread of the disease and possible relationships between genotype and host species (Beyer & Turnbull, 2009).

It is not known whether different *B. anthracis* genotypes (strain types) can account for prevalence in different host species and temporal-spatial distributions. Dominant strains can be identified with molecular techniques (Keim *et al.,* 2000; Le Fleche *et al.,* 2001; Lista *et al.,* 2006; Van Ert *et al.*, 2007; Beyer & Turnbull, 2009) and therefore studies can be undertaken to determine the phenotypic and genotypic basis of their prevalence, including virulence, ability to sporulate, tenacity of spores and their preferential spread by living vectors.

1.2.11 Determining Copy Numbers for MLVA

In MLVA study, PCR amplicons targeting VNTR's are viewed elecrophoretically, on two different platforms (i.e. agarose and capillary) to determine the amplicon size. Agarose gel electrophoresis is a method commonly practiced in molecular laboratories. The equipment and reagents needed for this technique are readily available. Using this platform for MLVA requires long migrations of PCR amplicons to visualise differences in amplicon size (Vergnaud & Pourcel, 2009). The copy number is then calculated using the PCR amplicon size and a comparison table (Appendix: Tables 2 and 3) where the flanking sequence was subtracted from PCR product size and using size of the tandem repeat (in bp) to determine the copy number of tandem repeat unit. The copy number represents the number of times the repeat is produced for each locus and hence describes the allele for that locus. It is necessary to include both a reference strain and molecular marker for every 6 samples tested on a gel to minimise errors originating from gel distortions or poor migrations (Vergnaud & Pourcel, 2009). The accuracy of this method decreases with repeat units smaller than 8 bp as it becomes difficult to differentiate such small differences in band size. The $34F_2$ Sterne vaccine strain (with its known amplicon sizes) was included in every run for comparison and to increase the accuracy of copy number assignments. The Sterne VNTR fingerprint is represented in the first column of Appendix Table 2 and 3.

The capillary electrophoresis technique uses capillaries loaded with a gel matrix. Fluorescently labelled amplicons are run at 500V/cm through the gel matrix and the fluorescence wavelengths are recorded. These results are displayed as an electropherogram

(Buel *et al.*, 2001). This method is clearly able to differentiate between amplicons which differ by even 2 bp; however the results are not consistent across all genetic analyzer machines (Lindstedt, 2005). The make and model as well as capillary length and diameter influence the readings (Buel *et al.,* 2001) as indicated in Appendix Table 4. This is further complicated by differences in the analysis software used to view the electropherograms. The results from the 2 platforms are therefore also not directly comparable without adjustment. The $34F_2$ vaccine strain was used to "calibrate" the genetic analyzer and analysis software so that the copy numbers would be automatically assigned and comparable to the agarose technique.

1.3 Objectives

MLVA is currently the method of choice in the typing of *B. anthracis,* but due to the debate over the level of differentiation offered by the different marker panels, their ability to discriminate between geographically close outbreaks and the predominance of genotypes within an outbreak, the objectives of this study were:

1. To investigate the resolution of the MLVA8, MLVA15, MLVA25 as well as MLVA31 using *B. anthracis sensu lato* strains from southern Africa.

2. To investigate the resolution, accessibility and cost of using agarose and automated capillary electrophoresis to separate MLVA markers.

3. To determine the "dominant" genotypes in southern Africa that favour the longevity and broad distribution of an outbreak strain, as well as, the variation within South Africa.

4. To examine the value of stained blood smears (as it is a readily available sample source) and whether the possibility of retrospective studies using this resource can be explored.

References

Achtman, M., Morelli, G., Zhu, P., Wirth, T., Diehl, I., Kusecek, B., *et al*. (2004). Microevolution and history of the plague *Bacillus, Yersinia pestis*. *Proceedings of the National Academy of Sciences of the United States of America, 101*(51), 17837-17842.

Andersen, G. L., Simchock, J. M., & Wilson, K. (1996). Identification of a region of genetic variability among *Bacillus anthracis* strains and related species. *Journal of Bacteriology, 178*(2).

Ash, C., & Collins, M. D. (1992). Comparative analysis of 23S ribosomal RNA gene sequences of *Bacillus anthracis* and emetic *Bacillus cereus* determined by PCR-direct sequencing. *FEMS Microbiology Letters, 73*(1-2), 75-80.

Ash, C., Farrow, J. A., Dorsch, M., Stackebrandt, E., & Collins, M. D. (1991). Comparative analysis of *Bacillus anthracis, Bacillus cereus*, and related species on the basis of reverse transcriptase sequencing of 16S rRNA. *International Journal of Systematic Bacteriology, 41*(3), 343-346.

Bales, M. E., Dannenberg, A. L., Brachman, P. S., Kaufmann, A. F., Klatsky, P. C., & Ashford, D. A. (2002). Epidemiologic response to anthrax outbreaks: Field investigations, 1950-2001. *Emerging Infectious Diseases, 8*(10), 1163-1174.

Beyer, W., & Turnbull, P. C. (2009). Anthrax in animals. *Molecular Aspects of Medicine, 30*(6), 481-489.

Braack, L. E., & De Vos, V. (1990). Feeding habits and flight range of blow-flies *(Chrysomyia* spp.) in relation to anthrax transmission in the Kruger National Park, South Africa. *The Onderstepoort Journal of Veterinary Research, 57*(2), 141-142.

Buel, E., LaFountain, M., Schwartz, M., & Walkinshaw, M. (2001). Evaluation of capillary electrophoresis performance through resolution measurements. *Journal of Forensic Sciences, 46*(2), 341-345.

Ciammaruconi, A., Grassi, S., Santis, R. d., Faggioni, G., Pittiglio, V., D'Amelio, R., *et al.* (2008). Fieldable genotyping of *Bacillus anthracis* and *Yersinia pestis* based on 25-loci multi locus VNTR analysis. *BMC Microbiology, 8*(21).

Clegg, S. B., Turnbull, P. C., Foggin, C. M., & Lindeque, P. M. (2007). Massive outbreak of anthrax in wildlife in the Malilangwe wildlife reserve, Zimbabwe. *The Veterinary Record, 160*(4), 113-118.

Coker, P. R., Smith, K. L., Fellows, P. F., Rybachuck, G., Kousoulas, K. G., & Hugh-Jones, M. E. (2003). *Bacillus anthracis* virulence in guinea pigs vaccinated with anthrax vaccine

adsorbed is linked to plasmid quantities and clonality. *Journal of Clinical Microbiology, 41*(3), 1212-1218.

Coker, P. R., Smith, K. L., & Hugh-Jones, M. E. (2002). Antimicrobial susceptibilities of diverse *Bacillus anthracis* isolates. *Antimicrobial Agents and Chemotherapy, 46*(12), 3843-3845.

Daffonchio, D., Borin, S., Frova, G., Gallo, R., & Mori, E. (2000). A randomly amplified polymorphic DNA marker specific for the Bacillus cereus group is diagnostic for *Bacillus anthracis*. *Applied Environmental Microbiology, 65*, 1298.

Daffonchio, D., Raddadi, N., Merabishvili, M., Cherif, A., Carmagnola, L., Brusetti, L., *et al*. (2006). Strategy for identification of *Bacillus cereus* and *Bacillus thuringiensis* strains closely related to *Bacillus anthracis. Applied and Environmental Microbiology, 72*(2), 1295-1301.

De Vos, V. (1990). The ecology of anthrax in the Kruger National Park, South Africa. *Salisbury Medical Bulletin,* 68S: 19-23.

De Vos, V. & Turnbull, P.C. (1994) Anthrax. In: Coetzer JA, Thomson GR, Tustin RC, editors. Infectious diseases of livestock, with special reference to Southern Africa, 2nd ed. Vol. 3. Cape Town (South Africa): Oxford University Press Southern Africa. p. 1788–818.

Denoeud, F., & Vergnaud, G. (2004). Identification of polymorphic tandem repeats by direct comparison of genome sequence from different bacterial strains: A web-based resource. *BMC Bioinformatics, 5*, 4.

Dixon, T. C., Fadl, A. A., Koehler, T. M., Swanson, J. A., & Hanna, P. C. (2000). Early *Bacillus anthracis*-macrophage interactions: Intracellular survival survival and escape. *Cellular Microbiology, 2*(6), 453-463.

Dixon, T. C., Meselson, M., Guillemin, J., & Hanna, P. C. (1999). Anthrax. *The New England Journal of Medicine, 341*(11), 815-826.

Duesbery, N. S., Webb, C. P., Leppla, S. H., Gordon, G. M., & Klimpel, K. R. (1998). Proteolytic inactivation of MAP-kinase-kinase by anthrax lethal factor. *Science, 280*, 734.

Ezzell, J. W., & Abshire, T. G. (1992). Serum protease cleavage of *Bacillus anthracis* protective antigen. *Journal of General Microbiology, 138*, 543.

Ezzell, J. W.,Jr, & Abshire, T. G. (1988). Immunological analysis of cell-associated antigens of *Bacillus anthracis*. *Infection and Immunity, 56*(2), 349-356.

Farrar, W. E. (1994). Anthrax: Virulence and vaccines. *Annals of Internal Medicine, 121*(5), 379-380.

Farrar, W. E. (1995). Anthrax: From Mesopotamia to molecular biology. *Spring;58(2):35-8,*

Fasanella, A., Galante, D., Garofolo, G., & Jones, M. H. (2010). Anthrax undervalued zoonosis. *Veterinary Microbiology, 140*(3-4), 318-331.

Fouet, A., Smith, K. L., Keys, C., Vaissaire, J., Le Doujet, C., Levy, M., *et al*. (2002). Diversity among french *Bacillus anthracis* isolates. *Journal of Clinical Microbiology, 40*(12), 4732-4734.

Fulmer, P. A. (2003). Susceptibility of *Bacillus anthracis* to gamma and cherry bacteriophage. (Magister Scientae, Lousiana State University and Agricultural and Mechanical College).

Gilfoyle, D. (2006). Anthrax in South Africa: Economics, experiment and the mass vaccination of animals, c. 1910-1945. *Medical History, 50*(4), 465-490.

Green, B. D., Battisti, L., Koehler, T. M., Thorne, C. B., & Ivins, B. E. (1985). Demonstration of a capsule plasmid in *Bacillus anthracis*. *Infection and Immunity, 49*(2), 291-297.

Hambleton, P., Carman, J. A., & Melling, J. (1984). Anthrax: The disease in relation to vaccines. *Vaccine, 2*(2), 125-132.

Hambleton, P., & Turnbull, P. C. (1990). Anthrax vaccine development: A continuing story. *Advances in Biotechnological Processes, 13*, 105-122.

Henry, L. (2001). Inhalational anthrax: Threat, clinical presentation, and treatment. *Journal of the American Academy of Nurse Practitioners, 13*(4), 164-8; quiz 169-70.

Hoffmaster, A. R., Novak, R. T., Marston, C. K., Gee, J. E., Helsel, L., Pruckler, J. M., *et al*. (2008). Genetic diversity of clinical isolates of *Bacillus cereus* using multilocus sequence typing. *BMC Microbiology, 8*, 191.

Hoffmaster, A. R., Ravel, J., Rasko, D. A., Chapman, G. D., Chute, M. D., Marston, C. K., *et al*. (2004). Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. *Proceedings of the National Academy of Sciences of the United States of America, 101*(22), 8449-8454.

Hugh-Jones, M., Keim, P., Smith, K., & Coker, P. (2000). Chromosomal genetic markers of *Bacillus anthracis* and related epidemiology. *Proceedings of the 9th Symposium of the International Society for Veterinary Epidemiology and Economics, Breckenridge, Colorado, USA, August 6-11 2000,* Id 62.

Hugh-Jones, M., & Blackburn, J. (2009). The ecology of *Bacillus anthracis*. *Molecular Aspects of Medicine, 30*(6), 356-367.

Hugh-Jones, M. E., & de Vos, V. (2002). Anthrax and wildlife. *Revue Scientifique Et Technique (International Office des Epizootics), 21*(2), 359-383.

Hyytia-Trees, E., Lafon, P., Vauterin, P., & Ribot, E. M. (2010). Multilaboratory validation study of standardized multiple-locus variable-number tandem repeat analysis protocol for shiga toxin-producing *Escherichia coli* O157: A novel approach to normalize fragment size data between capillary electrophoresis platforms. *Foodborne Pathogens and Disease, 7*(2), 129-136.

Jackson, P. J., Walthers, E. A., Kalif, A. S., Richmond, K. L., Adair, D. M., Hill, K. K., *et al*. (1997). Characterization of the variable-number tandem repeats in vrrA from different *Bacillus anthracis* isolates. *Applied and Environmental Microbiology, 63*(4), 1400-1405.

Keim, P., Klevytska, A. M., Price, L. B., Schupp, J. M., Zinser, G., Smith, K. L., *et al*. (1999). Molecular diversity in *Bacillus anthracis*. *Journal of Applied Microbiology, 87*(2), 215-217.

Keim, P., Price, L. B., Klevytska, A. M., Smith, K. L., Schupp, J. M., Okinaka, R., *et al*. (2000). Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *Journal of Bacteriology, 182*(10), 2928-2936.

Keim, P., Van Ert, M. N., Pearson, T., Vogler, A. J., Huynh, L. Y., & Wagner, D. M. (2004). Anthrax molecular epidemiology and forensics: Using the appropriate marker for different evolutionary scales. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases, 4*(3), 205-213.

La force, F. M. (1978). Woolsorters' disease in England. *Bulletin of the New York Academy of Medicine, 54*(10), 956-963.

La Force, F. M. (1994). Anthrax. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America, 19*(6), 1009-13; quiz 1014.

Le Fleche, P., Hauck, Y., Onteniente, L., Prieur, A., Denoeud, F., Ramisse, V., *et al*. (2001). A tandem repeats database for bacterial genomes: Application to the genotyping of *Yersinia pestis* and *Bacillus anthracis*. *BMC Microbiology, 1*, 2.

Leppla, S. H. (1982). Anthrax toxin edema factor: A bacterial adenylate cyclase that increases cyclic AMP concentrations in eukaryotic cells. *Proc.Natl.Acad.Sci.USA, 79*, 3162.

Leppla, S. H. (1991). Purification and characterization of adenylyl cyclase from *Bacillus anthracis*. *Methods in Enzymology, 195*, 153-168.

Leppla, S. H., Robbins, J. B., Schneerson, R., & Shiloach, J. (2002). Development of an improved vaccine for anthrax. *The Journal of Clinical Investigation, 110*(2), 141-144.

Lindeque, P. M., & Turnbull, P. C. (1994). Ecology and epidemiology of anthrax in the Etosha National Park, Namibia. *Onderstepoort J.Vet.Res., 61*, 71.

Lindstedt, B. A. (2005). Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis, 26*(13), 2567-2582.

Lista, F., Faggioni, G., Valjevac, S., Ciammaruconi, A., Vaissaire, J., le Doujet, C., *et al*. (2006). Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. *BMC Microbiology, 6*, 33.

Liu, C. Q., Nuttall, S. D., Tran, H., Wilkins, M., Streltsov, V. A., & Alderton, M. R. (2008). Construction, crystal structure and application of a recombinant protein that lacks the collagen-like region of *Bcl*A from *Bacillus anthracis* spores. *Biotechnology & Bioengineering, 99*(4), 774-782.

Liu, H., Bergman, N. H., Thomason, B., Shallom, S., Hazen, A., Crossno, J., *et al*. (2004). Formation and composition of the *Bacillus anthracis* endospore. *Journal of Bacteriology, 186*(1), 164-178.

Mock, M., & Fouet, A. (2001). Anthrax. *Annual Review of Microbiology, 55*(1), 647-671.

O'Brien J, Friedlander, A., Dreier, T., Ezzell. J, & Leppla, S. (1985). Effects of anthrax toxin components on human neutrophils. *Infect.Immun., 47*, 306.

Odendaal, M. W., Pieterson, P. M., de Vos, V., & Botha, A. D. (1991). The antibiotic sensitivity patterns of *Bacillus anthracis* isolated from the Kruger National Park. *The Onderstepoort Journal of Veterinary Research, 58*(1), 17-19.

OIE (2008). *Manual of diagnostic tests and vaccines for terrestrial animals*. Paris, world organisation for animal health (OIE).

OIE handistatus II: www.oie.int/hs2.

Pannifer, A. D., Wong, T. Y., Schwarzenbacher, R., Renatus, M., Petosa, C., Bienkowska, J., *et al*. (2001). Crystal structure of the anthrax lethal factor. *Nature, 414*(6860), 229-233.

Parry, J. M., Turnbull, P. C. B., & Gibson, J. R. (1983). A colour atlas of *Bacillus* species. London: Wolfe Medical Publications.

Petosa, C., Collier, R. J., Klimpel, K. R., Leppla, S. H., & Liddington, R. C. (1997). Crystal structure of the anthrax toxin protective antigen. *Nature, 385*, 833.

Radnedge, L., Agron, P. G., Hill, K. K., Jackson, P. J., Ticknor, L. O., Keim, P., *et al*. (2003). Genome differences that distinguish *Bacillus anthracis* from *Bacillus cereus* and *Bacillus thuringiensis*. *Applied and Environmental Microbiology, 69*(5), 2755-2764.

Rasko, D. A., Rosovitz, M. J., Oekstad, O. A., Fouts, D. E., Jiang, L., Cer, R. Z., *et al*. (2007). Complete sequence analysis of novel plasmids from emetic and periodontal *Bacillus cereus* isolates reveals a common evolutionary history among the B. cereus-group plasmids, including *Bacillus anthracis* pXO1. *Journal of Bacteriology, 189*(1), 52-64.

Read, T.D., Salzberg, S.L., Pop, M., Shumway, M., Umayam, L., Jiang, L., Holtzapple, E., Busch, J.D., Smith, K.L., Schupp, J.M., Solomon, D., Keim, P., Fraser, C.M., (2002). Comparative genome sequencing for discovery of novel polymorphisms in *Bacillus anthracis.* Science 296, 2028-2033.

Redmond, C., Baillie, L. W., Hibbs, S., Moir, A. J., & Moir, A. (2004). Identification of proteins in the exosporium of *Bacillus anthracis*. *Microbiology (Reading, England), 150*(Pt 2), 355-363.

Saggese, M. D., Noseda, R. P., Uhart, M. M., Deem, S. L., Ferreyra, H., Romano, M. C., *et al*. (2007). First detection of *Bacillus anthracis* in feces of free-ranging raptors from central Argentina. *Journal of Wildlife Diseases, 43*(1), 136-141.

Schuch, R., & Fischetti, V. A. (2009). The secret life of the anthrax agent *Bacillus anthracis*: Bacteriophage-mediated ecological adaptations. *PloS One, 4*(8), e6532.

Singh, Y., Klimpel, K. R., Arora, N., Sharma, M., & Leppla, S. H. (1994). The chymotrypsin-sensitive site, FFD315, in anthrax toxin protective antigen is required for translocation of lethal factor. *J.Biol.Chem., 269*, 29039.

Smith, K. L., De Vos, V., Bryden, H., Price, L. B., Hugh-Jones, M. E., & Keim, P. (2000). *Bacillus anthracis* diversity in Kruger National Park. *Journal of Clinical Microbiology, 38*(10), 3780-3784.

Smith, S.H., Keppie, J., Stanley, J.L. (1955). The chemical basis of the virulence of *Bacillus anthracis*. *Br.J.Exp.Pathol., 36*, 460.

Sterne, M. (1967). Distribution and economic importance of anthrax. *Federation Proceedings, 26*(5), 1493-1495.

Sterne, M. (1937). Avirulent anthrax vaccine. *Onderstepoort Journal of Veterinary Science, 21*, 41.

Turnbull, P. C. B. (1991). Anthrax vaccines: Past, present and future. *Vaccine, 9*, 533.

Turnbull, P. C. B. (2000). Current status of immunization against anthrax: Old vaccines may be here to stay for a while. *Curr.Opin.Infect.Dis., 13*, 113.

Turnbull, P. C. B. (Ed.). (1998). WHO, EMC/ZDI/98 guidelines for the surveillance and control of anthrax in humans and animals *(fourth edition).*

Turnbull, P. C. B., Hutson, R. A., Ward, M. J., Jones, M. N., & Quinn, C. P. (1992). *Bacillus anthracis* but not always anthrax. *J.Appl.Bact., 72*, 21.

Turnbull, P. C. B., Leppla, S. H., Broster, M. G., Quinn, C. P., & Melling, J. (1988). Antibodies to anthrax toxin in humans and guinea pigs and their relevance to protective immunity. *Med.Microbiol.Immunol., 177*, 293.

Turnbull, P. C., Diekmann, M., Kilian, J. W., Versfeld, W., De Vos, V., Arntzen, L., *et al*. (2008). Naturally acquired antibodies to *Bacillus anthracis* protective antigen in vultures of southern africa. *The Onderstepoort Journal of Veterinary Research, 75*(2), 95-102.

van Schaik, W., Prigent, J., & Fouet, A. (2007). The stringent response of *Bacillus anthracis* contributes to sporulation but not to virulence. *Microbiology, 153*(Pt 12), 4234-4239.

Van Ert, M. N., Easterday, W. R., Huynh, L. Y., Okinaka, R. T., Hugh-Jones, M. E., Ravel, J., et al. (2007). Global genetic population structure of *Bacillus anthracis*. *PLoS ONE, 2*(5), e461

Vergnaud, G., & Pourcel, C. (2009). Multiple locus variable number of tandem repeats analysis. *Methods in Molecular Biology (Clifton, N.J.), 551*, 141-158.

Vilas-Boas, G. T., Peruca, A. P., & Arantes, O. M. (2007). Biology and taxonomy of *Bacillus cereus*, *Bacillus anthracis*, and *Bacillus thuringiensis*. *Canadian Journal of Microbiology, 53*(6), 673-687.

Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews, 51*(2), 221-271.

CHAPTER 2:

A comparative study between agarose and capillary gel electrophoresis using 31 multiple locus variable number tandem repeats analysis (MLVA) for *Bacillus anthracis sensu lato*

Abstract

Anthrax caused by *Bacillus anthracis* has received a great deal of attention in recent years due to its negative association with biological warfare. For this reason there has been concerted research in typing strains using multiple locus variable number tandem repeats (VNTRs) analyses (MLVA) for the rapid and reliable differentiation of anthrax strains. In this study, 113 *B. anthracis sensu lato* strains from southern Africa were analyzed using the MLVA31 panel that were separated using the affordable yet laborious agarose gel electrophoresis and the rapid but costly capillary electrophoresis. PCR products of 31 VNTR markers were measured using the two methods. They produced near identical results for 24 of the 31 VNTR loci. The 113 *B. anthracis sensu lato* strain cluster analysis generated from unweighted pair group method using arithmetric averages (UPGMA) were identical for the agarose – and capillary electrophoresis when using these 24 VNTR loci. However the cluster analysis of the 113 strains UPGMA using 31 VNTR's had more than 92% similarities for agarose electrophoresis and more than 78% similarity for capillary electrophoresis. The resolution using agarose gel electrophoresis does not allow the accurate separation of 6 VNTR loci with tandem repeat consisting of 6 bp or less, therefore this technique is not sufficient to characterise *B. anthracis* strains using 31 MLVA panel. However agarose gel electrophoresis using only a 25 MLVA panel (excluding small tandem repeat VNTR loci) is sufficient to type *B. anthracis* strains for the purpose of epidemiological study and is the most cost effective and appropriate technique for the average African / developing country laboratory. Although separation of the 31 MLVA panel using capillary electrophoresis is not cost effective, it is a rapid (since loci are amplified using multiplex PCR) and accurate for *B. antrhacis* typing method, provided a 50cm or longer capillary is employed.

2.1 Introduction

Anthrax has received a lot of attention in recent years, most commonly due to its association with biological warfare (Hoffmaster *et al.,* 2008). Due to this negative view of *Bacillus anthracis*, the causal agent of anthrax, there has been a concerted drive in the research for the development of fast and reliable DNA fingerprinting methods to trace intentionally released strains back to their origin (Keim *et al.,* 2000; Takahashi *et al.,* 2004; Van Ert *et al.,* 2007). MLVA is one such method and is in fact currently the technique of choice because of its reproducibility. It is a PCR-based technique with few steps and adaptable for automation that relies on the polymorphism of tandem repeated DNA sequences (VNTR) that is known as micro- or minisattelites for discriminations among strains (Keim *et al.,* 2000). MLVA uses PCR to amplify VNTR loci that differ in length due to variation in the number of times that a VNTR sequence repeats. The PCR fragment size is determined using agarose or capillary electrophoresis and the copy numbers of the VNTR unit is calculated by subtracting the size of the flanking regions from the amplicon size (http://mlva.u-psud.fr/MLVAnet). MLVA is therefore used to differentiate anthrax strains and to evaluate the diversity of *B. anthracis* from different geographical areas (Keim *et al.,* 2000; 2004; Van Ert *et al.,* 2007).

The development and use of an MLVA marker to differentiate *B. anthracis* strains started when Andersen *et al.* (1996) identified the VrrA locus that differentiated *B. anthracis* strains due to polymorphism of a VNTR. Keim *et al.* (2000) identified 8 VNTR loci, which included the VrrA marker, and were amplified in multiplexed PCR followed by automated capillary electrophoresis. A further 7 markers were included by Van Ert *et al.* (2007), which gave rise to the MLVA15 panel for automated electrophoresis. Similarly, Lista *et al.* (2006) had improved on the MLVA8 set with 17 chromosomal markers, which made up the MLVA25 panel used for agarose gel electrophoresis. The combined panels (MLVA15 and MLVA25) add up to 31 VNTR loci due to identical target regions in both the 15 and 25 VNTR panels. This 31 VNTR panel is currently used at the University of Hohenheim in Germany to optimize common protocols with regards to laboratory platforms and reagents for automated typing (Beyer & Turnbull, 2009).

The analysis of VNTR using agarose gel electrophoresis is considered feasible for the average African / developing country laboratory since most laboratories have this basic equipment. Although it is a simple and common technique with many applications, it requires long migrations for accurate analyses of VNTR copy numbers during MLVA typing (Lista *et al.,* 2006). It is difficult/ nearly impossible to multiplex VNTR loci using agarose electrophoresis due to overlapping amplicon sizes for the different allele ranges (Le Flèche *et al.,* 2001).

Capillary electrophoresis is the adaptation of traditional gel electrophoresis using a capillary with a replaceable matrix which allows amplicons with similar charge-to-mass ratios to be resolved by size at a much higher rate than conventional gel electrophoresis (Buel *et al.,* 2001). Fluorescence can be used to detect amplicons and multiplex PCR is possible since different fluorophores are available to label different VNTR markers (Lindstedt, 2005). New fluorophores are also bound to emerge to even further advance multiplexing capabilities as fluorophore chemistry improves (Lindstedt, 2005). There is some argument over whether agarose gel electrophoresis can provide sufficient resolution when compared to the other techniques (Yokoyama *et al.*, 2006; Lista *et al.*, 2006; Jenkins *et al.*, 2010). In this study, we investigated the use of 31 VNTR markers to type *B. anthracis* isolates from southern Africa using both the capillary and agarose electrophoresis methods to determine the comparative value of each.

2.2 Materials and Methods

2.2.1 Sample History

Isolates of 113 *B. anthracis sensu lato* from various parts of southern Africa were used in this study (Appendix Table 1), which includes the *B. anthracis* $34F_2$ Sterne attenuated vaccine strain (BA#113) that serves as a reference strain during typing. The samples were cultured at 37°C on blood agar medium from isolates stored in 10 % glycerol at -80°C. Most of the cultures (BA#001-091) were collected by Dr Valerius de Vos in the Kruger National Park (KNP) in South Africa over a period of 4 decades during anthrax outbreaks and surveillance in South Africa. The 2 Zambian samples (Ba#46-47) were isolated by Victor Siamudala and another 21 *B. anthracis sensu lato* DNA (BA#092-112) were obtained from

the bacteriology section of Agriculture Research Council-Onderstepoort Veterinary Institute (ARC-OVI), Onderstepoort, South Africa.

2.2.2 DNA Isolation

DNA was isolated using the Qiagen Gram-positive bacteria protocol of the Qiagen QIAmp DNA mini kit (Qiagen) and Zymo Research Bacterial DNA kit (Zymo Research Cooperation) using BashingBeads™ according to the manufacturer's instructions. With the aid of an inoculation loop, several colonies were removed from the culture plate and DNA was extracted using the Qiagen QIAmp® DNA Mini kit. The DNA was concentrated in 100 µl of Buffer AE instead of 200 µl. The DNA elute was further diluted to 1:100 (~1 ng/µl) with AE Buffer and stored at -20°C. With the Zymo Research fungal/bacterial DNA kit (Zymo Research Cooperation), approximately 10^9 bacterial colonies were suspended in 200 µl of PBS (phosphate buffered saline) using an inoculation loop and DNA was extracted according to the manufacturer's instructions. DNA was suspended in 100 µl of DNA elution buffer and stored at -20°C at a concentration of approximately 1 ng/µl and used undiluted.

2.2.3 VNTR amplification

Some of the isolates (n=38) were characterised using the MLVA25 panel (Le Fléche *et al.,* 2001; Lista *et al.,* 2006) at the Institute of Genetics and Microbiology, University Paris-South Orsay, France in collaboration with Gilles Vergnaud and Christine Pourcel to facilitate knowledge transfer regarding MLVA and for inter-laboratory comparisons.

2.2.3.1. Monoplex PCR for Discrimination of Tandem Repeats on Agarose Gel

VNTR markers targeting 31 loci (Table 2.1) were amplified using 4 ng DNA in a final volume of 15 µl containing 1x PCR ExSel reaction buffer (JMR Holdings), 2 mM $MgSO_4$, 0.2 mM dNTP, 20 pM primers (Table 2.1) and 0.125 U High Fidelity *Taq* polymerase (JMR Holdings). Table 2.1 indicates PCR conditions used for each VNTR primer. PCR condition 1 consisted of 95°C initial enzyme activation for 5 minutes followed by 35 cycle of 95°C denaturation for 20 seconds, 60°C annealing for 30 seconds and 72°C elongation for 5 minutes for primers amplifying products smaller than 600 bp and 7 minutes for primers amplifying products

larger than 600 bp (Table 2.2). PCR conditions 2 consisted of 95°C initial enzyme activation for 5 minutes, followed by 35 cycles of 95°C denaturation for 20 seconds, 56°C annealing for 30 seconds and 72°C elongation for 5 minutes.

2.2.3.2 Multiplex PCR for Discrimination of Tandem Repeats using Capillary Electrophoresis

The 31 VNTR markers with fluorescently labelled forward primer, were multiplexed into 7 reactions (Table 2.1) containing 2 ng DNA in a final volume of 15 µl with 1x reaction buffer, 2 mM $MgCl_2$, 0.2 mM dNTP and the appropriate concentration of oligonucleotide (where the cumulative concentration of all oligonucleotides in a multiplex is 20 pM) and 0.125 U Promega Go*Taq* Hot Start Polymerase (Whitehead Scientific). The PCR condition was as follows: 96°C initial enzyme activation for 5 minutes, followed by 30 cycles of 96°C denaturation for 20 seconds, 58°C annealing for 30 seconds and 72°C elongation for 7 minutes.

2.2.3.3 Sequencing

Bams 15 PCR products of isolates Ba#113 (Sterne $34F_2$ reference strain), Ba#035 and Ba#094 were purified using a NucleoSpin® Extraction II kit (Macherey-Nagel GmbH) and bi-directionally sequenced using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems) and 3130XL Genetic Analyzer (Applied Biosystems). Sequence fragments were assembled using Staden package version 1.6.0 (Staden, 1996) and a homology search of each sequence was done against whole genome sequences of *Bacillus anthracis* from the National Centre for Biotechnology Information (NCBI; www.ncbi.nlm.nih.gov).

2.2.4 Electrophoresis

2.2.4.1 Agarose Gel Electrophoresis

The choice of agarose and its concentration were determined for each locus according to the length of the repeat unit (Table 2.1). PCR products were loaded in sets of 7 samples

between Hyperladder II (Bioline) or O'Generuler 100 bp plus ladder (Fermentas), with the first well containing the $34F_2$ reference strain PCR product as indicated by Vergnaud & Pourcel (2009). Agarose gels were run at 200 V for 3 hours in 1x TBE. Agarose gels were stained using ethidium bromide and visualized under UV light. The gel image was saved as (TIFF files) and imported into Bionumerics version 6.0 (Applied Maths) that enables amplicon size calculations.

2.2.4.2 Capillary Electrophoresis

The MLVA PCR products were separated by the use of 31 cm capillary on an ABI 3130 Genetic Analyzer (Applied Biosystems). The allele sizes of multiplex PCR products were determined by comparison to the internal standard 1200LIZ (Applied Biosystems) using STRand Analysis Software version 2.3.106 *(*http://www.vgl.ucdavis.edu/STRand). *Bacillus anthracis* 34F2 Sterne strain (BA#113) was included as a reference strain to help with calibrating of allele size determination between separation methods. Minor corrections were made since every genetic analyzer used for capillary electrophoresis produces a different amplicon read (Appendix Table 4). The copy number assignments for calibration of capillary reads are therefore based on Appendix Table 2 and 3 in order to make the capillary and agarose data comparable.

Table 2.1 Information regarding VNTR markers used in this study.

VNTR locus *	Forward primer 5'-3'	Reverse primer 5'-3'	Size of repeat unit (bp)	Range of fragment sizes (bp)	PCR conditions	Multiplex PCR reaction (M) **	Fluorescent Label **	Percentage (w/v) agarose gel specification ***
BAMS21	TGTAGTGCCAGATTTGTCTTCTGTA	CAAATTTTGAGATGGGAGTTTTACT	45	541 - 766	1	M5	PET	2%
BAMS51	ATTTCCTGAAGCAGGTTGTGTT	TGCATCTAACAATGCAGAACAA	45	358 - 538	1	M1	PET	
BAMS23	CCTGTTGCTCCTAGTGATTTCTTAC	CGGTCTGTCTCTATTATTCAGTGGT	42	399 - 693	1	M4	VIC	
BAMS24	CGTCACGTACCATTTAATGTTGTTA	CTTCTACTTCCGTACTTGAAATTGG	42	469 -511	1	M7	NED	
BAMS5	ATTATTAGCAGGGGCCTCTCCTGCATTACC	GCAGGAAGAACAAAAGAAACTAGAAGAGCA	39	229 - 424	1	M2	VIC	
BAMS34	TGTGCTAAATCATCTTGCTTGG	CAGCAAAATCAATCGAATCAAA	39	230 - 581	1	M6	PET	
BAMS44	GCACTTGAATATTTGGCGGTAT	GCGAATTAATTGCTCCTCAAAT	39	183 - 573	1	M1	VIC	
BAMS22	ACCGTTAATTCACGTTTAGCAGA	ATCAAAAATTCTTGGCAGACTGA	36	519 - 1041	1	M1	FAM	
BAMS28	TATTAAACCAGGCGTTACTTACAGC	CTCTGTTGTAACAAAATTTCCGTCT	24	373 - 505	1	M4	NED	3%
BAMS1	AGTTCAAGCGCCAGAAGGTTATGAGTTATC	GTTGAGCATGAGAGGTACCTTGTCCTTTTT	21	296 - 611	1	M3	PET	
vrrC2[a]	GTCTTTCCATTAATCGCGCTCTATC	CCAGAAGAAGTGGAACCTGTAGCAC	18	528 - 604	1	M4	PET	
BAMS3	TCCTCCCTGAGAACTGCTATCACCTTTAAC	GCAGCAACAGAAAACTTCTCTCCAATAACA	15	429 - 654	1	M1	NED	
BAMS25	TGAAAGATCTTGAAAAACAAGCATT	CCGAATACGTAAGAAATAAATCCAC	15	376 - 391	1	M5	VIC	
vrrA[a]	GCGCGTTTCGTTTGATTCATAC	CACAACTACCACCGATGGCACA	12	289 - 338	1	M5	FAM	
BAMS53	CATATTTTCACCTTAATTTTGGAAG	GAGGTGTGTTAGGTGGGCTTAC	12	322 - 346	1	M4	VIC	
VNTR23	GTAATACGTATGGTTCATTCCC	TTTAGAAACGTTATCACGCTTA	12	170 - 208	1	M2	NED	
vrrB1[a]	GATGAGTTTGATAAAGAATAGCCTGTG	ATAGGTGGTTTTCCGCAAGTTATTC	9	184 - 292	1	M4	FAM	
vrrB2[a]	CCCAAGGTGAAGATTGTTGTTGA	CACAGGCTATTCTTTATCAAACTCATC	9	135 -198	1	M2	NED	
vrrC1[a]	CATTTCCTCAAGTGCTACAGGTTC	GAAGCAAGAAAGTGATGTAGTGGAC	9	364 - 688	1	M3	VIC	
BAMS13	CTAGTGCATTTGACCCTAATCTTGT	AATTGAGAAATTGCTGTACCAAACT	9	337 - 868	1	M7	FAM	
BAMS15	GTGTACATGTTGATTCATGCTGTTT	GTATTTCCCCCAGATACAGTAATCC	9	409 - 643	1	M2	FAM	
BAMS30	CAGAAAATATTGGACCTACCTTCC	AGCTAATCACCTACAACACCTGGTA	9	268 - 925	1	M6	FAM	
BAMS31	GGAGTACTGTTTGTTGAATGTTGTTT	GCTGTATTTATCGAGCTTCAAAATCT	9	331 - 1087	1	M3	FAM	
VNTR 16[b]	GAATAATAAGGGTTCTCATGGTAT	CTCTTGAAAATATAAAACGCA	8	137 - 346	2	M2	PET	3% MS
VNTR17[b]	GATCGTACAACAGCAATTATCAT	TAGGTAAACAAATTTTCGTAATC	8	366 - 453	2	M3	NED	
VNTR35[b]	GTCCTGAAATAAATGCTGAAT	AAATAATATGTTCCTTTTGCTG	6	102 - 126	1	M6	NED	
CG3[a]	AAATAATATGTTCCTTTTGCTG	TGTCGTTTTACTTCTCTCTCCAATAC	5	153 - 160	1	M1	FAM	
pXO1[a]	TCTAGAATTAGTTGCTTCATAATGGC	CAATTTATTAACGATCAGATTAAGTTCA	3	120 - 144	1	M1	PET	
VNTR19[b]	GAAATATTTTATTAAACATGCTTTCCATCC	GTGATGAAATCGGACAAGTTAGGAG	3	91 - 134	2	M5	NED	
pXO2[a]	GTGTGATGAACTCCGACGACA	TCATCCTCTTTTAAGTCTTGGGT	2	133 - 155	1	M6	VIC	
VNTR 12[b]	GCATATAATTGCACCTCATCTAG	CGTACGAAGTAGAAGTCATTAA	2	106 -120	2	M3	NED	

* Reference to VNTR markers from (a) Keim *et al.* (2000); (b) Van Ert *et al.* (2007) and rest from Le Fleche *et al.* (2001).

** Multiplex PCR reaction with labelled fluorophore used to mark forward primer in each multiplex PCR reaction separated using capillary electrophoresis

***Monoplex PCR separated using agarose electrophoresis with agarose gel percentage indicated. Small tandem repeat unit loci were separated using 3% molecular screening (MS)agarose, which improves resolution of 500bp and less DNA fragments .

2.2.5 MLVA data analysis

A database was created for *B. anthracis* using BioNumerics version 6.0 software for analysis of copy numbers of the VNTR units (rounded to the first decimal place) using agarose gel images. Allele sizes for the MLVA25 panel were converted into copy numbers using Appendix Table 2 as reported by Le Fleche *et al.* (2001), Lista *et al.* (2006) and MLVA web-service (http://mlva.u-psud.fr) at the Institute of Genetics and Microbiology, University Paris-South Orsay, France. The remaining 6 VNTR markers were converted into copy numbers (Appendix Table 3) by blasting primer sequences against the Sterne strain in GPMS (Genomes, Polymorphism and Minisatellites) website (http://minisatellites.u-psud.fr) to determine expected fragment size (Denoeud & Vergnaud, 2004; van Ert *et al.,* 2007). Genetic distances were calculated using similarity coefficient and cluster analysis were done by UPGMA with BioNumerics 6.0 software.

The copy number was calculated for each marker according to the algorithm described by Vergnaud & Poucel (2009) using allele sizes determined with capillary electrophoresis. Genetic distances were calculated using similarity coefficient and cluster analyses were done by UPGMA with BioNumerics 6.0 software.

2.3 Results and Discussion

The first multilocus VNTR analysis used 8 markers, MLVA8, to type *B. anthracis* strains using 6 chromosomal and 2 plasmid marker loci (Keim *et al.,* 2000). The MLVA8 panel was not sufficient since most French and Italian strains were assigned to 2 genotypes (Keim *et al.,* 2000; Fouet *et al.,* 2002; Fasenella *et al.,* 2005). Therefore, additional markers were developed to produce the MLVA15 and MLVA25 panels. In this study we combined the MLVA 15 and 25 panels that share 9 VNTR loci within the MLVA31 panel, to determine their resolution using agarose and capillary electrophoresis on 113 *B. anthracis sensu lato* isolates. These 113 *B. anthracis sensu lato* isolates were mostly collected from the anthrax endemic regions in the Northern Cape Province and KNP in South Africa. There are two isolates each from both the Luangwa Valley in Zambia (Ba#46 and Ba#47); from a dairy in Rondebosch (Ba#44 and Ba#45) in the Western Cape Province, South Africa and from

Ngamiland, Botswana (Ba#94 and Ba#103). A single isolate (Ba#104) was obtained from the Kingdom of Lesotho as well.

Cluster analyses of 31 MLVA data of the 113 *B. anthracis sensu lato* isolates separated by agarose electrophoresis indicated more than 92% similarity (Appendix Figure 1), whereas capillary electrophoresis had only more than 78% similarity (Appendix Figure 2). Our results therefore indicated that agarose electrophoresis was most suitable for cluster analysis of the 31 VNTR data of *B. anthracis* strains. However this might be due to inaccurate variation of small tandem repeat units (< 6 bp) when separated on MS agarose that is discussed later. With the agarose and capillary electrophoresis methods, cluster analysis of 24 of the 31 VNTR loci produced almost identical dendogram trees based on the copy numbers calculated using UPGMA for both methods (Appendix Figure 3). The 7 VNTR loci that caused discrepancies included VNTR 12 (2 bp tandem repeat unit); VNTR19 (3 bp repeat unit); CG3 (5 bp repeat unit); pXO1 (3 bp repeat unit); pXO2 (2 bp repeat unit); VNTR 35 (6 bp repeat unit) and Bams 13 using the electrophoresis methods.

2.3.1 Capillary Gel Electrophoresis

With capillary analysis the 31 VNTR markers were multiplexed whereas with agarose electrophoresis the PCR was monoplexed. Multiplex PCR was possible as all VNTR markers can be labelled with different fluorophores. These fluorescent markers are necessary, as capillary electrophoresis works through measuring of different wavelengths emitted by fluorophore labelled molecules after laser illumination as they move through the capillary at ~500V/cm (Lindstedt, 2005). The 31 VNTR markers used in this study were multiplexed into 7 reactions so that overlapping amplicon ranges could be distinguished through their different light emission. These are perceived as colours in real time, through software, which display the readings in the form of an electropherogram (Beckman Coulter, 1991). All DNA sequencers are however, not consistent in the way that an amplicon size will be perceived, because this is dependent on the make and model, capillary length and diameter as well as the matrix within which the amplicons are moving (Buel *et al.,* 2001). It is recommended that an internal size marker be used to increase the accuracy with which the amplicons would be read and binned according to copy number or alternatively a reference

strain used (Lindstedt *et al.,* 2004; Lindstedt, 2005). The amplicon sizes cannot be directly compared to those of agarose gel electrophoresis because the 3-dimensional morphology, charge and fluorescent label all play a role in the rate at which the molecule travels within the capillary matrix (Beckman Coulter, 1991; also observed in Appendix Table 4). For this reason, the Sterne $34F_2$ reference strain VNTR loci were analysed first. The STRand® software database for *B. anthracis* was set up using the fsa files of this run and each reading frame set up according to the copy number values for Sterne (Appendix: Tables 2 and 3). This was to ensure that the agarose gel results and the capillary gel results would be comparable.

The most obvious benefit of using capillary gel electrophoresis is its ability to distinguish the difference between small repeat units, even as small as 2bp. This allows for the inclusion of larger VNTR marker panels targeting smaller amplicons. This will decrease run times to half an hour or even less and increase the throughput of samples. There is also the benefit of using less DNA and primer as it is a more sensitive method. A single multiplex PCR reaction required as little as 2ng of DNA and included up to 6 VNTR markers without same fluorophore overlap (Appendix Figure 4).

Certain challenges using the capillary method became evident. The use of the 31 cm capillary proved to be limiting in this study for the Bams 13 VNTR locus, the amplicons were too large to be read accurately (Appendix Figure 5). The LIZ 1200bp size standard would have proven sufficient for typing all loci had a 50 cm capillary been employed. Running the capillary for a prolonged period did not improve the results for Bams 13 although, other large amplicons (e.g. Bams 30 and 31) were more clearly visible after an extended run. It then became necessary for Bams 13 to be excluded from the cluster analysis in comparing the 2 electrophoresis methods. According to Buel *et al.* (2001), resolution is dependent on the size standard's peak width and peak spacing; therefore, the number of bases in a DNA fragment is inversely proportionate to the resolution. The resulting peak overlaps can sometimes also result in false reads. The dissipating accuracy of resolution can be observed in Appendix Figure 6. This method is highly dependent on the use of a standard reference strain or means of interpreting the results, which is uniform across all makes and models of genetic analyzers. This is the only way in which interlaboratory comparisons can be performed with ease (Hyytia-Trees *et al.,* 2010).

2.3.2 Agarose Gel Electrophoresis

Agarose gel electrophoresis is a commonly used method in the average laboratory. For use in MLVA, the only possible added expense would be the purchase of a 30cm gel tank that will allow longer agarose gel migrations. This is necessary for accurate discrimination of copy number, especially when discerning amplicons with small repeat units. It is also essential to include a reference strain that should be interspersed within the samples separated on agarose gel for increased accuracy. This prevents the incorrect assignment of copy number due to gel distortions. One disadvantage of this method is its limitation in accurately typing amplicons with small (≤ 6bp) repeat units which has already been discussed (Le Flèche *et al.*, 2001; Lista *et al*, 2006).

An example of the inability to accurately determine copy numbers at VNTR loci with small repeat units (≤6 bp) is shown in Figure 2.1. Molecular screening (MS) agarose was used; where at 3% (w/v) is reported to give a DNA fragment/amplicon size with an equivalent resolution to gels made with polyacrylamide at a concentration of 8%. However, the capillary electrophoresis (7% polyacrylamide) gave less variation than the MS agarose. This can be attributed to the 3 dimensional morphology of the amplicon as it moves through the different mediums. The MS agarose is run for a longer time at lower volts than capillary electrophoresis. The band separated by MS agarose and visualized with UV contains a large surface area that needs to be pin pointed on the gel image to determine the amplicon size. This was achieved with the aid of BioNumerics software. The middle of the band was used to pin point the band size, which resulted in the large variation / inaccuracy obtained with MS agarose compared to capillary electrophoresis (Table 2.2). VNTR loci with 3 bp and less do not produce accurate copy number results using agarose electrophoresis and should therefore be excluded when using agarose gel electrophoresis. However, these small repeat unit VNTR loci include the plasmid VNTR loci, pXO1 and pXO2. Therefore, excluding the plasmid loci will create bias since pathogenicity factor occurs on plasmids. Therefore, we suggest that the presence or absence (present (1) or absent (0 or -2) of one or both plasmids still needs to be indicated and analysed as part of the MLVA data.

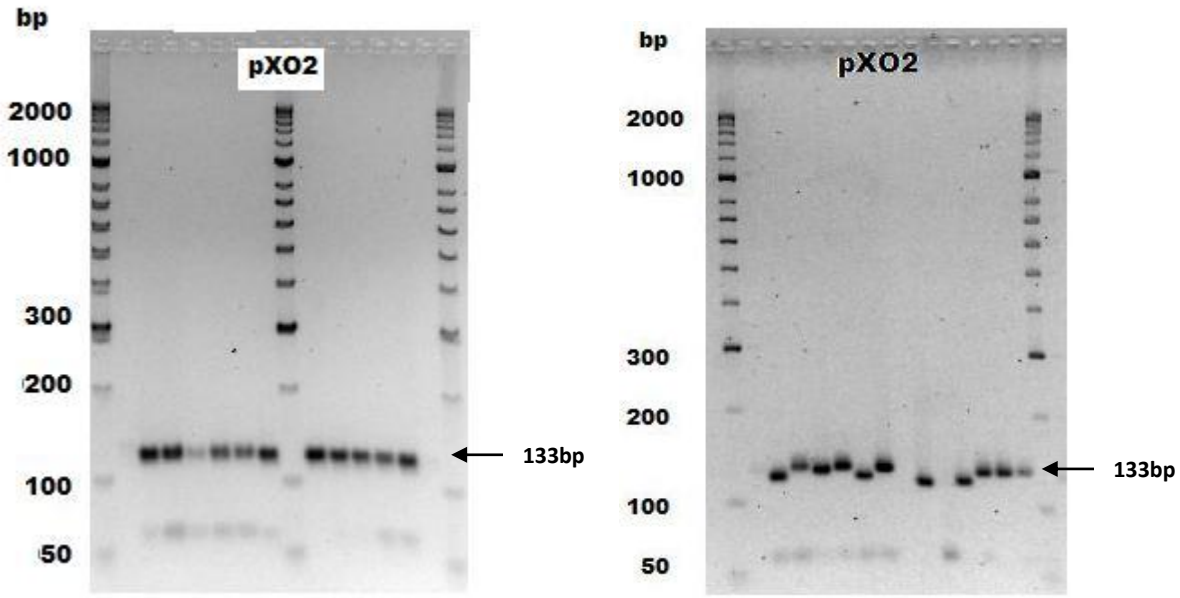


Figure 2.1: Separations of VNTR pXO2 locus on 3% multi-purpose agarose (left) and 3% MS agarose (right).

Table 2.2: Example of copy numbers determined for 11 *Bacillus anthracis* isolates using agarose and capillary electrophoresis at VNTR loci with small tandem repeats to highlight the inability to accurately determine the copy number at these VNTR loci with less than 6 base pair tandem repeat units (taken from Appendix Table 5).

Sample number	VNTR12(2 bp repeat unit)		VNTR19 (3 bp repeat unit)		CG3 (5 bp repeat unit)		pXO1 (3 bp repeat unit)		pXO2 (2 bp repeat unit)	
	Agarose	Capillary	Agarose	Capillary	Agarose	Capillary	Agarose	Capillary	Agarose	Capillary
Ba#020	7	5	5	6	2	2	6	4	10	12
Ba#021	9	5	7	7	2	1	6	4	5	4
Ba#022	8	5	6	6	2	1	7	4	5	4
Ba#023	8	5	7	7	3	2	7	4	6	5
Ba#024	8	5	9	12	3	2	7	2	6	5
Ba#025	8	5	6	6	3	2	7	2	6	5
Ba#026	9	5	7	7	3	1	7	2	7	10
Ba#027	9	5	7	7	3	2	7	4	6	5
Ba#028	9	5	4	4	3	1	7	4	6	1
Ba#029	7	5	4	4	2	2	7	4	1	1
Ba#030	7	5	4	4	3	1	6	4	1	2

The MS agarose compared to muti-purpose agarose (Figure 2.1) has its disadvantages as discussed above, but MS agarose still improved the resolution to allow accurate

determination of small repeat units that range between 5-8 bp. Note that in the multi-purpose agarose in Figure 2.1, all the samples would be assigned the same copy number whereas the MS agarose allows discrimination between the isolates. MS agarose is 0.7 times more expensive than the multi-purpose agarose. Despite this additional cost, the agarose electrophoresis technique is still a third of the cost of the capillary technique but time consuming as PCR needs to be done in duplicate to eliminate human error. Copy numbers can easily be determined using agarose gel electrophoresis (Figure 2.1) using the conversion table to determine copy numbers (Appendix: Table 2 and 3), which is a great advantage in laboratories that do not have access to expensive analysis software like Bionumerics.

Another discrepancy which became evident is that the Sterne reference strain did not always type according to the published fingerprint. This has been the cause for differences in copy number assignment between the capillary and agarose methods like with Bams 15 (Table 2.3). On closer examination of the copy number produced for Bam 15, the agarose and capillary electrophoresis methods show a similar profile (Table 2.3). There is a consistent difference of 7 repeat units in copy number assignations between the agarose and capillary electrophoresis methods making this a reliable marker.

Table 2.3: Example of similar profiles of copy numbers for Bams 15 locus for 20 *Bacillus anthracis* isolates using agarose and capillary electrophoresis (taken from Appendix Table 5).

Sample ID	Bams15		Sample ID	Bams15		Sample ID	Bams15	
	Agarose	Capillary		Agarose	Capillary		Agarose	Capillary
Ba#001	24	31	Ba#011	43	50	Ba#021	43	50
Ba#002	41	48	Ba#012	43	50	Ba#022	45	52
Ba#003	41	48	Ba#013	43	50	Ba#023	45	52
Ba#004	41	48	Ba#014	43	50	Ba#024	43	50
Ba#005	41	48	Ba#015	43	50	Ba#025	43	50
Ba#006	24	31	Ba#016	43	50	Ba#026	43	50
Ba#007	41	48	Ba#017	46	53	Ba#027	43	50
Ba#008	44	51	Ba#018	46	53	Ba#028	44	51
Ba#009	44	51	Ba#019	43	50	Ba#029	48	55
Ba#010	43	50	Ba#020	43	50	Ba#030	47	54

As per (Appendix: Table 2) and observed on CEQ Beckman (unpublished data), Bams 15 should amplify a region 418 bp for Sterne strain, which translates to 24 copy units, but a PCR product of approximately 600 bp was observed on agarose gel which translates to 45 copy units (data not shown). For this reason, the Sterne reference strain and 2 random samples targeting VNTR locus Bams 15 were sequenced to determine whether the discrepancy between the expected and observed results was valid or not. Alignment of the forward and reverse sequences produced a sequence of 609 bp that was edited to 591 bp reliable sequence data (Figure 2.2). Figure 2.2 illustrates the sequence and BLAST results indicating 100% homology between our reference strain sequence and Sterne strain Genbank accession no. AE017225.1. The copy number unit of Sterne reference strain must be corrected from 24 to 45 copy numbers in conversion table (Appendix Table 2) to allow future accurate use of data. However, the use of the Sterne reference strain in the agarose gel technique does not influence copy numbers determination of other isolates. It is merely a control to ensure that PCR conditions and gel homogeneity are optimal to ensure as much accuracy in copy number assignments as is humanly possible. The sequence results of other 2 samples verified the calculated copy number from PCR product size to be accurate when comparing it to sequence (Appendix: Figures 7 & 8).

Sterne_Bams 15_FASTA

atgttgattcatgctgtttTagttgcgatggaacagttcctaagctgggtccaacaggtccgacgggagcaacgggagcaacaggagcaacaggagtaacg
ggagtaacgggagtaacaggagcgacaggaataacgggagcgacaggaataacaggagcaacaggaataacgggagtaacgggtccaacgggagcga
caggaataacaggagcaacgggagcaacgggtccaacgggagcaacaggtccgacgggagcgacaggaataacgggagcaacaggagcaacaggagc
aacaggagcaacgggagtaacgggagtaacaggagcgacaggaataacaggagcaacaggaataacgggagtaacgggtccaacgggagcgacagga
ataacaggagcaacgggtccaacgggagcgacaggaataacgggagcgacaggtccaacgggagtcacaggtacaagtattacggcgacgtatgcattt
gcaaataatacgtcagggacagcgatatcggtacttcttggtggaacaaatgtcccgcttccaaataatcaaaatatcggtccgggattactgtatctgggg
ggaaata

Sequences highlighted in blue are the regions where the forward and reverse primers bind

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
AE017225.1	Bacillus anthracis str. Sterne, complete genome	1092	2327	99%	0.0	100%	
CP001215.1	Bacillus anthracis str. CDC 684, complete genome	1033	2570	99%	0.0	98%	
CP001598.1	Bacillus anthracis str. A0248, complete genome	568	2057	99%	2e-161	99%	
AE017334.2	Bacillus anthracis str. 'Ames Ancestor', complete genome	568	2057	99%	2e-161	99%	
AE016879.1	Bacillus anthracis str. Ames, complete genome	568	2057	99%	2e-161	99%	

>gb|AE017225.1| Bacillus anthracis str. Sterne, complete genome
Length=5228663

Sort alignments for this subject sequence by:
E value Score Percent identity
Query start position Subject start position

Features in this part of subject sequence:
conserved hypothetical protein, collagen triple helix rep...

Score = 1092 bits (591), Expect = 0.0
Identities = 591/591 (100%), Gaps = 0/591 (0%)
Strand=Plus/Minus

```
Query  2        TGATTCATGCTGTTTTAGTTGCGATGGAACAGTTCCTAAGCTGGGTCCAACAGGTCCGAC  61
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333474  TGATTCATGCTGTTTTAGTTGCGATGGAACAGTTCCTAAGCTGGGTCCAACAGGTCCGAC  4333415

Query  62       GGGAGCAACGGGAGCAACAGGAGCAACAGGAGTAACGGGAGTAACGGGAGTAACAGGAGC  121
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333414  GGGAGCAACGGGAGCAACAGGAGCAACAGGAGTAACGGGAGTAACGGGAGTAACAGGAGC  4333355

Query  122      GACAGGAATAACGGGAGCGACAGGAATAACAGGAGCAACAGGAATAACGGGAGTAACGGG  181
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333354  GACAGGAATAACGGGAGCGACAGGAATAACAGGAGCAACAGGAATAACGGGAGTAACGGG  4333295

Query  182      TCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGAGCAACGGGTCCAACGGGAGCAAC  241
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333294  TCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGAGCAACGGGTCCAACGGGAGCAAC  4333235

Query  242      AGGTCCGACGGGAGCGACAGGAATAACGGGAGCAACAGGAGCAACAGGAGCAACAGGAGC  301
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333234  AGGTCCGACGGGAGCGACAGGAATAACGGGAGCAACAGGAGCAACAGGAGCAACAGGAGC  4333175

Query  302      AACGGGAGTAACGGGAGTAACAGGAGCGACAGGAATAACAGGAGCAACAGGAATAACGGG  361
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333174  AACGGGAGTAACGGGAGTAACAGGAGCGACAGGAATAACAGGAGCAACAGGAATAACGGG  4333115

Query  362      AGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGTCCAACGGGAGCGAC  421
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333114  AGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGTCCAACGGGAGCGAC  4333055

Query  422      AGGAATAACGGGAGCGACAGGTCCAACGGGAGTCACAGGTACAAGTATTACGGCGACGTA  481
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333054  AGGAATAACGGGAGCGACAGGTCCAACGGGAGTCACAGGTACAAGTATTACGGCGACGTA  4332995

Query  482      TGCATTTGCAAATAATACGTCAGGGACAGCGATATCGGTACTTCTTGGTGGAACAAATGT  541
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4332994  TGCATTTGCAAATAATACGTCAGGGACAGCGATATCGGTACTTCTTGGTGGAACAAATGT  4332935

Query  542      CCCGCTTCCAAATAATCAAAATATCGGTCCGGGGATTACTGTATCTGGGGG  592
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4332934  CCCGCTTCCAAATAATCAAAATATCGGTCCGGGGATTACTGTATCTGGGGG  4332884
```

Figure 2.2: The 100% identity of Blast search between Sterne sequence (Ba#113 indicated as Query) and Sterne sequence AE017225.1 (indicated as Sbjct). The repeat unit (highlighted in green) of Bams 15 in isolates Ba#113 translates to 45 copy units. The sequence was edited and blasted in NCBI nucleotide database.

Lista *et al.,* 2006 reported similar differences at VNTR loci Bams 15, Bams 31, Bams 3 and Bams 30 when comparing correspondence between agarose and capillary gel. The observed differences for Bams 30, for which the offset in observed and expected sizes increased with allele size, were clarified after sequencing and it was determined that it was because the size of the repeat unit was incorrectly assigned initially. This has since been corrected. The cause for the differences in the others has not been elucidated in full as yet, except to note it as possibly part of the hypermutable behaviour of VNTR repeat units or due to the three dimensional morphology of the fragment as it moves through the different gel media (DeForce *et al.,* 1998; Lindstedt, 2005).

2.4 Conclusion

Agarose gel electrophoresis and capillary gel electrophoresis produce results which are identical for amplicons with repeat units between 9bp and 24bp. Both techniques have their own draw-backs at the extremes of the fragment length range. Capillary gel electrophoresis does indeed produce a more accurate resolution when using VNTR loci which target smaller amplicons. The capillary technique produces faster results and has a higher through-put, but access to the appropriate machinery, reagents and funding can be a limiting factor for most institutions in Africa. There is also a concern regarding lab training, software and technical support in the proper use of this technique, which could lead to false reads and improper data generation.

Agarose gel electrophoresis is the more feasible method and more laboratories are already equipped with the necessities. It is a time consuming and labour intensive method, but in the end, for this study, the agarose technique proved to be more robust in correctly assigning unit copy number results (as the amplicons are viewed independently from one another) and results are interlaboratory comparable without any adjustments. According to the various studies (Keim *et al*, 2000; Le Fleche *et al.,* 2001; Lista *et al.*, 2006) this is after all one of the primary reasons why MLVA typing is so sought after, because new isolates can be compared to isolates fingerprinted in the past without bias.

References

Andersen, G. L., Simchock, J. M., & Wilson, K. (1996). Identification of a region of genetic variability among *Bacillus anthracis* strains and related species. *Journal of Bacteriology, 178*(2)

Beckman Coulter (1991). Introduction to capillary electrophoresis handbook. http://web.campbell.edu/faculty/fetterman/PHSC%20410/Introduction%20to%20CE%20-%20CE%20Primer1.pdf

Beyer, W., & Turnbull, P. C. (2009). Anthrax in animals. *Molecular Aspects of Medicine, 30*(6), 481-489.

Buel, E., LaFountain, M., Schwartz, M., & Walkinshaw, M. (2001). Evaluation of capillary electrophoresis performance through resolution measurements. *Journal of Forensic Sciences, 46*(2), 341-345.

Deforce, D. L. D., Millecamps, R. E. M., Van Hoofstat, D., & Van den Eeckhout, E. G. (1998). Comparison of slab gel electrophoresis and capillary electrophoresis for the detection of the fluorescently labeled polymerase chain reaction products of short tandem repeat fragments. *Journal of Chromatography A, 806*(1), 149-155.

Denoeud, F., & Vergnaud, G. (2004). Identification of polymorphic tandem repeats by direct comparison of genome sequence from different bacterial strains: A web-based resource. *BMC Bioinformatics, 5*, 4.

Fasanella, A., Van Ert, M., Altamura, S. A., Garofolo, G., Buonavoglia, C., Leori, G., *et al*. (2005). Molecular diversity of *B*acillus anthracis in Italy. *Journal of Clinical Microbiology, 43*(7), 3398-3401

Fouet,A.; Smith,K.L.; Keys,C.; Vaissaire,J.; Le Doujet,C.; Levy,M.; Mock,M.; Keim,P. (2002). Diversity among French Bacillus anthracis isolates. Journal of Clinical Microbiology, 40, 12, 4732-4734

Hoffmaster, A. R., Novak, R. T., Marston, C. K., Gee, J. E., Helsel, L., Pruckler, J. M., *et al.* (2008). Genetic diversity of clinical isolates of *Bacillus cereus* using multilocus sequence typing. *BMC Microbiology, 8*, 191.

http://mlva.u-psud.fr/MLVAnet

http://www.vgl.ucdavis.edu/STRand

Hyytia-Trees, E., Lafon, P., Vauterin, P., & Ribot, E. M. (2010). Multilaboratory validation study of standardized multiple-locus variable-number tandem repeat analysis protocol for shiga toxin-producing *Escherichia coli* O157: A novel approach to normalize fragment size data between capillary electrophoresis platforms. *Foodborne Pathogens and Disease, 7*(2), 129-136.

Jenkins, A. O., Venter, E. H., Hutamo, K., & Godfroid, J. (2010). Comparison of the capillary and agarose electrophoresis based multiple locus VNTR (variable number of tandem repeats) analysis (MLVA) on *Mycobacterium bovis* isolates. *Veterinary Microbiology, 145*(1-2), 172-176.

Keim, P., Price, L. B., Klevytska, A. M., Smith, K. L., Schupp, J. M., Okinaka, R., *et al.* (2000). Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *Journal of Bacteriology, 182*(10), 2928-2936.

Keim, P., Van Ert, M. N., Pearson, T., Vogler, A. J., Huynh, L. Y., & Wagner, D. M. (2004). Anthrax molecular epidemiology and forensics: Using the appropriate marker for different evolutionary scales. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases, 4*(3), 205-213.

Le Fleche, P., Hauck, Y., Onteniente, L., Prieur, A., Denoeud, F., Ramisse, V., *et al.* (2001). A tandem repeats database for bacterial genomes: Application to the genotyping of yersinia pestis and Bacillus anthracis. *BMC Microbiology, 1*, 2.

Lindstedt, B. A. (2005). Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis, 26*(13), 2567-2582.

Lindstedt, B. A., Vardund, T., Aas, L., & Kapperud, G. (2004). Multiple-locus variable-number tandem-repeats analysis of *Salmonella enterica* subsp. *enterica* serovar *typhimurium* using PCR multiplexing and multicolor capillary electrophoresis. *Journal of Microbiological Methods, 59*(2), 163-172.

Lista, F., Faggioni, G., Valjevac, S., Ciammaruconi, A., Vaissaire, J., le Doujet, C., *et al.* (2006). Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. *BMC Microbiology, 6*, 33.

Staden, R. (1996). The staden sequence analysis package. *Molecular Biotechnology, 5*(3), 233-241.

Takahashi, H., Keim, P., Kaufmann, A. F., Keys, C., Smith, K. L., Taniguchi, K., *et al.* (2004). *Bacillus anthracis* incident, Kameido, Tokyo, 1993. *Emerging Infectious Diseases, 10*(1), 117-120.

Van Ert, M. N., Easterday, W. R., Simonson, T. S., U'Ren, J. M., Pearson, T., Kenefic, L. J., *et al.* (2007). Strain-specific single-nucleotide polymorphism assays for the *Bacillus anthracis* ames strain. *Journal of Clinical Microbiology, 45*(1), 47-53.

Vergnaud, G., & Pourcel, C. (2009). Multiple locus variable number of tandem repeats analysis. *Methods in Molecular Biology (Clifton, N.J.), 551*, 141-158.

www.ncbi.nlm.nih.gov

Yokoyama, E., Kishida, K., Uchimura, M., & Ichinohe, S. (2006). Comparison between agarose gel electrophoresis and capillary electrophoresis for variable numbers of tandem repeat typing of *Mycobacterium tuberculosis*. *Journal of Microbiological Methods, 65*(3), 425-431.

Chapter 3

Anthrax: to be or not to be, that is the question

Abstract

Bacillus anthracis together with *B. cereus* and *B. thuringiensis* form *B. cereus sensu stricto* within the *B. cereus* group. There has been some debate over whether to classify the 3 members of the *B. cereus sensu stricto* as distinct species or as subspecies of *B. cereus group. Bacillus cereus sensu stricto* share a highly conserved chromosome, but differ in pathogenicity that is predominantly plasmid–encoded. *Bacillus antrhacis sensu lato/B. anthracis*–like organisms have recently been found to cause anthrax-like disease in people, chimpanzees or gorillas. Multiple-locus variable number tandem repeats analysis (MLVA) is known to reveal the genetic relationships within *B. anthracis*, which is important to differentiate the alleles that are present amongst the *B. cereus sensu stricto* group. Therefore, the aim of this study was to determine if MLVA could differentiate between *B. anthracis* and *B. anthracis*-like organisms. The 31MLVA panel clustered *B. anthracis* isolates from the Northern Cape into 3 sub-clusters with most Northern Cape isolates grouping in the main subcluster, a border territory sub-cluster and the smallest sub-clade consisting of 2 isolates. A minimal spanning tree indicated that these 2 isolates group closer to a *B. cereus* strain than the rest of the Northern Cape isolates. Microbiology tests indicated that one isolate was non-motile, non-haemolytic, penicillin sensitive and phage sensitive and therefore *B. anthracis;* while the other isolate was non-motile, non-haemolytic, penicillin resistant and gamma-phage resistant and therefore classified as non-anthrax. However, these isolates were from Lesotho and an unknown geographical region indicating MLVA alone may not be sufficient in resolving the species debate, but is an effective tool in determining genetic and geographical distance and identifying isolates with anomalies that require further study.

3.1 Introduction

The *Bacillus cereus* group is the most homogenous group of the genus *Bacillus* and its members include *B. cereus*, *B. anthracis, B. mycoides, B. pseudomycoides, B. thuringiensis* and *B. weihenstephanensis* (Kaneko *et al.,* 1978; Helgason *et al.,* 2000a). The *B. cereus* group is closely related in phenotype and genotype despite their varied aetiology (Daffonchio *et al.,* 2000). They are routinely classified according to microbiological techniques as aerobic, rod shaped and gram-positive bacteria commonly found in soil (Vilas-Boas *et al.,* 2007).

Bacillus cereus sensu stricto consist of three pathogenic spore forming species, namely *B. cereus, B. anthracis* and *B. thuringiensis,* which share a highly conserved chromosome but differ in pathogenicity that are mostly plasmid–encoded. *Bacillus cereus* is a food borne pathogen whereas *B. thuringiensis* is an insect pathogen with plasmid-encoding insecticidal crystal proteins and *B. anthracis,* the causal agent of anthrax, has plasmid-encoding toxin and capsule proteins. *Bacillus cereus sensu stricto* have other virulent factors encoded on the chromosome; like *B. cereus* that is motile via peritrichous flagella, composed of frilly edged colonies and displays haemolytic activity on sheep blood agar. *Bacillus thuringiensis* similarly presents motility, haemolytic activity and penicillin sensitivity (Vilas-Bôas *et al.,* 2007). In comparison, *B. anthracis* is characteristically non-motile, displays capsule production, lacks haemolytic activity and is sensitive to penicillin when grown on blood agar (Klee *et al.,* 2006a; Vilas-Boas *et al.,* 2007).

Diagnosis of *B. anthracis* from clinical samples has been considered to be quite straightforward, based on clinical symptoms together with routine microbiological techniques (Klee *et al.,* 2006a). However, in rare cases some discrepancies have been observed where the microbiological diagnosis and the clinical symptoms did not correspond and resulted in uncertainty of a definitive diagnosis for the causative organism (Lalitha & Thomas, 1997; Liang & Yu, 1999; Klee *et al.,* 2006a; Hoffmaster *et al.,* 2006; Leendertz *et al.,* 2006; Marston *et al.,* 2006; Klee *et al.,* 2010). Plasmid cured *B. anthracis* isolates are difficult to identify through standard microbiological methods as they tend to be phage resistant and in other rare cases *B. cereus* isolates have been found to be phage sensitive and non-haemolytic. Microbiological and molecular methods were used to elucidate these discrepancies (Turnbull *et al.,* 1999; Marston *et al.,* 2006; Klee *et al.,* 2010).

Bacillus anthracis is a highly monophyletic clade as determined with molecular techniques including sequencing, amplified fragment length polymorphisms (AFLP), MLVA, single nucleotide polymorphisms assay (SNP's), and multilocus sequence typing (MLST) (Keim *et al.,* 2000; Klee *et al.,* 2006a; b; Lista *et al.,* 2006; Marston *et al.,* 2006; van Ert *et al.,* 2007). *Bacillus cereus* and *B. anthracis* share identical 16S rRNA sequences (Ash *et al.,* 1991) and demonstrate only two nucleotide differences in 23S rRNA sequences (Ash & Collins, 1992). As indicated, the chromosomes are highly conserved amongst *B. cereus senso stricto* and species–specific pathogenicity were often plasmid encoded, but these genes cannot always be linked to pathogenicity solely through PCR and/ or sequence analysis (Rasko *et al.,* 2007). Some researchers believe that *B. cereus*, *B. anthracis* and *B. thuringiensis* should belong to a single species *B. cereus sensu lato,* because they have acquired plasmids which code for specific phenotypic characters such as pathogenicity and the chromosomes have a similar gene content and order (Daffonchio *et al.,* 2000; Rasko *et al.,* 2007).

It seems that when a *B. cereus sensu lato* strain adapts to a different ecological environment, selection of the particular genome traits characterizing its biodiversity becomes possible. *Bacillus cereus* isolates responsible for severe cases of pneumonia were found to harbour *B. anthracis* toxin and/or capsule virulence genes *pagA*, *lef*, *cya*, *capA, capB* and *capC* located on the plasmids (Hoffmaster *et al.*, 2006, 2008). Similarly, 2 isolates obtained from chimpanzees in Côte d'Ivoire and Cameroon that apparently died of anthrax, but when isolated, displayed motility on growth medium (Leendertz *et al.,* 2006) and that these strains residing between *B. cereus* and *B. anthracis* should be referred to as *B. cereus / B. anthracis sensu lato* until phylogenetic relationships and phenotypic characteristics can be firmly established (Okinaka *et al.,* 2006).

Despite the presence of multiple, well-established phenotypic and molecular markers to define *B. cereus*, *B. thuringiensis*, and *B. anthracis* there are often isolates that are mis-designated due to unusual properties or lack of sufficient information. *Bacillus* strains isolated from chimpanzee with clinical symptoms of anthrax were only designated *B. cereus* var *anthracis* after sequencing revealed them to contain a *B. cereus* chromosome and *B. anthracis* plasmids (Klee *et al.*, 2006a; 2006b; Okinaka *et al.*, 2006; Klee *et al.*, 2010). In this

study we typed isolates identified as *B. anthracis* mainly from the Northern Cape region with a view to determine if the MLVA assay would be able to corroborate bacteriological results and be able to differentiate between *B. anthracis* and *B. anthracis*-like isolates. The phylogentic analysis found some '*B. anthracis*' isolates to be similar to *B. cereus* and these isolates are discussed.

3.2 Materials and Methods

3.2.1 Isolate

A single *B. cereus* isolate (stored in glycerol at -80°C) obtained from a clinical case at the Bacteriology department at the Faculty of Veterinary Science, University of Pretoria, Onderstepoort, South Africa was plated out onto blood agar and incubated for 8 hours at 37°C. The culture was observed to be creamy in appearance with frilly edged colonies characteristic of *B. cereus* and was not susceptible to gamma-phage. Besides the knowledge that the isolate was submitted for diagnostic tests, from a clinical case, the origin and related history of the isolate is not known. Other *B. cereus / B. anthracis sensu lato* isolates included in this study are listed in Table 3.1.

Table 3.1: *Bacillus* isolates and their related history of isolation.

Key	SAMPLE number *	YEAR ISOLATED	LOCATION	Isolated from
Ba#000	*B. cereus*	unknown	Unknown	Unknown
Ba#092	3275-2D	2009	Kimberley	Soil
Ba#093	2110	1997	Herbert	Ovine
Ba#094	6103-6B	1998	Sendelingsdrif, Richtersveld NP	Caprise
Ba#095	6461-SP	1998	Ngomiland, Botswana	Elephant
Ba#096	3080-5A	2009	Kimberley	Bovine
Ba#097	6461-SP2	1998	Sendelingsdrif, Richtersveld	Caprise
Ba#098	5838	1998	Kimberley	Rooi hartebees
Ba#099	3080-1B	2009	Kimberley	Bovine
Ba#100	2991-2B	2009	Kimberley	Ovine
Ba#101	7693	1995	Unknown	Bovine
Ba#102	4980	1995	Vaalbos National Park	Kudu

Table 3.1: continued...

Key	SAMPLE ID	YEAR ISOLATED	LOCATION	Isolated from
Ba#103	6103-6D	1998	Botswana	Elephant
Ba#104	8334	1999	Maseru, Lesotho	Giraffe
Ba#105	6461-SP1	1998	Sendelingsdrif, Richtersveld NP	Caprise
Ba#106	2991-1B	2009	Kimberley	Ovine
Ba#107	3008-1A	2008	Kimberley	Soil
Ba#108	3122-2B	2009	Kimberley	Gemsbok
Ba#109	3080-3B	2009	Kimberley	Bovine
Ba#110	3079-1C	2009	Kimberley	Gemsbok
Ba#111	3132-1B	2009	Kimberley	Kudu
Ba#112	6057	1998	Kimberley	Wildebees
Ba#113	Sterne	N/A	N/A	-

**** All isolates identified as *B. anthracis sensu lato* except for *B. cereus*.**

Figure 3.1: Map of Northern Cape Province where *B. anthracis / B. cereus sensu lato* isolates were isolated.

3.2.2 MLVA assay

Bacillus cereus / B. anthracis senu lato DNA of each isolate (Table 3.1) was extracted using the QIAGEN QIAmp® DNA Mini kit (Qiagen) following the manufacturer's instructions for extraction of genomic DNA from Gram-positive bacteria as indicated in Chapter 2 section 2.2.2. The 31 VNTR loci (Table 2.1) were amplified as previously indicated in chapter 2 section 2.2.3.1. Any locus that did not produce a PCR fragment was optimized to ensure accuracy of the results. PCR amplicons were analysed by agarose gel electrophoresis and the copy numbers assigned as elucidated in Chapter 2 sections 2.2.4.1 and 2.2.5. Unweighted pair group method using arithmetic averages (UPGMA) cluster analysis were used to illustrate the population genetics (Figure 3.3) and Minimum Spanning Trees (MST) based on the similarity matrices of maximum UPGMA score and minimum sum based on MLVA data using Bionumeric version 6.0 (Figure 3.4 and 3.5).

3.2.3 Microbiology tests

Isolates (Ba#101 and Ba#104) were selected for further microbiology tests conducted at the ARC-OVI Reference laboratory, Onderstepoort, South Africa. These tests included motility, haemolytic activity, penicillin sensitivity and phage sensitivity.

3.3 Results

The *B. cereus* isolate (Ba#000) did not amplify at the loci of the plasmids (pXO1 and pXO2), Bams 30 and VNTR 17. The 31 VNTR loci profile for this isolate (Ba#000) is depicted in Figure 3.2 with the PCR fragment size and translated copy number at each loci for the *B. cereus* listed in Table 3.2.

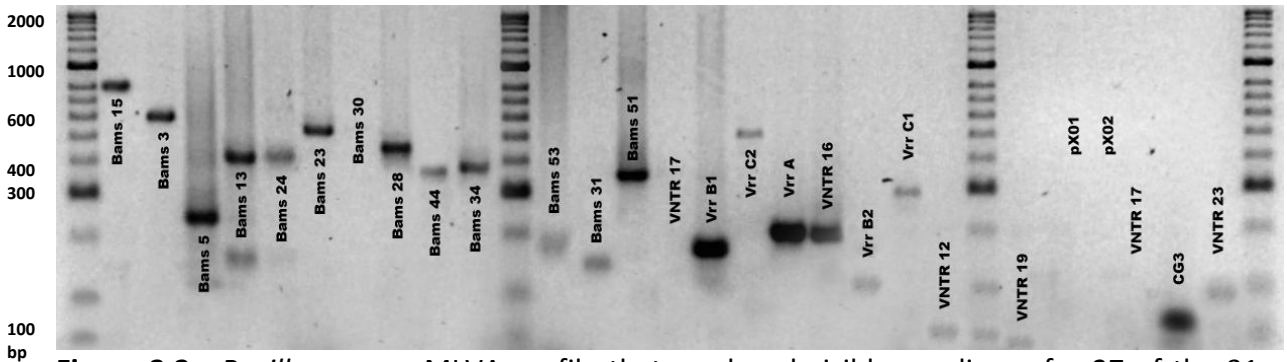


Figure 3.2: *Bacillus cereus* MLVA profile that produced visible amplicons for 27 of the 31 VNTR loci on a 3% agarose gel run at 200V for 3 hours.

Table 3.2: The fragment size and its corresponding copy number for each VNTR locus of the *Bacillus cereus* isolate obtained using gel electrophoresis. There was no amplification for Bams 30, pXO1, pXO2 and VNTR 17.

VNTR locus	repeat unit copy #	Fragment size (bp)	VNTR locus	repeat unit copy #	Fragment size (bp)
Bams 1	16	485	Bams 53	1	262
Bams 3	41	774	VrrA	8	290
Bams 5	6	346	VrrB1	22	276
Bams 13	37	517	VrrB2	11	198
Bams 15	76	886	VrrC1	33	400
Bams 21	12	766	VrrC2	21	604
Bams 22	12	663	CG3	1	153
Bams 23	11	651	pXO1	-	-
Bams 24	9	511	pXO2	-	-
Bams 25	17	451	VNTR12	1	109
Bams 28	15	517	VNTR16	1	219
Bams 30	-	-	VNTR17	-	-
Bams 31	6	250	VNTR19	3	93
Bams 34	8	417	VNTR23	4	196
Bams 44	9	456	VNTR35	2	104
Bams 51	9	493			

Twenty-two *B. cereus / B. anthracis sensu lato* isolates, including the Sterne vaccine strain (BA#113) which lacks the pXO2 plasmid, were analyzed using 31 VNTR loci (Appendix Table 5). Strain Ba#104 lacked both of the *B. anthracis* plasmids, pXO1 and pXO2 (indicated by -2 in Appendix Table 5) as well as Bams 30 and 34. Strains Ba#92 and 104 did not amplify the pXO1 locus.

The cluster analysis of MLVA31 data using UPGMA (Figure 3.3) grouped the isolates in 2 clades and the *B. cereus* (Ba#000) isolate clustered separately. The major clade, *B. anthracis* isolates (which also include Sterne strain Ba#113), consisted of 2 sub-clades (Figure 3.3). The first sub-clade included the majority of isolates mainly from the Northern Cape (indicated as Kimberely region in Figure 3.5) with the second, smaller border territory sub-clade, which included isolates from Sendelingsdrift and Botswana (BA#94, 95, 97, 103 and 105) (Figure 3.3; 3.4 and 3.5).

Like the Sterne vaccine strain, Ba#92 and Ba#108 are also attenuated strains however; they lack the plasmid pXO1 while Sterne lacks pXO2 and all group in the larger sub-clade. The isolates Ba#101 and Ba#104 clustered in second clade despite the fact that Ba#104 did not

amplify either of the plasmids pXO1 and pXO2 nor the Bams 30 and Bams 34 VNTR loci, while Ba#101 amplified all 31 VNTR loci (Appendix Table 5, Figure 3.3). Microbiology tests performed on isolates from the smaller, second clade consisting of Ba#101 and Ba#104 indicated that strains were non-motile, non-hemolytic and sensitive to penicillin. However, Ba#101 was phage sensitive and Ba#104 was phage resistant. According to microbiology test classification, Ba#101 is *B. anthracis* and Ba#104 is non-*B. anthracis*.

Isolates cluster according to their geographic origin using UPGMA and MST analyses. Isolates originating from Kimberly, Herbert and Vaalbos National Park clustered together in major sub-clade, while isolates (Ba#94, 95, 97, 103 and 105) originating from Sendelingsdrift and Botswana clustered together in a second, border territory sub-clade (Figure 3.3). Kimberley is the capital of the Northern Province located in the north western part of South Africa (Figure 3.1). Vaalbos National Park was situated on the Vaal River within the limits of Kimberly, but has since been closed and therefore no longer exists. Herbert is south-west from Kimberley in the Northern Cape Province. Namibia, Botswana and South Africa are neighbouring countries with common borders. Sendelingsdrift is a rugged, mountainous part of Richtersveld National Park. It falls on the South African side of the Orange River which forms the border to Namibia. Ngamiland is located in the North-Western part of Botswana and also shares a river border with Namibia.

The MST (Figure 3.5) demonstrates the minimum sum of weights of bacterial genetic distances between isolate edges (i.e. relatedness of isolates). The distance between the *B. anthracis* group and the *B. cereus* is nearly 40 minimum spanning units (MSU). The distance within the *B. anthracis* group between the Northern Cape region sub-clade and the Border territory sub-clade is less than 10 MSU, while the distance of isolates within each *B. anthracis* clade is approximately 1 MSU or less. Isolate Ba#104 is 15 MSU and 22 MSU from the Northern Cape sub-clade and border territory sub-clade, respectively. Isolate Ba#101 is 23 MSU from the Northern Cape sub-clade and 30 MSU from the border territory sub-clade.

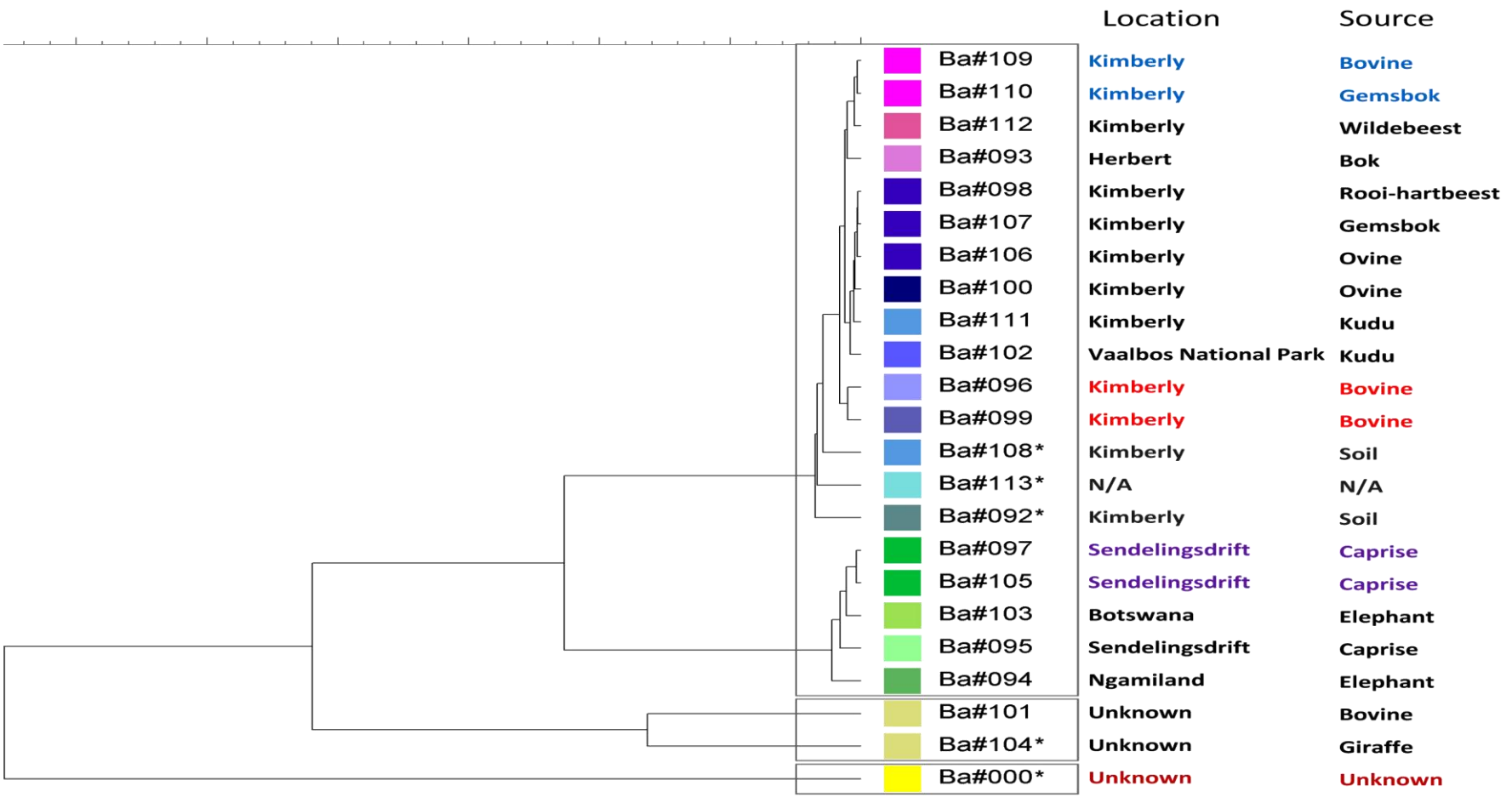


Figure 3.3: MLVA31 cluster analysis using UPGMA algorithm of *B. cereus*, 21 *B. anthracis* sensu lato and *B. anthracis* Sterne vaccine strain according to geographical region. The geographical and host information are included**.** Attenuated isolates are indicated by (*)

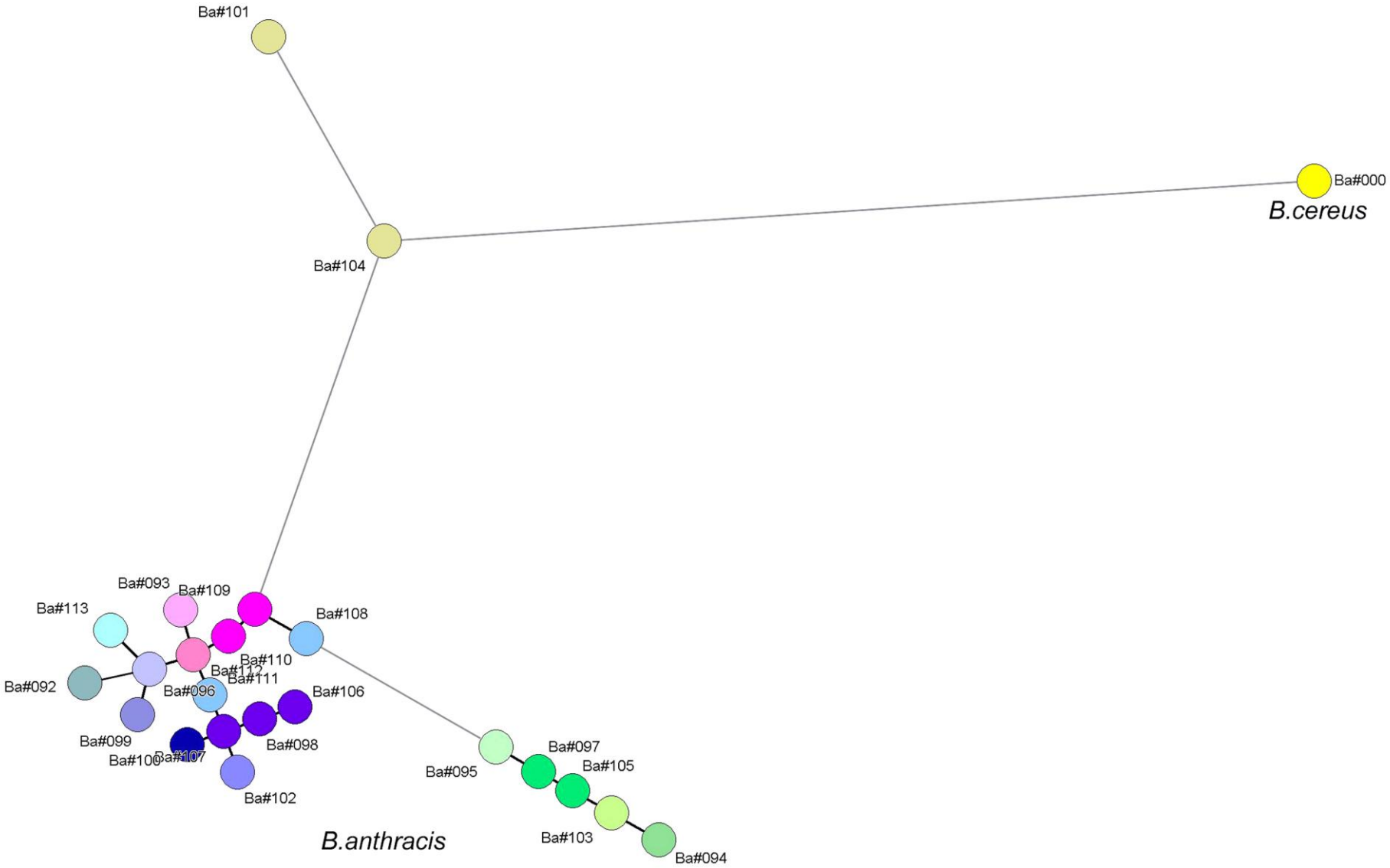


Figure 3.4: MST based on maximum score UPGMA illustrating the distinct isolate clusters.

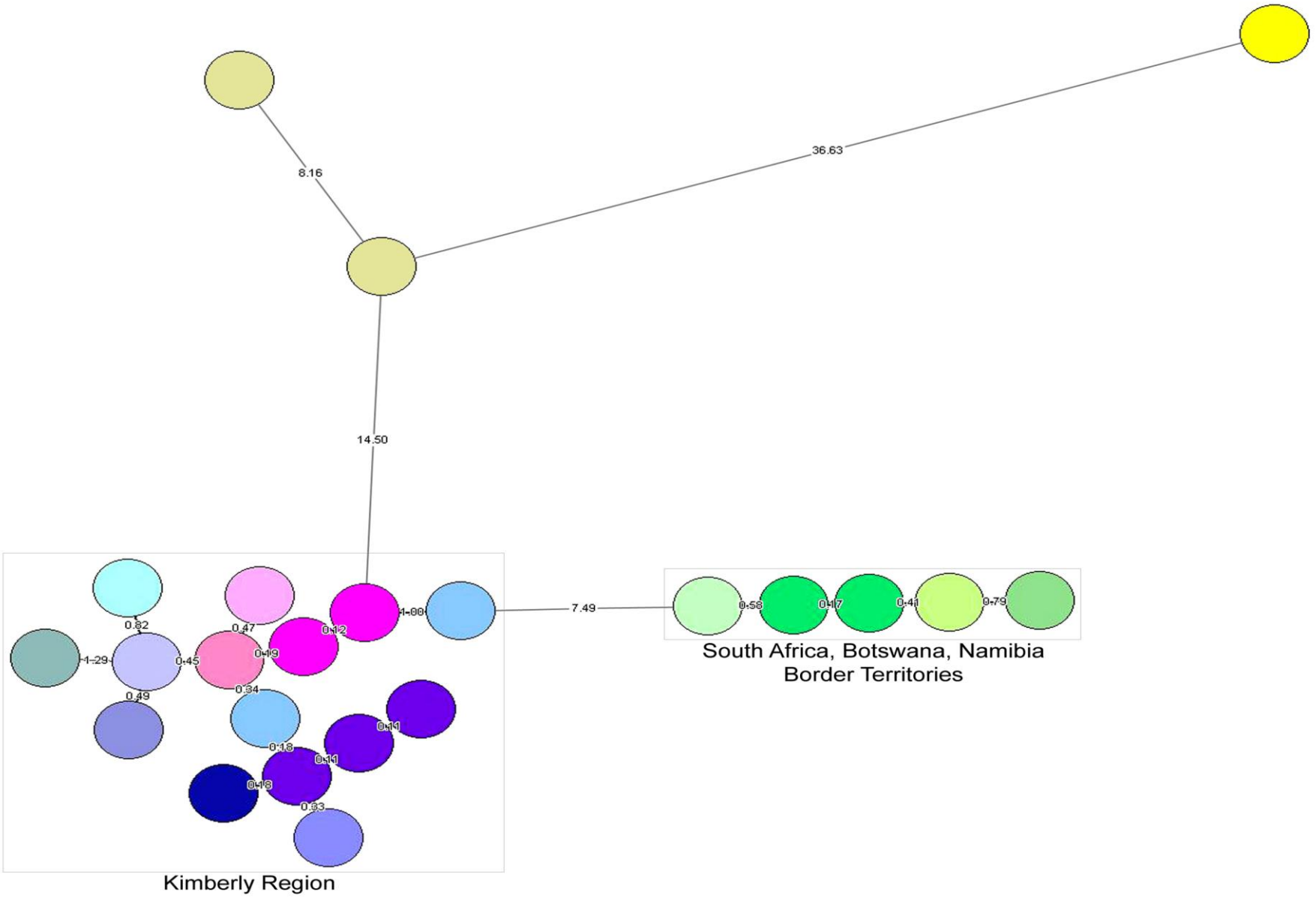


Figure 3.5: MST illustrating the minimum sum of weights of bacterial genetic distances between isolates.

3.4 Discussion

Bacillus cereus group species have been defined by biochemical and phenotypic / morphological characteristics and/or the ability for genetic transfer (sexual reproduction) (Godreuil *et al.,* 2005; Klee *et al.,* 2006b). There has been a great deal of debate on how to define a species in the field within the *B. cereus* group. With the advent of molecular techniques, it is hoped that the bacterial species definition could be refined (Woese, 1987; Cohan, 2002; Godreuil *et al.,* 2005). The 16S and 23S rRNA sequences are not able to differentiate members of the *B. cereus* group (Ash *et al.,* 1991; Ash & Collins,1992). It has been proposed that *B. anthracis* be considered a distinct species from *B. cereus* and *B. thuringiensis* because it is more monomorphic in nature than the other two bacteria (Ticknor *et al.,* 2001; Radnedge *et al.,* 2003).

According to Leski *et al.* (2009) the *bclABCDE* genes are useful in discriminating between members of the *B. cereus* group. Some of the *bcl* genes are target regions of VNTR markers namely Bams13 (*bcl*A), Bams30 (*bcl*B), Bams31 (*bcl*C) and Bams15 (*bcl*D) (*Lista et al.,* 2006). The *bcl* genes can be used to specifically differentiate *B. anthracis* from other members of the *B. cereus* group (Leski *et al.,* 2009). The Bams 13 and Bams 30 VNTR loci are 2 of the most polymorphic loci and code for genes involved in the exosporium coat (Lista *et al.,* 2006). The Bams 30 locus which corresponds to the *bcl*B region, did not amplify the *B. cereus* (BA#000) and Ba#104 isolates and Ba#101 had a small copy number compared to the other isolates giving credence to a hypothesis that these isolates belong under the banner of *B. cereus sensu lato* (Table 3.2; Appendix Table 5).

The VrrA locus was also believed to be a good candidate for discrimination between members of the genus *Bacillus* and of geographically different *B. anthracis* isolates in particular. Although it does not amplify across all isolates in the genus *Bacillus*, it amplifies more consistently amongst the members of the *B. cereus* group than CG3 and the other Vrr loci (Anderson *et al.,* 1996; Keim *et al.,* 2000; Klee *et al.,* 2006a; Marston *et al.,* 2006). It was noted by Keim *et al* (2000) that the VrrA amplicon sizes differed markedly between *B. anthracis* and other members of the *B. cereus* group. The VrrA locus amplified for all 23 isolates (including *B. cereus)* tested in this study and the sizes were comparable to those

commonly found in *B. anthracis*, but this could be attributed to the close geographic origin of these isolates. Similar results were observed by Klee *et al.* (2006a) where the *B. cereus* var. *anthracis* strains had amplicon sizes within the same range of *B. anthracis* isolates.

Horizontal transfer has been reported amongst soil borne *B. cereus* group that in turn greatly affects the phenotypic feature of the isolates (Gonzales *et al.,* 1982; Helgason et al., 2000a; 2000b; Hoffmaster *et al.,* 2006; Klee et al., 2010). The pXO1-like plasmids vary in size and copy number and are widespread throughout the *B. cereus* group. Pannucci *et al.* (2002) found a few *B. cereus* and *B. thuringiensis* isolates with sequences between 80% and 98% similar to that of pXO1. Similarly, Hu *et al.* (2006) found that half of ubiquitous *B. cereus* and *B. thuringiensis* isolates tested were PCR positive for conserved fragments of the pXO1 plasmid. The pXO1-like plasmids along with the chromosomal genes appear to form identifiable subgroups of *B. cereus* that can be associated with certain disease presentations. It is believed that the pXO1-like plasmids have co-evolved with their chromosome to further improve pathogenesis and/or niche adaptation of an organism (Rasko *et al.,* 2007). It therefore appears that pathogenic *B. cereus* harbour specialized plasmids associated with its clinical and metabolic phenotype that presents forms of emetic disease and anthrax-like pneumonia, depending on the genes it has managed to acquire (Rasko *et al.,* 2007).

Furthermore some *B.* cereus and *B. thuringiensis* isolates were found to be encapsulated (Sue *et al.,* 2006; Cachat *et al.,* 2008). It is believed that *B. cereus* and *B. thuringiensis* contains pXO2-derived genes as alleles coding for the capsule are found on the pXO2 plasmid of *B. anthracis.* It has been suggested that PCR-based screening methods are limited by the prerequisite of conserved primer sequences (Beyer & Bohm, 2003) and that pXO2-like plasmids may be more commonly distributed in the *B. cereus* group than initially believed (Cachat *et al.,* 2008). Sue *et al.* (2006) isolated *B. cereus* isolates displaying capsule production from severe human pneumonia cases. The study was able to demonstrate that varying levels of capsule expression is influenced by culturing conditions. The authors caution against using just one phenotypic test in the classic differentiation between *B. cereus* and *B. anthracis.*

MLVA typing clustered isolates according to their geographical region. The major clade, *B. anthracis* isolates are predominantly from Kimberly. The border territory isolates clustered in a sub-clade with the Northen Cape region isolates in another sub-clade (Fig. 3.3). Isolates Ba#101 and Ba#104 form a clade indicated as *B. anthracis / B. cereus sensu lato* group within the *B. anthracis* clade and groups between *B. anthracis* and *B. cereus* (Fig. 3.3). Isolate Ba#101 amplified all of the 31 MLVA loci whereas isolate Ba#104 only amplified 27 MLVA loci (Appendix Table 5). The MST indicated that these two isolates form a subgroup (Figure 3.4) and genetic distances with MST shows that Ba#104 is more closely related to *B. anthracis* (14.54) than *B. cereus* (36.63) (Figure 3.5) despite microbiology tests indicating it as non-*B. anthracis*. MST results indicate that these two isolates were most probably not from the Northern Cape Province as Ba#101 was unknown and Ba#104 was from Lesotho.

It can also be argued that pathogenic isolates from the *B. cereus* group appear to be more similar to each other than environmental isolates (Ticknor *et al.,* 2001). Since the *B. cereus* isolate (Ba#000) was from a clinical case, it is plausible it would bear a closer resemblance to other pathogenic isolates from the *B. cereus* group (Ticknor *et al.,* 2001). Identification of *B. anthracis* from environmental samples has been found to be much more difficult, because frequently, soil borne *Bacillus* species are present and borderline isolates might share some bacteriological similarities with *B. anthracis* (Klee *et al.,* 2006b).

There are other phenotypic characters like haemolytic and non-haemolytic activity of *B. anthracis* isolates from Poland, which were found to be indistinguishable on a molecular level, when using MLVA (Gierczynski *et al.,* 2006; 2009). Due to the ability of members of *the B. cereus* group to exchange genetic material, the utilization of molecular methods alone makes it difficult to identify a *Bacillus* strain as the genetic exchange between *Bacillus* strains, in the absence of genetic barriers, rules out this kind of classification at the species and subspecies level (Manzano *et al.,* 2009). While MLVA is a powerful molecular epidemiological method for tracing back the origin of strains, it is dependent on reliable background data before any sound conclusions can me made (Pilo *et al.,* 2008). Klee *et al* (2006) suggested using microbiology tests together with molecular tests identifying

plasmids and conserved chromosomal regions unique to each species. Non typical isolates should be further characterized using molecular techniques. The isolates in our study should be characterized using microbiology tests and those non-typical isolates should be further characterized using molecular techniques.

3.5 Conclusion

The MLVA results were able to highlight *B. anthracis / B. cereus sensu lato* isolates which require further study. The non-anthrax isolate Ba#104 grouped separately with the anthrax isolate Ba#101. This clustering could be due to unique ecological niche/geographical location of these isolates; as the results clearly show that MLVA is able to trace these isolates back to their geographical origin. For this study, MLVA has proven itself to be an excellent tool in determining genetic and geographical distance and the origin of isolates in spatial distribution models; however, it was not able to conclusively resolve the anthrax species question. MLVA still provides valuable data when linked to bacteriological tests which can definitively identify species or species-variants. Additional molecular methods other than MLVA need to be investigated and coupled with bacteriological and biochemical tests to identify *B. cereus / B. anthracis sensu lato* accurately from *B. anthracis* isolates. In the eternal words of The Bard, “to be or not to be” is still the question.

References

Andersen, G. L., Simchock, J. M., & Wilson, K. (1996). Identification of a region of genetic variability among *Bacillus anthracis* strains and related species. *Journal of Bacteriology, 178*(2)

Ash, C., & Collins, M. D. (1992). Comparative analysis of 23S ribosomal RNA gene sequences of *Bacillus anthracis* and emetic *Bacillus cereus* determined by PCR-direct sequencing. *FEMS Microbiology Letters, 73*(1-2), 75-80.

Ash, C., Farrow, J. A., Dorsch, M., Stackebrandt, E., & Collins, M. D. (1991). Comparative analysis of *Bacillus anthracis, Bacillus cereus*, and related species on the basis of reverse transcriptase sequencing of 16S rRNA. *International Journal of Systematic Bacteriology, 41*(3), 343-346.

Beyer, W., Böhm, R. (2003). Pitfalls in the diagnosis of *B. anthracis* from environmental samples –lessons from investigations of suspicious specimens. Mexico.

Cachat, E., Barker, M., Read, T. D., & Priest, F. G. (2008). A *Bacillus thuringiensis* strain producing a polyglutamate capsule resembling that of *Bacillus anthracis*. *FEMS Microbiology Letters, 285*(2), 220-226.

Cohan, F. M. (2002). What are bacterial species? *Annual Review of Microbiology, 56*, 457-487.

Daffonchio, D., Borin, S., Frova, G., Gallo, R., & Mori, E. (2000). A randomly amplified polymorphic DNA marker specific for the *Bacillus cereus* group is diagnostic for *Bacillus anthracis*. *Appl.Environ.Microbiol., 65*, 1298.

Gierczynski, R., Zasada, A. A., Kaluzewski, S., & Jagielski, M. (2006). Molecular characterisation of the haemolytic isolates of *Bacillus anthracis* originated in Poland. [Molekularna charakterystyka Hemolizujacych szczepow *Bacillus anthracis* izolowanych w Polsce] *Medycyna Doswiadczalna i Mikrobiologia, 58*(4), 339-346.

Gierczynski, R., Jakubczak, A., & Jagielski, M. (2009). Extended multiple-locus variable-number tandem-repeat analysis of *Bacillus anthracis* strains isolated in poland. *Polish Journal of Microbiology / Polskie Towarzystwo Mikrobiologow = the Polish Society of Microbiologists, 58*(1), 3-7.

Godreuil, S., Cohan, F., Shah, H., & Tibayrenc, M. (2005). Which species concept for pathogenic bacteria? an E-debate. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases, 5*(4), 375-387.

Gonzales, J. M. J., Brown, B. S., & Carlton, B. C. (1982). Transfer of *Bacillus thuringiensis* plasmids coding for d-endotoxin among strains of *Bacillus thuringiensis* and *Bacillus cereus*. *Proc.Natl.Acad.Sci.U.S.A., 79:6951–6955*

Helgason, E., Caugant, D. A., Olsen, I., & Kolsto, A. B. (2000a). Genetic structure of population of *Bacillus cereus* and *B. thuringiensis* isolates associated with periodontitis and other human infections. *Journal of Clinical Microbiology, 38*(4), 1615-1622.

Helgason, E., Okstad, O. A., Caugant, D. A., Johansen, H. A., Fouet, A., Mock, M., et al. (2000b). *Bacillus anthracis, Bacillus cereus, and Bacillus thuringiensis*--one species on the basis of genetic evidence. *Applied and Environmental Microbiology, 66*(6), 2627-2630.

Hoffmaster, A. R., Hill, K. K., Gee, J. E., Marston, C. K., De, B. K., Popovic, T., et al. (2006). Characterization of *Bacillus cereus* isolates associated with fatal pneumonias: Strains

are closely related to Bacillus anthracis and harbor B. anthracis virulence genes. *Journal of Clinical Microbiology, 44*(9), 3352-3360.

Hoffmaster, A. R., Novak, R. T., Marston, C. K., Gee, J. E., Helsel, L., Pruckler, J. M., et al. (2008). Genetic diversity of clinical isolates of *Bacillus cereus* using multilocus sequence typing. *BMC Microbiology, 8*, 191.

Hu, X., Hansen, B. M., Hendriksen, N. B., & Yuan, Z. (2006). Detection and phylogenic analysis of one anthrax virulence plasmid pXO1 conservative open reading frame ubiquitous presented within *Bacillus cereus* group strains. *Biochemical and Biophysical Research Communications, 349*(4), 1214-1219.

Kaneko, T., Nozaki, R., & Aizawa, K. (1978). Deoxyribonucleic acid relatedness between *Bacillus anthracis* and *Bacillus cereus* and *Bacillus thuringiensis*. *Microbiol.Immunol., 22*, 639.

Keim, P., Price, L. B., Klevytska, A. M., Smith, K. L., Schupp, J. M., Okinaka, R., et al. (2000). Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *Journal of Bacteriology, 182*(10), 2928-2936.

Klee, S. R., Brzuszkiewicz, E. B., Nattermann, H., Bruggemann, H., Dupke, S., Wollherr, A., et al. (2010). The genome of a *Bacillus* isolate causing anthrax in chimpanzees combines chromosomal properties of B. cereus with *B. anthracis* virulence plasmids. *PloS One, 5*(7), e10986.

Klee, S. R., Nattermann, H., Becker, S., Urban-Schriefer, M., Franz, T., Jacob, D., et al. (2006a). Evaluation of different methods to discriminate *Bacillus anthracis* from other bacteria of the *Bacillus cereus* group. *Journal of Applied Microbiology, 100*(4), 673-681.

Klee, S. R., Ozel, M., Appel, B., Boesch, C., Ellerbrok, H., Jacob, D., et al. (2006b). Characterization of *Bacillus anthracis*-like bacteria isolated from wild great apes from Cote d'Ivoire and Cameroon. *Journal of Bacteriology, 188*(15), 5333-5344.

Lalitha, M. K., & Thomas, M. K. (1997). Penicillin resistance in *Bacillus anthracis*. *Lancet, 349*(9064), 1522.

Le Fleche, P., Hauck, Y., Onteniente, L., Prieur, A., Denoeud, F., Ramisse, V., et al. (2001). A tandem repeats database for bacterial genomes: Application to the genotyping of *Yersinia pestis* and *Bacillus anthracis. BMC Microbiology, 1*, 2.

Leendertz, F. H., Yumlu, S., Pauli, G., Boesch, C., Couacy-Hymann, E., Vigilant, L., et al. (2006). A new *Bacillus anthracis* found in wild chimpanzees and a gorilla from west and central Africa. *PLoS Pathogens, 2*(1), e8.

Leski, T. A., Caswell, C. C., Pawlowski, M., Klinke, D. J., Bujnicki, J. M., Hart, S. J., et al. (2009). Identification and classification of *bcl* genes and proteins of *Bacillus cereus* group organisms and their application in *Bacillus anthracis* detection and fingerprinting. *Applied and Environmental Microbiology, 75*(22), 7163-7172.

Liang, X., & Yu, D. (1999). Identification of *Bacillus anthracis* strains in China. *Journal of Applied Microbiology, 87*(2), 200-203.

Lista, F., Faggioni, G., Valjevac, S., Ciammaruconi, A., Vaissaire, J., le Doujet, C., et al. (2006). Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. *BMC Microbiology, 6*, 33.

Manzano, M., Giusto, C., Lacumin, L., Cantoni, C., & Comi, G. (2009). Molecular methods to evaluate biodiversity in *Bacillus cereus* and *Bacillus thuringiensis* strains from different origins. *Food Microbiology, 26*(3), 259-264.

Marston, C. K., Gee, J. E., Popovic, T., & Hoffmaster, A. R. (2006). Molecular approaches to identify and differentiate *Bacillus anthracis* from phenotypically similar *Bacillus* species isolates. *BMC Microbiology, 6*, 22.

Okinaka, R., Pearson, T., & Keim, P. (2006). Anthrax, but not *Bacillus anthracis*? *PLoS Pathogens, 2*(11), e122.

Pannucci, J., Okinaka, R. T., Sabin, R., & Kuske, C. R. (2002). Bacillus anthracis pXO1 plasmid sequence conservation among closely related bacterial species. *Journal of Bacteriology, 184*(1), 134-141.

Pilo, P., Perreten, V., & Frey, J. (2008). Molecular epidemiology of *Bacillus anthracis*: Determining the correct origin. *Applied & Environmental Microbiology, 74*(9), 2928-2931.

Radnedge, L., Agron, P. G., Hill, K. K., Jackson, P. J., Ticknor, L. O., Keim, P., et al. (2003). Genome differences that distinguish *Bacillus anthracis* from *Bacillus cereus* and *Bacillus thuringiensis*. *Applied and Environmental Microbiology, 69*(5), 2755-2764.

Rasko, D. A., Rosovitz, M. J., Oekstad, O. A., Fouts, D. E., Jiang, L., Cer, R. Z., et al. (2007). Complete sequence analysis of novel plasmids from emetic and periodontal *Bacillus cereus* isolates reveals a common evolutionary history among the *B. cereus*-group plasmids, including *Bacillus anthracis* pXO1. *Journal of Bacteriology, 189*(1), 52-64.

Sue, D., Hoffmaster, A. R., Popovic, T., & Wilkins, P. P. (2006). Capsule production in *Bacillus cereus* strains associated with severe pneumonia. *Journal of Clinical Microbiology, 44*(9), 3426-3428.

Ticknor, L. O., Kolsto, A. B., Hill, K. K., Keim, P., Laker, M. T., Tonks, M., et al. (2001). Fluorescent amplified fragment length polymorphism analysis of Norwegian *Bacillus cereus* and *Bacillus thuringiensis* soil isolates. *Applied and Environmental Microbiology, 67*(10), 4863-4873.

Turnbull, P. C. (1999). Definitive identification of *Bacillus anthracis*--a review. *Journal of Applied Microbiology, 87*(2), 237-240.

Van Ert, M. N., Easterday, W. R., Huynh, L. Y., Okinaka, R. T., Hugh-Jones, M. E., Ravel, J., et al. (2007). Global genetic population structure of *Bacillus anthracis. PLoS ONE, 2*(5), e461.

Vilas-Boas, G. T., Peruca, A. P., & Arantes, O. M. (2007). Biology and taxonomy of *Bacillus cereus, Bacillus anthracis*, and *Bacillus thuringiensis*. *Canadian Journal of Microbiology, 53*(6), 673-687.

Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews, 51*(2), 221-271.

Chapter 4:

Genotyping of southern African *Bacillus anthracis sensu lato* strains using multilocus variable number of tandem repeat analyses

Abstract

Genotyping of southern African *Bacillus anthracis sensu lato* isolates was done using multilocus variable number of tandem repeat analysis (MLVA) using the MLVA8, MLVA15 and MLVA25 panels. The latter two panels were combined and we also report on the MLVA31 panel. These panels were evaluated as a tool for epidemiological studies for southern African outbreaks. The 113 *B. anthracis sensu lato* isolates were differentiated using the 8, 15, 25 and 31 MLVA panels into 52, 61, 72 and 80 genotypes. The MLVA15 panel was unable to differentiate between *B. anthracis sensu lato* and *B. cereus*, which makes it less ideal for genotyping purposes. Cluster analysis using the 25 and 31 MLVA panel were very similar. The MLVA panels all separated the isolates into 2 lineages namely A and B clades, but the larger MLVA panels correlated the isolate more towards the geographic origin of the outbreak. The MLVA genotyping technique proved to be an excellent tool in tracing an isolate back to its geographic origin, but was not able to reveal any insights into neither preferential species infection nor the predominance of a single genotype during an outbreak. This is most likely due to the paucity of isolates collected during a single outbreak (which makes finding any statistically significant trends difficult) rather than an unfitness of the typing technique.

4.1 Introduction

Anthrax is a highly infectious disease of domestic and wild animals and humans are also susceptible. This disease affects mammals, where, ruminants and hindgut fermenters have shown themselves to be most vulnerable. Carnivores, suids and primates (including humans) are more resistant (Hugh-Jones & De Vos, 2002; Fortini *et al.,* 2007; Beyer & Turnbull, 2009). The disease is characterised by sudden death and the presence of the bacterium *Bacillus anthracis* in the blood and body fluids (Keim *et al.,* 2000). Anthrax occurs worldwide, however in endemic areas, such as southern Africa, cyclic outbreaks are the norm amongst wildlife (De Vos, 1990), especially in its large, free roaming wild ungulate populations (Smith *et al.,* 2000).

Outbreaks are usually set in motion when an infected animal sheds anthrax spores upon its death, which are then spread along the rivers. Even amongst the susceptible animals, there seems to be a disproportionate infection of a particular herbivorous species within a region. This could be due to stocking densities, herd behaviour and/or spread by living vectors. In recent years, buffalo (*Syncerus caffer*) have had the largest recorded anthrax-related mortalities in the Kruger National Park (KNP), South Africa (Strauss, 2010). The kudu (*Tragelaphus imerbis; T. strepisceros*) has accounted for more than 50% of anthrax cases during past outbreaks. This is in contrast to Etosha National Park, Namibia where zebras (*Equus burchelli*) account for almost half the anthrax deaths recorded while kudu barely make up 10% of the recorded mortalities (De Vos, 1990; Lindeque & Turnbull, 1994).

The development of modern molecular techniques has made it possible to trace outbreaks back to their origin and even distinguish strains related to bioterrorist attacks from wild-type strains (Keim *et al.,* 2000; Okinaka *et al.,* 2008; Van Ert *et al.,* 2007). Assays such as single nucleotide polymorphisms (SNP) and multi-locus sequence typing (MLST) have proven to contribute toward the epidemiology of anthrax, but the resolution offered by these techniques is limited in the case of *B. anthracis,* because it has high homogeneity amongst sequences (Denoeud & Vergnaud, 2004; Van Ert *et al.,* 2007; Okinaka *et al.,* 2008; Beyer & Turnbull, 2009). The discovery of tandem repeat sequences which differ in their repetition hence producing varied length polymorphisms, demonstrated that even highly monomorphic bacteria such as *B. anthracis* could be differentiated through identification of such polymorphic genomic regions (Keim *et al.,* 2000). When multiple loci of these variable

tandem repeat regions (VNTRs), indicated as MLVA, was targeted it enabled distinguishing different outbreaks by different strains, to trace an outbreak back to its possible origin and to track the routes of transmission of an outbreak strain within and between animal populations (Keim *et al.,* 2000; Le Fleche *et al.,* 2001; Lista *et al.,* 2006; Beyer & Turnbull, 2009).

Keim *et al.* (2000) used MLVA8, which was extended to MLVA15 by Van Ert *et al.* (2007) and later to MLVA25 (Le Fleche *et al.,* 2001). The extension of VNTR loci increases the resolution power, e.g. with the MLVA8 from 160 isolates were resolved into 30 genotypes whereas with the MLVA25 the same isolates were resolved into 67 genotypes (Lista *et al.,* 2006). A combination of the MLVA15 and MLVA25 markers resulted in MLVA31 that Beyer & Turnbull (2009) reported to be superior. For the purpose of this study, the resolution of these different panels, were juxtaposed to ascertain the value of each in understanding the epidemiology of anthrax. The MLVA31 was used to determine the relationship, if any, between genotypes and their point of origin.

4.2 Materials and Methods

Bacillus cereus / B. anthracis sensu lato isolates are listed in Appendix Table 1 and Table 3.2 in Chapter 3. The copy number based on tandem repeats from agarose and capillary electrophoresis based on MLVA8, MLVA15, MLVA25 and MLVA31 (Table 4.1 indicates the VNTR loci used for each) were managed using the Bionumerics version 6.0 (Applied Maths). These were generated as previously described using capillary and agarose electrophoresis (see chapter 2). UPGMA (unweighted pair group method using arithmetic averages) cluster analysis and similarity matrices were done using Bionumeric version 6.0. Isolates were considered the same, if they had exactly the same copy number profile for all loci. Two control isolates were used, namely *B. cereus (*BA#000) and *B. anthracis* Sterne strain (Ba#113).

The genotype and map coordinates at collection or carcass locations of each *B. anthracis* isolate was plotted on a map, where possible, to establish if there is a visible distribution pattern regarding geographic origin and isolate phylogeny / relatedness (Appendix Table 1). Before the collection sites from the KNP isolates could be plotted, it was necessary to translate the 4 digit military grid code (used to document the isolation location) into latitude and longitude co-ordinates. This was achieved through a map key Appendix Figure 9: A (KNP

North) and B (KNP South)) provided by the geographic information systems (GIS) and remote sensing division of SANparks. Appendix Table 1 tabulates the collection details for the 113 isolates designated Ba#001 to Ba#113 which includes the co-ordinates. The genotypes were plotted using ArcGIS version 9.3 (Esri).

Table 4.1: The MLVA panels and the VNTR loci of each panel

VNTR Panels						
MLVA 8	**MLVA 15**	**MLVA 25**		**MLVA 31**		
VrrA	VrrA	VrrA	Bams 21	VrrA	Bams 21	VNTR 12
VrrB1	VrrB1	VrrB1	Bams 22	VrrB1	Bams 22	VNTR16
VrrB2	VrrB2	VrrB2	Bams 23	VrrB2	Bams 23	VNTR 17
VrrC1	VrrC1	VrrC1	Bams 24	VrrC1	Bams 24	VNTR 19
VrrC2	VrrC2	VrrC2	Bams 25	VrrC2	Bams 25	VNTR 23
CG3	CG3	CG3	Bams28	CG3	Bams28	VNTR 35
pXO1	pXO1	pXO1	Bams 30	pXO1	Bams 30	
pXO2	pXO2	pXO2	Bams 31	pXO2	Bams 31	
	Bams 1	Bams 1	Bams 34	Bams 1	Bams 34	
	VNTR 12	Bams 3	Bams 44	Bams 3	Bams 44	
	VNTR16	Bams 5	Bams 51	Bmas 5	Bams 51	
	VNTR 17	Bams 13	Bams 53	Bams 13	Bams 53	
	VNTR 19	Bams 15		Bams 15		
	VNTR 23					
	VNTR 35					

4.3 Results and Discussion

4.3.1 MLVA8

The MLVA8 panel consist of the first VNTR for *B. anthracis,* VrrA (Andersen *et al*., 1996), which was found in a variable region consisting of 12 bp tandem repeat in an open reading frame (ORF) of a 30-kDa protein (Jackson *et al.,* 1997), VNTRs in pXO1 and pXO2 plasmids and 5 chromosomal VNTRs were found with amplified fragment length polymorphisms (AFLP) (Keim *et al.,* 2000). The MLVA8 panel (Table 4.1) was developed to be separated with capillary electrophoresis (Keim *et al.,* 2000; 2004).

In the study by Keim *et al.* (2000), a small percentage of isolates were found to be attenuated and were excluded from the phylogenetic study to evaluate isolates only containing both plasmids. Out of a sample size of 426, 89 genotyopes were resolved in 2 major genetic lineages designated A and B. The A cluster subdivided into 4 subclades and members of the A cluster are found worldwide. The B clade was almost exclusively composed of South African and Mozambique isolates. Isolates originating in Zambia and Namibia, as well as Sterne ($34F_2$) were found scattered amongst the A3 clade isolates.

With the MLVA8 panel, 52 different genotypes amongst 113 isolates could be differentiated (Table 4.2). Cluster analysis using the MLVA8 panel identifies 2 major genetic lineages (clade A and B) amongst the southern Africa *B. anthracis sensu lato* isolates (Figure 4.1). The B clade consist of *B. anthracis* isolates mainly from the 1975 anthrax outbreak in KNP in Pafuri region (most northern parts of KNP; Ba#017, 018 and 022) and the rest of the isolates group in various subclades within the A clade. Attenuated isolates Ba#092 and 108 from Kimberley region in Northern Cape form a subclade of A with the rest of the isolates from KNP, Border regions, Rondebosch, Zambia and Northern Cape regions grouping in the other subclade. The other attenuated strain Ba#104 clustered in the same subclade with the attenuated Sterne strain. However, the subclades of the A clade in our study was different from those found in Keim *et al.* (2000) as our study contained only southern African isolates compared to their global collection of isolates. *Bacillus cereus* is an out-group and excluded from genotyping. It is clear from this study that the MLVA8 resolution is too limited as it can only differentiate between isolates of great diversity and is not able to distinguish between isolates originating from different geographic locations, like those of Zambia, Northern Cape and KNP regions.

Table 4.2: Genotypes of *Bacillus anthracis sensu lato* isolates produced with MLVA8 panel

Isolate	genotypes	Clade	Isolate	genotypes	Clade
Ba#030	1	A1	Ba#073	15	A1
Ba#033	1	A1	Ba#109	15	A1
Ba#034	2	A1	Ba#112	16	A1
Ba#0101	3	A3	Ba#110	17	A1
Ba#036	4	A1	Ba#108	18	A2
Ba#037	4	A1	Ba#105	19	A1
Ba#038	5	A1	Ba#103	20	A1
Ba#095	6	A1	Ba#102	21	A1
Ba#099	6	A1	Ba#098	22	A1
Ba#107	6	A1	Ba#097	23	A1
Ba#106	6	A1	Ba#096	24	A1
Ba#104	7	A3	Ba#094	25	A1
Ba#113	7	A3	Ba#093	26	A1
Ba#081	8	A1	Ba#092	27	A2
Ba#080	8	A1	Ba#079	28	A1
Ba#068	9	A1	Ba#078	29	A1
Ba#071	9	A1	Ba#077	30	A1
Ba#058	10	A1	Ba#076	31	A1
Ba#063	11	A1	Ba#065	32	A1
Ba#066	11	A1	Ba#064	33	A1
Ba#067	11	A1	Ba#062	34	A1
Ba#069	11	A1	Ba#060	36	A1
Ba#070	11	A1	Ba#052	37	A1
Ba#072	11	A1	Ba#053	37	A1
Ba#001	12	A1	Ba#055	37	A1
Ba#006	12	A1	Ba#047	40	A1
Ba#009	12	A1	Ba#046	41	A1
Ba#010	12	A1	Ba#044	42	A1
Ba#011	12	A1	Ba#045	42	A1
Ba#012	12	A1	Ba#061	43	A1
Ba#013	12	A1	Ba#027	44	A1
Ba#014	12	A1	Ba#026	45	A1
Ba#015	12	A1	Ba#024	46	A1
Ba#016	12	A1	Ba#023	47	A1
Ba#039	12	A1	Ba#022	48	B
Ba#043	12	A1	Ba#021	49	A1
Ba#017	13	B	Ba#002	50	A1
Ba#018	13	B	Ba#007	51	A1
Ba#005	14	A1	Ba#008	52	A1

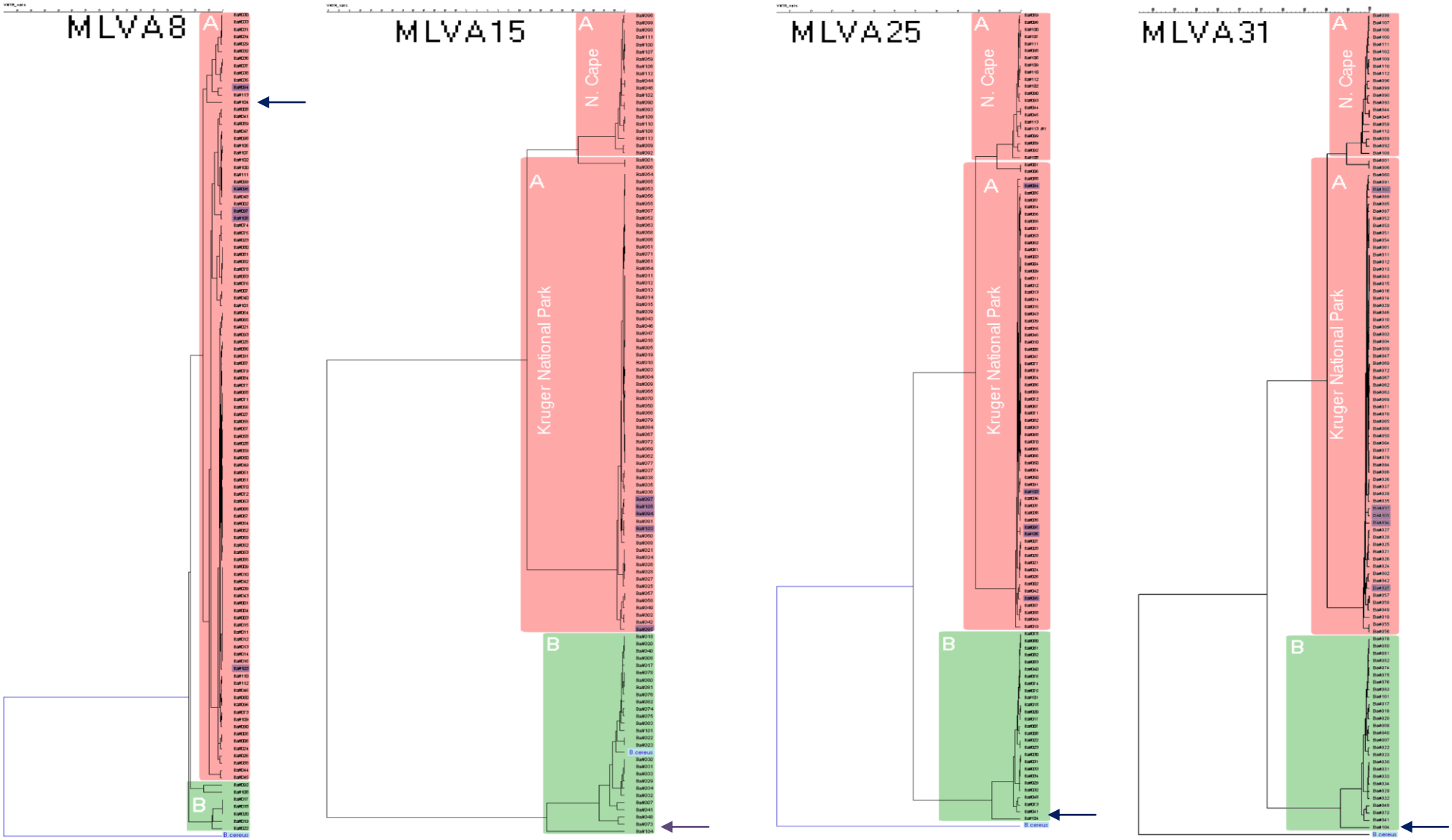


Figure 4.1: UPGMA dendograms for the MLVA8 (left), MLVA15 (middle left), MLVA25 (middle right) and MLVA31 (right) demonstrating the genetic distance between lineages. In the MLVA 8, 25 and 31 panels *B. cereus* (in blue) is an out-group. The isolates highlighted in purple are from the border territories. Isolate Ba#104 is indicated with an arrow. * See pages 128-131 for more detail.

4.3.2 The MLVA15 panel

The success of the VNTR 8 panel became the basis for the targeting of other VNTR loci. As the MLVA8 panel was not sufficient for molecular forensic approaches (Lista *et al.,* 2006; Van Ert *et al.,* 2007), additional markers were developed. Van Ert *et al* (2007) added 7 additional VNTR markers to produce the MLVA15 panel (Table 4.1) that was also separated using capillary electrophoresis (automated system). This larger panel differentiate the isolates into 3 major lineages (A, B and C) with the A lineages consisting of worldwide dispersed members, whereas the rarer genotypes were located in the B and C lineages. Members of the B clade were solely from the KNP. The C branch also appeared to have significantly lower evolutionary rates, which was indicative of fewer infective cycles (Van Ert *et al.,* 2007). The authors also tested the same 400 samples using the MLVA8 and MLVA15 panels that produced 89 and 221 genotypes, respectively (Keim *et al.,* 2000; Van Ert *et al.,* 2007). Most of the unique genotypes were resolved in the shallow branches within lineages (Van Ert *et al.,* 2007).

For our study, the 113 isolates were differentiated into 61 genotypes (Table 4.3 and Figure 4.1). This panel only differentiated a further 9 genotypes into two lineages A and B. The MLVA15 panel did not group *B. cereus* as an outgroup, but grouped in the B clade with Ba#17, 18, 20, 22 that grouped in B clade using 8 VNTR markers. Attenuated strain Ba#104 that lacked both plasmids like the *B. cereus* isolate (Ba#000) grouped in the B clade but in a separate subclade (Figure 4.1). The attenuated strains Ba# 92, 108 and 113 lacking one of the plasmids all clustered together in subclade of A with Northern Cape region isolates mainly from Kimberley and Rondebosch. Most of the KNP, Zambia and border regions grouped in another subclade of A (highlighted in purple in Figure 4.1).

The MLVA8 panel consist of 6 chromosomal markers that are also common to other members of the *B. cereus* group, however it was believed that they produce allele profiles that differentiate them from anthrax (i.e. allowing identification of anthrax and non-anthrax species) (Lista *et al.,* 2006). However Klee *et al*. (2006) characterized *B. cereus* isolates with anthrax plasmids pXO1 and pXO2. As most studies did not include non-anthrax bacilli as controls in their molecular typing, it is difficult to determine whether these systems are truly able to differentiate between members of the *B. cereus g*roup. The MLVA15 panel increased the resolution slightly from the MLVA8 panel but could not differentiate *B. cereus* from *B. anthracis* isolates.

Table 4.3: Genotypes of *Bacillus cereus / B. anthracis sensu lato* isolates produced with MLVA15 panel

Isolate	genotype	clade	Isolate	genotype	clade	Isolate	genotype	clade
Ba#104	1	B2	**Ba#088**	27	A1	**Ba#061**	43	A1
Ba#073	2	B1	**Ba#060**	28	A1	**Ba#081**	44	A1
Ba#048	3	B1	**Ba#103**	29	A1	**Ba#071**	45	A1
Ba#007	4	B1	**Ba#091**	29	A1	**Ba#051**	45	A1
Ba#041	5	B1	**Ba#094**	30	A1	**Ba#088**	46	A1
Ba#032	6	B1	**Ba#097**	31	A1	**Ba#068**	46	A1
Ba#034	7	B1	**Ba#105**	32	A1	**Ba#063**	46	A1
Ba#029	8	B1	**Ba#036**	33	A1	**Ba#052**	47	A1
Ba#033	9	B1	**Ba#035**	33	A1	**Ba#087**	47	A1
Ba#030	9	B1	**Ba#038**	34	A1	**Ba#055**	47	A1
Ba#031	9	B1	**Ba#037**	34	A1	**Ba#058**	47	A1
Ba#000	10	B1	**Ba#067**	35	A1	**Ba#053**	48	A1
Ba#023	11	B1	**Ba#072**	35	A1	**Ba#085**	48	A1
Ba#022	11	B1	**Ba#069**	35	A1	**Ba#054**	48	A1
Ba#101	12	B1	**Ba#062**	35	A1	**Ba#001**	49	A3
Ba#083	13	B1	**Ba#077**	35	A1	**Ba#006**	49	A3
Ba#074	14	B1	**Ba#084**	36	A1	**Ba#089**	50	A1
Ba#075	15	B1	**Ba#079**	37	A1	**Ba#092**	50	A1
Ba#082	16	B1	**Ba#066**	37	A1	**Ba#113**	51	A3
Ba#076	17	B1	**Ba#050**	37	A1	**Ba#108**	52	A3
Ba#078	18	B1	**Ba#065**	38	A1	**Ba#109**	53	A1
Ba#080	18	B1	**Ba#070**	38	A1	**Ba#110**	53	A1
Ba#081	18	B1	**Ba#009**	39	A1	**Ba#090**	54	A1
Ba#008	19	B1	**Ba#004**	40	A1	**Ba#093**	54	A1
Ba#040	20	B1	**Ba#003**	40	A1	**Ba#044**	55	A2
Ba#020	20	B1	**Ba#010**	40	A1	**Ba#045**	55	A2
Ba#018	20	B1	**Ba#019**	40	A1	**Ba#102**	56	A1
Ba#095	21	A1	**Ba#005**	41	A1	**Ba#098**	57	A3
Ba#042	22	A1	**Ba#016**	41	A1	**Ba#111**	57	A3
Ba#002	22	A1	**Ba#047**	41	A1	**Ba#100**	57	A1
Ba#049	23	A1	**Ba#046**	41	A1	**Ba#107**	57	A1
Ba#057	24	A1	**Ba#043**	41	A1	**Ba#059**	58	A1
Ba#058	24	A1	**Ba#039**	41	A1	**Ba#106**	59	A1
Ba#025	25	A1	**Ba#015**	41	A1	**Ba#112**	60	A3
Ba#027	26	A1	**Ba#013**	41	A1	**Ba#096**	61	A1
Ba#021	26	A1	**Ba#014**	41	A1	**Ba#099**	61	A3
Ba#024	26	A1	**Ba#012**	41	A1			
Ba#026	26	A1	**Ba#011**	41	A1			
Ba#028	26	A1	**Ba#064**	42	A1			

4.3.3 The MLVA25 panel

Le Fleche *et al.* (2001) described the use of database (http://minisatellites.u-psud.fr) of tandem repeats from available bacterial genome sequences, including *B. anthracis,* to facilitate characterization of polymorphic markers used for typing. Le Fleche *et al.* (2001) and Lista *et al.* (2006) used the MLVA25 panel (Table 4.1), which consisted of the original 8 MLVA panel (Keim *et al.,* 2000) and 17 additional chromosomal VNTR markers. The MLVA25 panel was applied to 160 *B. anthracis* isolates that differentiated 67 genotypes, whereas the MLVA8 panel identified 39 genotypes (Lista *et al.,* 2006). In their study the agarose and capillary electrophoresis platforms for the MLVA25 were compared, which identified discrepancies at Bams13 and Bams30. These discrepancies were due to increase discrimination power of capillary electrophoresis to agarose electrophoresis especially for tandem repeat units larger than 8 bp, like Bams13 and 30. The authors concluded that agarose electrophoresis was sufficient for typing VNTR loci with repeat units 9 bp or more in length (Lista *et al.,* 2006).

In our study, 72 genotypes of the 113 *B. anthracis sensu lato* isolates (Table 4.4) were resolved using the MLVA 25 panel. *Bacillus cereus* grouped as an out-group (Figure 4.1). Bams13 and 30 that code for the *Bcl* genes along with Bams15 and Bams31, which were directly connected in the coding of *Bcl* genes, and therefore involved in the expression of proteins related to the exosporium and hence spore viability (Leski *et al.,* 2009). It seems that these VNTR loci not included in MLVA15 panel enable differentiation of *B. cereus* from *B. anthracis sensu lato.*

The cluster analysis of the MLVA25 panel grouped the samples into lineages A and B, but more according to their geographic origin than the MLVA 8 and 15 panels. The Sterne strain grouped in a subclade of the A lineage as did the Northern Cape region isolates mainly from Kimberley. This panel was able to identify *B. cereus* as an outlier and grouped Ba#104 as a separate subclade in the B lineage (Figure 4.1).

It was expected that one particular genotype would prove more prominent or dominant in an outbreak. But thus far the genotype distributions have not identified a prominent or dominant genotype. This is probably due to an excess of data which then becomes “white noise” in the search for any obvious patterns. This is discussed in more detail in section 4.3.6

Table 4.4: Genotypes of *Bacillus anthracis sensu lato* identified with the MLVA25 panel

Isolate	genotype	Clade	Isolate	genotype	clade
Ba#100	1	A3	Ba#046	34	A1
Ba#107	1	A3	Ba#047	35	A1
Ba#044	2	A3	Ba#048	36	A1
Ba#045	2	A3	Ba#058	37	A1
Ba#052	3	A1	Ba#060	38	A1
Ba#053	3	A1	Ba#062	39	A1
Ba#055	3	A1	Ba#063	40	A1
Ba#061	3	A1	Ba#064	41	A1
Ba#004	4	A1	Ba#065	42	A1
Ba#036	5	A1	Ba#066	43	A1
Ba#037	5	A1	Ba#067	44	A1
Ba#018	6	B	Ba#068	45	A1
Ba#031	7	A1	Ba#069	46	A1
Ba#001	8	A2	Ba#070	47	A1
Ba#002	9	A1	Ba#071	48	A1
Ba#005	10	A1	Ba#072	49	A1
Ba#007	11	B	Ba#073	50	A1
Ba#008	12	B	Ba#076	51	B
Ba#009	13	A1	Ba#077	52	A1
Ba#010	14	A1	Ba#078	53	A1
Ba#011	15	A1	Ba#079	54	A1
Ba#012	16	A1	Ba#080	55	A1
Ba#013	17	A1	Ba#092	56	A3
Ba#014	18	A1	Ba#093	57	A1
Ba#015	19	A1	Ba#094	58	A1
Ba#016	20	A1	Ba#095	59	A1
Ba#017	21	B	Ba#096	60	A1
Ba#021	22	A1	Ba#097	61	A1
Ba#022	23	B	Ba#098	62	A1
Ba#023	24	B	Ba#099	63	A3
Ba#024	25	A1	Ba#102	64	A1
Ba#026	26	A1	Ba#103	65	A1
Ba#027	27	A1	Ba#105	66	A3
Ba#032	28	B	Ba#106	67	A1
Ba#033	29	B	Ba#108	68	A3
Ba#034	30	B	Ba#109	69	A3
Ba#038	31	A1	Ba#110	70	A1
Ba#039	32	A1	Ba#111	71	A3
Ba#043	33	A1	Ba#112	72	A1

4.3.4 The MLVA31 panel

The MLVA31 panel is composed of all the panels used up to date (i.e. MLVA8 + MLVA15 + MLVA25), where the MLVA8 loci are common to all the panels and the MLVA15 and 25 panels also have the Bams 1 locus in common (Table 4.1). The panels were designed using different separation methods and therefore protocols are being developed to standardise the use of reagents and technique in the characterization of isolates. A comparison of the discriminative power of the four MLVA systems, using 8, 15, 25, and 31-markers respectively clearly showed the superiority of the 31-marker MLVA (Figure 4.1). This was also indicated by Beyer & Turnbull (2009) that found amongst 391 isolates the resolution was 78, 98, 139 and 158 genotypes using MLVA8, 15, 25 and 31, respectively.

In our study the *B. anthracis sensu lato* isolates (113) produced 80 genotypes and grouped *B. cereus* as outgroup (Table 4.5 and Figure 4.1). On comparison of the other MLVA panels, it is noteworthy that the MLVA25 panel and the MLVA31 panel are almost identical except for the additional 7 genotypes exposed (Figure 4.1). This again, only contributes toward resolution of the shallow branches within lineages. It raised the question about the necessity of using such large panels. The resolution offered by the additional markers is only marginally greater and could be considered superfluous. VNTR panels should be streamlined to provide adequate information for epidemiological studies taking cost and labour hours into consideration. Keim *et al.* (2000) indicated that “it is not appropriate to simply utilize the most highly diverse markers for every analysis, as high levels of genetic diversity can also confound phylogenetic patterns.”

Table 4.5: Genotypes of *B. antracis sensu lato* isolates identified using the MLVA31 panel

Isolate	genotype	clade	Isolate	Genotype	clade	Isolate	genotype	clade
Ba#104	1	B1	Ba#035	27	A1	Ba#103	54	A1
Ba#041	2	A1	Ba#038	28	A1	Ba#091	55	A1
Ba#048	3	A1	Ba#036	29	A1	Ba#060	56	A1
Ba#073	4	A1	Ba#037	29	A1	Ba#001	57	A2
Ba#032	5	B	Ba#086	30	A1	Ba#006	57	A2
Ba#033	5	B	Ba#084	31	A1	Ba#108	58	A1
Ba#031	5	A1	Ba#079	31	A1	Ba#092	59	A1
Ba#030	5	A1	Ba#077	31	A1	Ba#089	60	A1
Ba#022	6	B	Ba#064	32	A1	Ba#113	61	A3
Ba#023	7	B	Ba#050	33	A1	Ba#044	62	A2
Ba#007	8	B	Ba#066	34	A1	Ba#045	62	A2
Ba#040	9	B	Ba#065	35	A1	Ba#090	63	A1
Ba#008	10	B	Ba#070	35	A1	Ba#093	64	A1
Ba#018	11	B	Ba#071	36	A1	Ba#099	65	A3
Ba#101	12	A1	Ba#068	37	A1	Ba#096	66	A1
Ba#017	12	B	Ba#063	37	A1	Ba#112	67	A1
Ba#083	13	A1	Ba#062	37	A1	Ba#110	68	A1
Ba#074	14	A1	Ba#072	37	A1	Ba#109	69	A3
Ba#075	15	A1	Ba#069	37	A1	Ba#102	70	A1
Ba#076	16	B	Ba#047	38	A1	Ba#111	71	A3
Ba#078	17	A1	Ba#004	39	A1	Ba#100	72	A3
Ba#080	17	A1	Ba#009	40	A1	Ba#106	73	A1
Ba#082	17	A1	Ba#003	41	A1	Ba#107	74	A3
Ba#055	18	A1	Ba#005	42	A1	Ba#098	75	A1
Ba#056	18	A1	Ba#010	43	A1	Ba#029	76	A1
Ba#019	19	A1	Ba#046	44	A1	Ba#034	77	A1
Ba#049	20	A1	Ba#039	45	A1	Ba#026	78	A1
Ba#057	21	A1	Ba#012	45	A1	Ba#011	79	A1
Ba#058	21	A1	Ba#043	45	A1	Ba#44	80	A3
Ba#095	22	A1	Ba#015	45	A1	Ba#45	80	A3
Ba#002	23	A1	Ba#016	45	A1			
Ba#004	24	A1	Ba#061	46	A1			
Ba#027	24	A1	Ba#054	47	A1			
Ba#028	24	A1	Ba#051	48	A1			
Ba#021	24	A1	Ba#053	49	A1			
Ba#025	24	A1	Ba#052	50	A1			
Ba#094	25	A1	Ba#087	51	A1			
Ba#097	26	A1	Ba#088	52	A1			
Ba#105	26	A3	Ba#085	53	A1			

4.3.5 Diversity and Distribution

The MLVA 31 VNTR panel produced the greatest resolution in terms of strain typing and was therefore used to demonstrate the genotype distribution for this study. Figure 4.2 shown here, illustrates that *B. anthracis sensu lato* clusters in lineages A and B as designated by Keim *et al*. (2000). The A lineage consist of the greater proportion of isolates and is further sub-divided. The first group (labelled 1) is made up of the Ba#001 and Ba#006 that also grouped separately using the MLVA 15 and 25 (Figure 4.1), Northern Cape isolates mainly from Kimberley, as well as, Rondebosch isolates (Ba#44 and 45) and isolate Ba#59 that clustered similarly using the MLVA15 and 25 panel. The second subclade consists of isolates mainly from the KNP and border regions (sub clustered together within the subcluster). The B clade included KNP isolates mostly from the northern parts of KNP (Figure 4.3) and the attenuated strain Ba#104 that clustered separately (Figure 4.1).

It is believed that soil conditions contribute to the persistence of anthrax spores in certain areas. Soils rich in organic matter and calcium that are alkaline in nature are considered ideal for spore survival (Hugh-Jones & De Vos, 2002; Hugh-Jones & Blackburn, 2009). In the KNP, Smith *et al.* (2000) indicated areas with soil calcium greater than 150 milliequivalents and a pH higher than 7.0 had anthrax death rates in the wildlife more than seven times higher than for those areas with lower soil values. In fact, Smith *et al.* (2000) determined that there are 2 independent foci areas in KNP where epidemics occurred. Although outbreaks occur almost throughout the park, the central region and the most

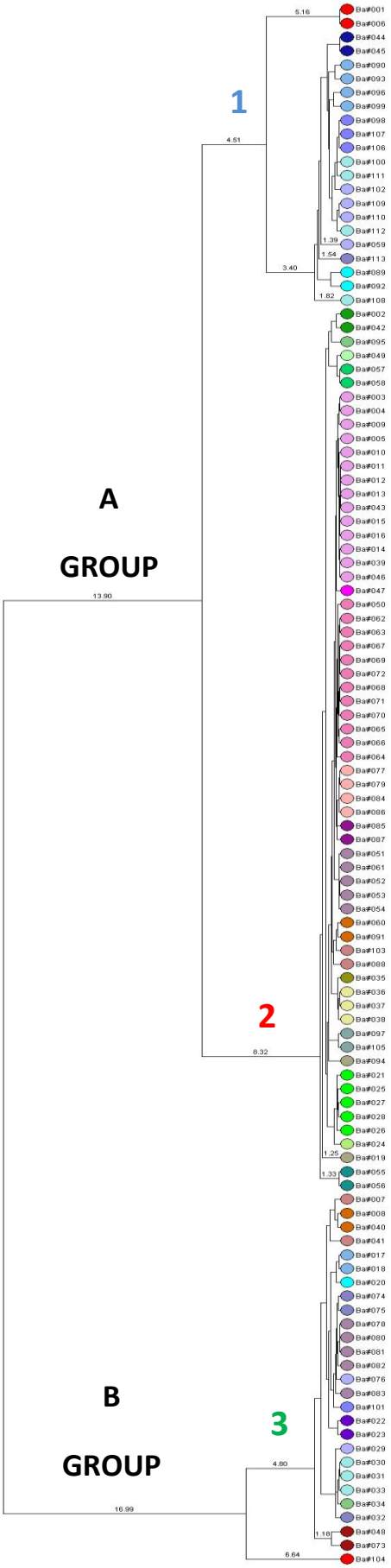


Figure 4.2: A maximum likelihood tree of the 113 *B.anthracis* strains for the MLVA 31 panel performed in Bionumerics version 6.0.

northern part of the park are most notable for anthrax cases. In the 2 foci areas, cluster analysis of MLVA8 panel indicated that clade A isolates were found throughout the KNP and also globally, whereas the B group isolates were only found in the northern part of KNP. Therefore, both clade A and B occur in the northern parts of KNP, while clade A isolates occur in the southern parts of KNP. The cluster patterns of the A and B clade isolates in KNP can be attributed to factors in the environment like soil pH and soil calcium content according to Smith *et al.* (2000) (Figure 4.4). In the KNP, clade A isolates had a wider geographical range and greater genetic diversity than the clade B isolates. Both A and B clade isolates occur in the larger overlap of the foci regions of the KNP (Smith *et al.,* 2000). Smith *et al.* (2000) and our study's results is consistent as MLVA cluster analysis indicate that the A and B clade are genetically dissimilar, where clade A isolates were found globally and clade B isolates are mainly restricted to southern Africa (Figure 4.3). The distribution of the A and B clade isolates is therefore believed to be due to an adaptive ability amongst the clade A isolates that is lacking in the branch B isolates (Smith *et al.,* 2000). It is this genotypic diversity which led Keim *et al.* (2000) to hypothesise that anthrax originated in southern Africa.

4.3.6 Temporal Distribution

The isolates typed in our study span 4 decades of collection from 1970 to 2009 (refer to Appendix Table 1). It is a general belief that due to the stable nature of anthrax spores, that many isolates originating from the same outbreak will have identical or a dominant genotype (Keim *et al.,* 2000; Hugh-Jones & De Vos, 2002; Lista *et al.,* 2006). Furthermore it has been observed, that there is a preferential species infection according to regions (De Vos, 1990). Hence cluster analysis was performed separately according to the year and animal species to determine any trends. Unfortunately, these analyses failed to provide any insight into any outbreak trends (data not shown). Even for the year 1990 that had the largest sample size (n=40), cluster and genotypic analysis was unable to highlight any 1 predominant genotype. The outbreak affected both kudu and buffalo equally. The only noteworthy outcome of the cluster analysis for that year is that the isolates belonged almost exclusively to the A clade of isolates (Figures 4.2 & 4.3) as also indicated by Smith *et al.* (2000).

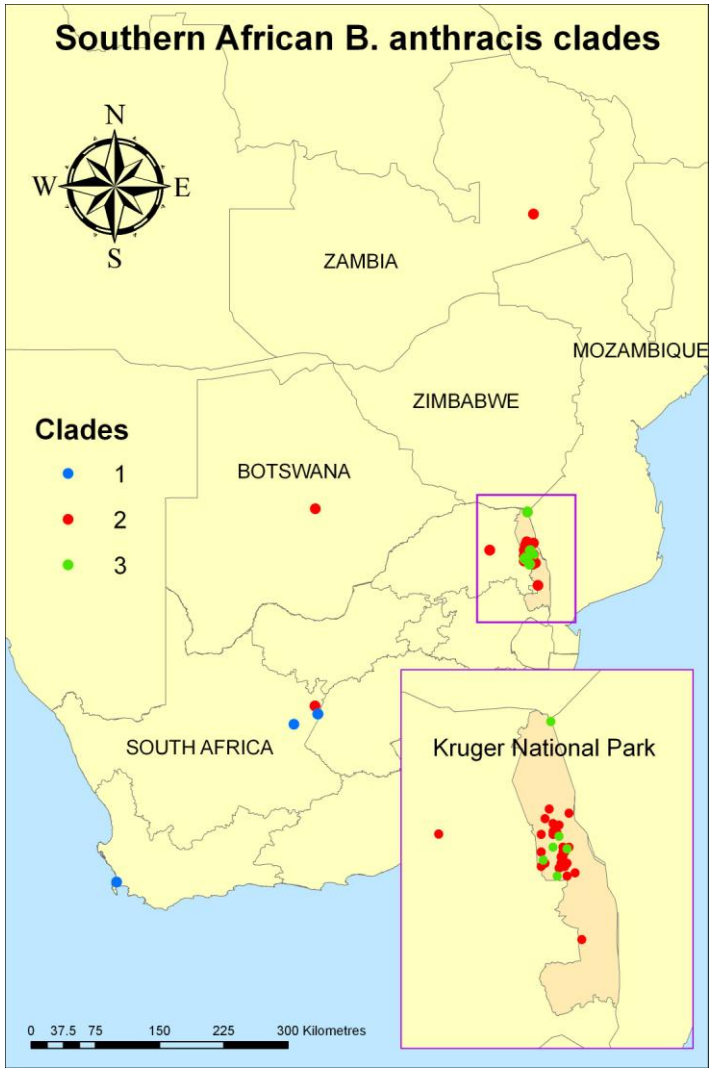


Figure 4.3: Distribution map of isolate collection sites. Clades 1 and 2 represent the A group isolates while clade 3 represents the B group isolates. Clade 1 is found almost exclusively in the Northern Cape, while clade 3 (B group) is found in the northern parts of Kruger National Park.

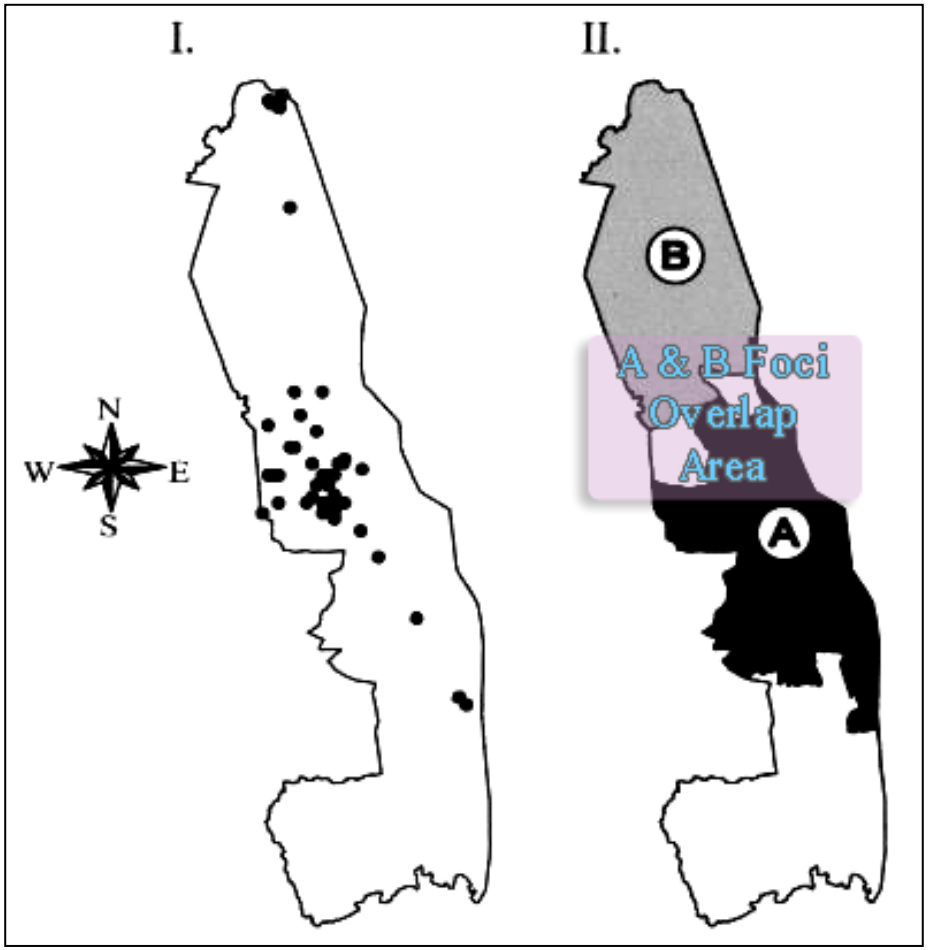


Figure 4.4: This map of the Kruger National Park (KNP) illustrated by Smith *et al.* (2000), where (I) shows the individual isolate locations within the KNP used for their study and (II) shows their findings for distribution of the A and B clade anthrax isolates using Bernoulli regression for space-time clustering.

Smith *et al*. (2000) attributed the paucity of B clade isolates to possible unusual rainfall patterns in that year. These authors were also unable to find a definitive trend from outbreaks between 1970 and 1981 stating that "this is in contrast to the expectation that a single outbreak will emanate from a single source and that pathogen isolates will have great similarity or even be indistinguishable" (Smith *et al.,* 2000).

While no obvious outbreak trend was discernible, studying some of the identical and/or genetically similar isolates provides an interesting narrative. For example Ba#001 was isolated downstream from the Gaisenga (KNP) in 1975 and Ba#006 was isolated from a hippo bone obtained at an unknown location in 1990 belonged to the same genotype and might indicate that the spores originated from the same location. Similarly, with the identical genotypic isolates Ba#044 and Ba#045 that originate from a dairy farm in Rondebosch which was shut down due to a suspected anthrax outbreak when the dairy was extended.

4.3.7 Allelic Frequencies

Genetic diversity is a compound measure that includes the number of allelic states as well as their frequency distribution within the population. It is affected by four processes: mutation, selection, genetic drift and recombination (Le Fleche *et al*., 2001). The mechanism that likely generates mutations in VNTR loci is slipped strand mis-pairing. These insertion-deletion (indel) mutations occur at varied rates in VNTR loci, ranging from less than 10^5 to greater than 10^4 in *B. anthracis* and exceeding 10^3 of other bacteria. This wide range of mutation rates across VNTR loci could probably be attributed to intrinsic, locus-specific properties, including length of the tandem array (Vergnaud & Denoeud, 2000; Keim *et al.,* 2004 Lista *et al.,* 2006).

The discriminatory power of each MLVA marker can be estimated by the number of alleles it is able to detect and by its diversity (Keim *et al.,* 2000). The closer the diversity value is to 1, the greater its diversity. For the frequencies in our study, the average was 0.658 for the agarose separation and 0.649 for the capillary separation. The VNTR loci with the highest diversity were Bams 13 with 0.923; Bams 30 with 0.917; Bams 31 with 0.872 and Bams 3 with 0.858. Bams 13; 30 and 31 all contain a 9 bp tandem repeat units while Bams 3 consists of a 15 bp tandem repeat unit. Besides their diversity, it was also discussed earlier that the

Bams 13 and Bams 30 code for a gene that is unique to anthrax spore survival. This necessitates the inclusion of these VNTR in any typing panel.

The pXO2 also had a high diversity index of 0.842 along with VNTR 19 with 0.809 that contain 2 and 3 bp tandem repeat units respectively. However, the allele frequency is influenced by the lack of plasmid(s) in the *B. cereus* and 6 attenuated *B. anthracis sensu lato* strains. VNTR loci with a low diversity and that can be considered for exclusion from future panels are: Bams 5 with 0.183, Bams 25 with 0.135, Bams 51 with 0.398 and Bams 24 with 0.288 since loci that cause 'noise' or extend genotypes should be excluded as indicated by Keim *et al*. (2000).

4.4 Conclusion

MLVA has proven to be a valuable tool in the epidemiological study of *B. anthracis.* All 4 of the MLVA panels were able to offer similar clustering of the isolates but the larger panels correlated more towards the geographic origin of the outbreak. As was expected, the larger the VNTR panel, the greater the resolution offered. An unexpected result was that the MLVA15 panel was unable to sufficiently differentiate between *B. anthracis* and *B. cereus*, which makes it less ideal for genotyping purposes.

The MLVA31 provided the most genotypic differentiation but the cluster analysis was nearly identical to that of the MLVA25 panel. Future research should investigate whether the additional VNTR could be superfluous or the MLVA31 panel can be reduced by excluding low frequency VNTR loci. However, with the latter the fact that very high levels of genetic diversity could also cause confusion and make trends difficult to spot should be kept in mind. For the purpose of epidemiological studies, a more streamlined panel which is just as capable of tracing an outbreak back to its source is better suited in terms of cost effectiveness as well as time saving.

This study did not provide any insight into the predominance of a single genotype in an outbreak or demonstrate any correlation to preferential species infection. This could be because the sample set contained too many variables. The samples spanned 4 decades of outbreaks and a multitude of sources which made finding trends difficult. Perhaps a large sample set from a single outbreak could provide more meaningful and statistically significant findings regarding the preferential infection of animal species and the predominance of particular genotypes within an outbreak.

References

Andersen, G. L., Simchock, J. M., & Wilson, K. (1996). Identification of a region of genetic variability among *Bacillus anthracis* strains and related species. *Journal of Bacteriology, 178*(2)

Beyer, W., & Turnbull, P. C. (2009). Anthrax in animals. *Molecular Aspects of Medicine, 30*(6), 481-489.

De Vos, V. (1990). The ecology of anthrax in the Kruger National Park, South Africa. *Salisbury Medical Bulletin,* 68S: 19-23.

Denoeud, F., & Vergnaud, G. (2004). Identification of polymorphic tandem repeats by direct comparison of genome sequence from different bacterial strains: A web-based resource. *BMC Bioinformatics, 5*, 4.

Fortini, D., Ciammaruconi, A., Santis, R. D., Fasanella, A., Battisti, A., D'Amelio, R., *et al.* (2007). Optimization of high-resolution melting analysis for low-cost and rapid screening of allelic variants of *Bacillus anthracis* by multiple-locus variable-number tandem repeat analysis. *Clinical Chemistry, 53*(7), 1377-1380.

Hugh-Jones, M., & Blackburn, J. (2009). The ecology of *Bacillus anthracis*. *Molecular Aspects of Medicine, 30*(6), 356-367.

Hugh-Jones, M. E., & De Vos, V. (2002). Anthrax and wildlife. *Revue Scientifique Et Technique (International Office of Epizootics), 21*(2), 359-383.

Jackson, P. J., Walthers, E. A., Kalif, A. S., Richmond, K. L., Adair, D. M., Hill, K. K., *et al.* (1997). Characterization of the variable-number tandem repeats in vrrA from different *Bacillus anthracis* isolates. *Applied and Environmental Microbiology, 63*(4), 1400-1405.

Keim, P., Pearson, T., & Okinaka, R. T. (2008). Evolution of *Bacillus anthracis*, causative agent of anthrax. *Evolutionary Biology of Bacterial and Fungal Pathogens,* 523-533.

Keim, P., Price, L. B., Klevytska, A. M., Smith, K. L., Schupp, J. M., Okinaka, R., *et al.* (2000). Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis. Journal of Bacteriology, 182*(10), 2928-2936.

Keim, P., Van Ert, M. N., Pearson, T., Vogler, A. J., Huynh, L. Y., & Wagner, D. M. (2004). Anthrax molecular epidemiology and forensics: Using the appropriate marker for different evolutionary scales. *Infection, Genetics and Evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases, 4*(3), 205-213.

Klee, S. R., Ozel, M., Appel, B., Boesch, C., Ellerbrok, H., Jacob, D., *et al.* (2006). Characterization of *Bacillus anthracis*-like bacteria isolated from wild great apes from Cote d'Ivoire and Cameroon. *Journal of Bacteriology, 188*(15), 5333-5344.

Le Fleche, P., Hauck, Y., Onteniente, L., Prieur, A., Denoeud, F., Ramisse, V., *et al. (*2001). A tandem repeats database for bacterial genomes: Application to the genotyping of *Yersinia pestis* and *Bacillus anthracis. BMC Microbiology, 1*, 2.

Leski, T. A., Caswell, C. C., Pawlowski, M., Klinke, D. J., Bujnicki, J. M., Hart, S. J., et al. (2009). Identification and classification of bcl genes and proteins of *Bacillus cereus* group organisms and their application in *Bacillus anthracis* detection and fingerprinting. *Applied and Environmental Microbiology, 75*(22), 7163-7172.

Lindeque, P. M., & Turnbull, P. C. (1994). Ecology and epidemiology of anthrax in the Etosha National Park, Namibia. *Onderstepoort J.Vet.Res., 61*, 71.

Lista, F., Faggioni, G., Valjevac, S., Ciammaruconi, A., Vaissaire, J., le Doujet, C., *et al.* (2006). Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. *BMC Microbiology, 6*, 33.

Okinaka, R. T., Henrie, M., Hill, K. K., Lowery, K. S., Van Ert, M., Pearson, T., et al. (2008). Single nucleotide polymorphism typing of *Bacillus anthracis* from Sverdlovsk tissue. *Emerging Infectious Diseases, 14*(4), 653-656.

Smith, K. L., De Vos, V., Bryden, H., Price, L. B., Hugh-Jones, M. E., & Keim, P. (2000). *Bacillus anthracis* diversity in Kruger National Park. *Journal of Clinical Microbiology, 38*(10), 3780-3784.

Strauss, L. (2010). Kruger Park News: Anthrax outbreak in Kruger. http://www.krugerpark.co.za/krugerpark-times-e-6-anthrax-outbreak-in-kruger-25276.html

Van Ert, M. N., Easterday, W. R., Huynh, L. Y., Okinaka, R. T., Hugh-Jones, M. E., Ravel, J., *et al.* (2007). Global genetic population structure of *Bacillus anthracis*. *PLoS ONE, 2*(5), e461.

Vergnaud, G., & Denoeud, F. (2000). Minisatellites: Mutability and genome architecture. *Genome Research, 10*(7), 899-907.

CHAPTER 5:

Preliminary evaluation of fixed blood smears as a source of DNA for use in the fingerprinting of *Bacillus anthracis*

Abstract

Bacillus anthracis, the causative agent of anthrax, is diagnosed in the field through observation of clinical symptoms and blood smears. Most often this is the only available sample of a reported anthrax case in developing countries. We therefore investigated the use of blood smears as a DNA source for further analsysis using a molecular typing method to examine strain dynamics among and within these outbreaks and epidemiology. In this study, 14 stained blood smear slides were evaluated as a source of DNA for the fingerprinting of anthrax using multilocus variable number of tandem repeat (VNTR) analyses (MLVA). Of the 14 stained blood smears, 4 were controls of which 1 was a freshly made smear from spiked blood and the rest were stained blood smear slides from positively identified anthrax clinical cases. Typing using MLVA31 panel was only successful from the freshly made stained blood smear slides. The remaining control slides could only be typed once subjected to GenomiPhi (amplifying of genome). The test slides only amplified markers smaller than 300 bp. Factors were identified that influenced fingerprinting / typing. The most significant were found to be the small amount of target DNA (anthrax spores) from one blood smear slide and conditions of collection and storage. The small target DNA was overcome by amplifying genomic DNA in the sample using GenomiPhi. However, the conditions of collection, storage and DNA isolation method are important factors that greatly influence typing and further investigation is required to determine the effect of each on the value of archival samples in fingerprinting and epidemiological studies.

5.1 Introduction

The diagnosis of anthrax, caused by *Bacillus anthracis,* requires the demonstration of encapsulated *B. anthracis* in smears of blood or tissues and growth of the organism on blood agar plates. Stained blood smears are commonly used as a diagnostic method in the initial confirmation of anthrax in the case of a suspicious death (Turnbull, 1998). As anthrax is a controlled animal disease in terms of the Animal Diseases Act (Act 35 of 1984) in South Africa, samples from 'suspect' animals should be sent to a reference laboratory (Agriculture Research Council - Onderstepoort Veterinary Institute (ARC-OVI), Onderstepoort, South Africa) equipped to comply with the OIE standard for identification for anthrax by means of culture.

On microscopic examination of a blood smear, confirmation of the presence of *Bacillus* can be obtained on observing Gram-positive rods. The cells are found in pairs or short chains often with square ends giving them a "bamboo" or 'box-car' (set of railway carriages) appearance. As the Gram stain is not suitable to visualize the capsulated *B. anthracis,* smears of the specimen must be stained with M'Fadyean's stain (polychrome methylene blue) that stain the capsule pink and the bacillus cells dark blue (Turnbull *et al.,* 1998; OIE, 2008). The diagnosis of anthrax in Africa / developing countries is predominantly achieved via microscopy to detect encapsulated *B. anthracis* with M'Fadyean's stain in smears of blood or tissue from fresh carcasses. The stained slides are stored (archived) and do not represent a hazard to possible infection (Amar *et al.*, 2002).

There have been several studies on the use of archival samples as a source of DNA for use in molecular diagnostics (Poljak *et al.,* 1995; Yokota *et al.,* 1995; Alger *et al.,* 1996; Amar *et al.,* 2001; Motazedian *et al.,* 2001; Amar *et al.,* 2002). These have met with differing levels of success depending on the target organism and the conditions of collection and storage for isolation and amplification of DNA from the samples. PCR has also been used in the downstream analysis of these samples, which normally become redundant after completion of the initial microscopic diagnosis (Poljak *et al.,* 1995; Yokota *et al.,* 1995; Alger *et al.,* 1996; Amar *et al.,* 2001; Motazedian *et al.,* 2001; Amar *et al.,* 2002). Yap *et al.* (1991) reported amplification directly on the slide. The use of PCR on archived material on stored slides are worth exploring in anthrax diagnoses where only stored slides are available and can be used

to offer information and make retrospective studies possible. In this study we first investigated the use of stained blood smears as DNA template to amplify the *Bacillus* specific 16S rRNA region with PCR. We then attempted to use 31 MLVA panel for strain typing that would enable molecular typing to provide more information regarding the origin of the outbreaks and epidemiology.

5.2 Materials and Methods

Test samples consisting of the Sterne vaccine strain were cultured and mixed with horse blood, fixed and Giemsa stained for use as a control for the assays. Also included were three clinical cases of stained smears from the Onderstepoort Veterinary Academic Hospital, South Africa (Caius; Sophie and Open-6-Giemsa) to test the viability of the DNA extraction and amplification techniques on old (several years) samples. Ten stained blood smears on microscope slides from anthrax field cases were obtained from the Central Veterinary Laboratory in Harare, Zimbabwe. The stained blood smears on microscope slides arrived with immersion oil present that started to dry. The sample details and related history were not provided.

The Giemsa and/or polychrome methylene blue stained blood smear slides were examined by microscopy and levels of capsulated *B. anthracis* were subjectively recorded for each smear. The blood smear on the microscope slide was scraped into a collection plate and transferred into a 1.5ml centrifuge tube using a sterile scalpel. The smear scraping was added to 180 µl ATL buffer. Lysozyme was then added in a concentration of 20mg/ml to maximise the DNA yield. The sample was incubated at 37°C for at least 30 minute. Thereafter the QIAGEN QIAmp® DNA Blood Mini kit was used and the manufacturer's instruction for DNA extraction from Giemsa stained smears was followed.

The 16S rRNA PCR was used to test for the presence of *Bacillus* DNA and for degradation as it amplified a large fragment of approximately 1680 bp. The 16S rRNA gene was amplified using 2.5µl DNA in a final volume of 15µl containing 1 X PCR ExSel High Fidelity reaction buffer (JMR Holdings), 2mM $MgSO_4$, 0.2mM dNTP, 0.4µM of each primer (67F-5'TGAAAACTGAACGAAACAAAC3' and 1671R- 5'CTCTCAAAACTGAACAAAACG AAA 3') (Sacchi *et al.,* 2002) and 0.125U *Taq* Polymerase (JMR Holdings). The PCR conditions consisted of

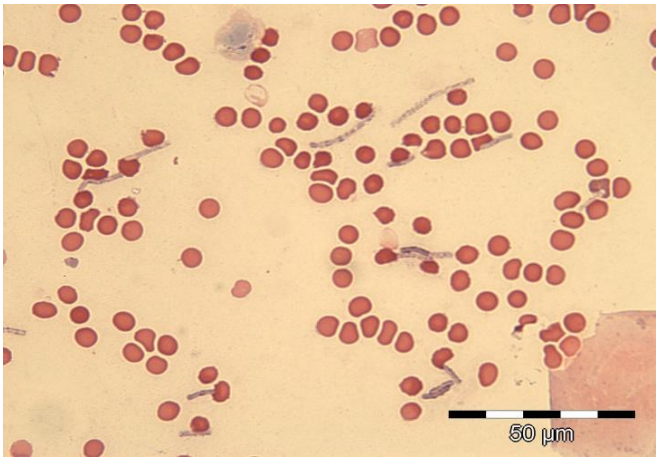
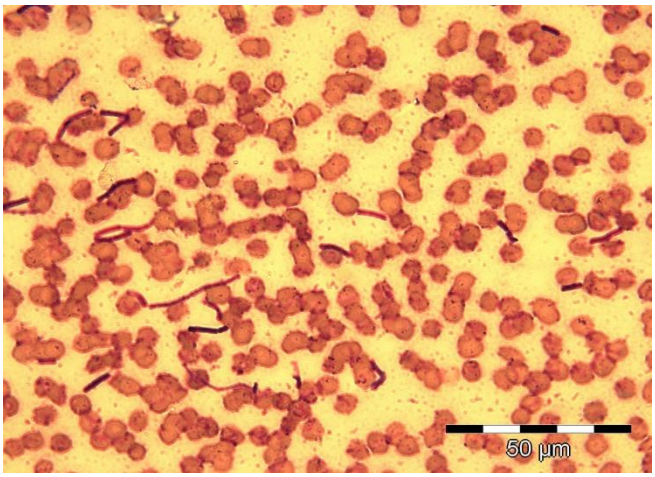
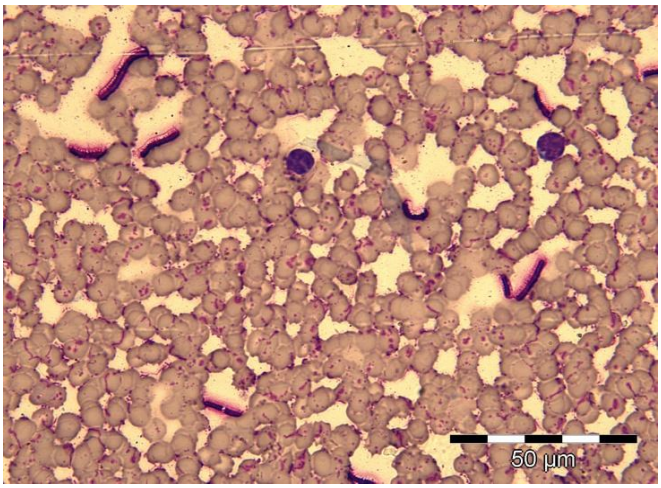
initial 95°C incubation for 5 minutes, followed by 35 cycles of 94°C for 15 seconds, 52°C for 15 seconds and 72°C for 90 seconds with a final elongation step at 72°C for another 5 minutes in an ABI Thermocycler 9700. As no PCR product was obtained a secondary PCR was done using 2.5 µl of primary PCR product using PCR reaction and conditions as indicated above. Amplification products were separated by electrophoresis using a 2% agarose gel and visualised by UV-light following ethidium bromide staining.

For PCR using VNTR markers (MLVA31) the GenomiPhi® Kit (AEC Amersham) was used in an attempt to increase the genomic DNA. Monoplex PCR of each VNTR locus was performed, as mentioned previously, in Chapter 2 section *2.2.3.1* using 4 µl of the GenomiPhi product instead of the original DNA. The amplification products were separated using gel electrophoresis as previously mentioned in section *2.2.4.1.* The multiplex PCR, where capillary electrophoresis is used to determine amplification product sizes, was performed, as mentioned previously, in section *2.2.3.2* using DNA extracted from smear and separated as previously mentioned in section *2.2.4.2.*

5.3 Results

The blood smear control samples as well as field samples from Zimbabwe were microscopically examined. The blood smear using blood spiked with the Sterne vaccine strain had a high level of bacilli by microscopy (data not shown) and the three anthrax positive blood smears from the Onderstepoort Hospital had low levels of bacilli (Table 5.1). Most of the field samples had low to medium levels of bacilli detected by microscopy (Table 5.1). However, a few field sample blood smears could not be examined using microscopy due to dried immersion oil on the slide.

Table 5.1: Microscopic images of stained blood smears of microscope slides indicating smear quality and level of Bacilli presence

Slide	ID#	Bacilli presence	Comments
	Sophie	Low	Unknown stain Stored in good condition
	Open 6 Giemsa	Low	Giemsa stained. Stored in good condition
	Caius	Low	Diff-Quick Stained. Stored in good condition

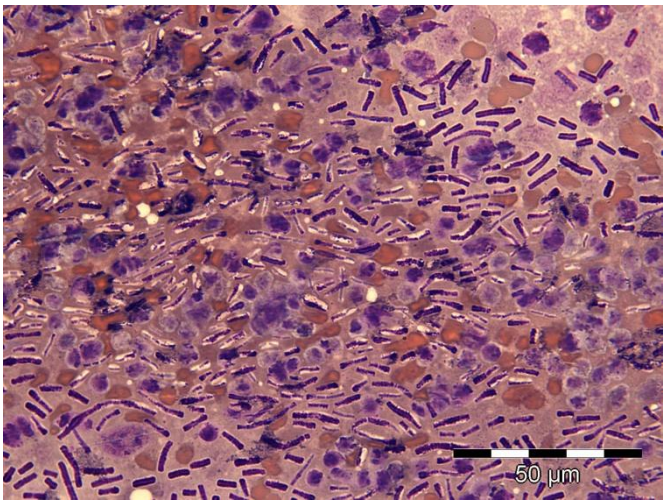
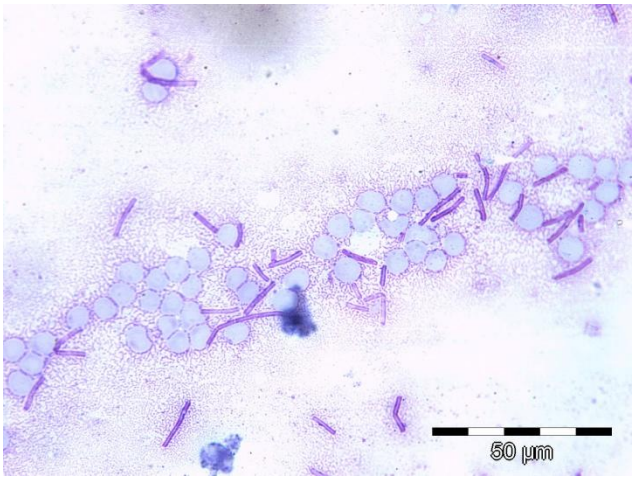
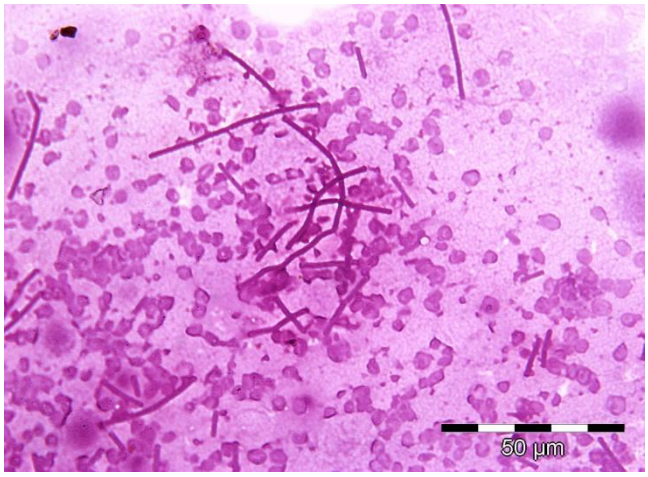
Slide	ID#	Bacilli presence	Comments
	Ba 1	High	Diff-Quick Stained. Blood smear was thick. Thickly coated in immersion oil and fibres
	anthrax	Medium	Diff-Quick stained. Slide was covered in dried mineral immersion oil
	1592	Medium	Long chain *Bacilli* were observed along with spores. Poorly Stored slide

Table 5.1: Continued...

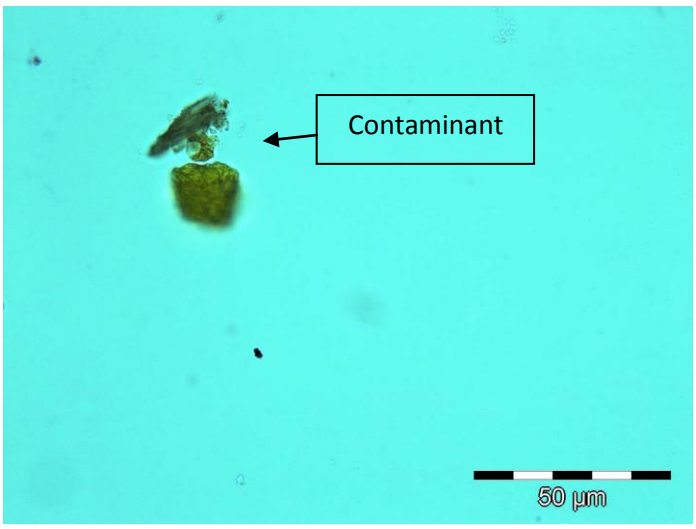
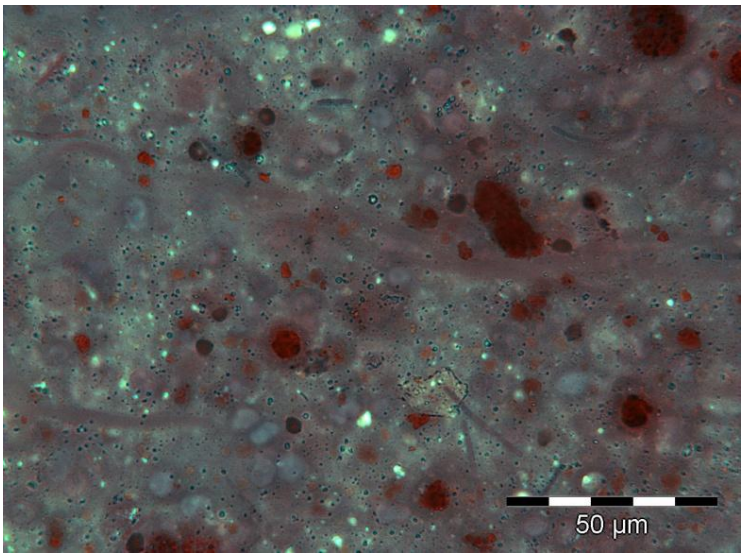
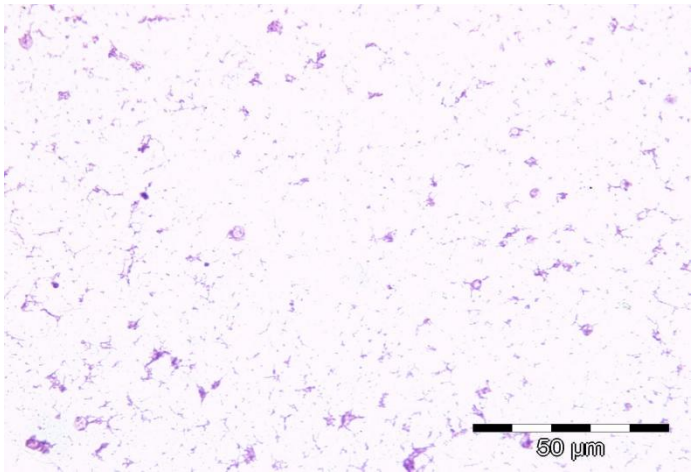
Slide	ID#	Bacilli presence	Comments
	1572	Assumed to be positive from field records	The slide was covered in embedded fibres and debris which obscured view
	1577	Low	The blood was thickly smeared. Slide was covered in dried mineral immersion oil
	1565	Medium	It was possible to observe the *Bacilli*, but the slide was thickly covered in dried mineral immersion oil

Table 5.1: Continued...

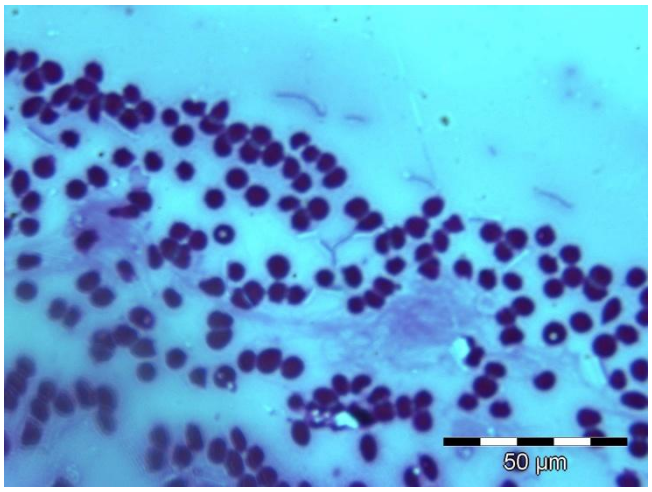
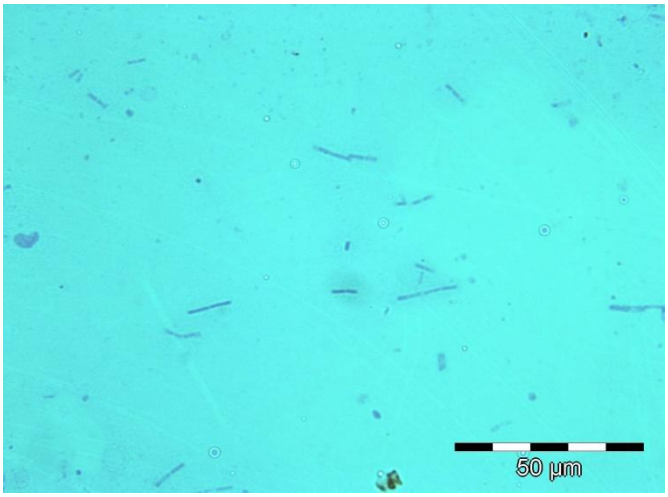
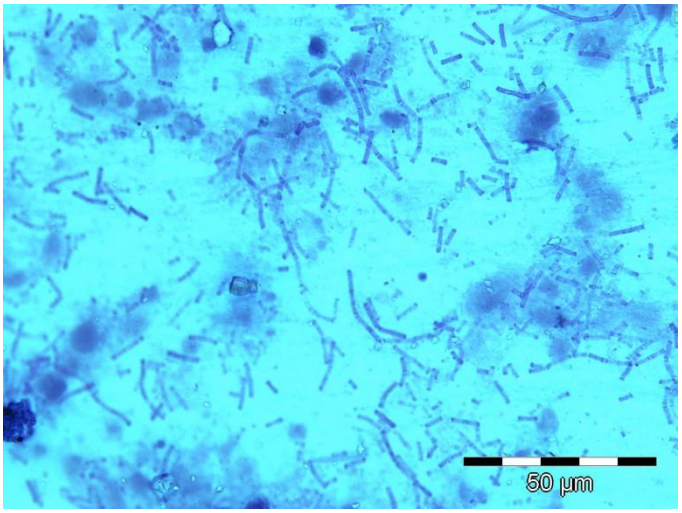
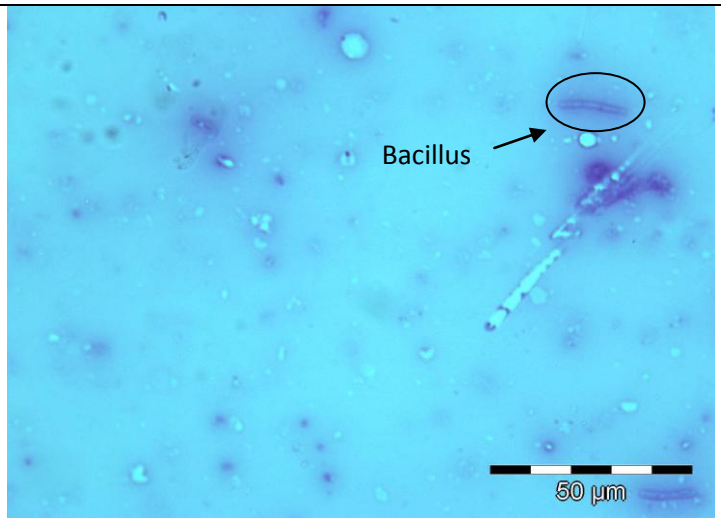
Slide	ID#	Bacilli presence	Comments
	1583	Low	Diff-Quick stained. Covered in dried mineral immersion oil
	1562	Low	Diff-Quick stained Covered in dried mineral immersion oil
	0125	High	Diff-Quick stained Covered in dried mineral immersion oil

Table 5.1: Continued...

Slide	ID#	Bacilli presence	Comments
	1584	Low	Diff-Quick stained Covered in dried mineral immersion oil

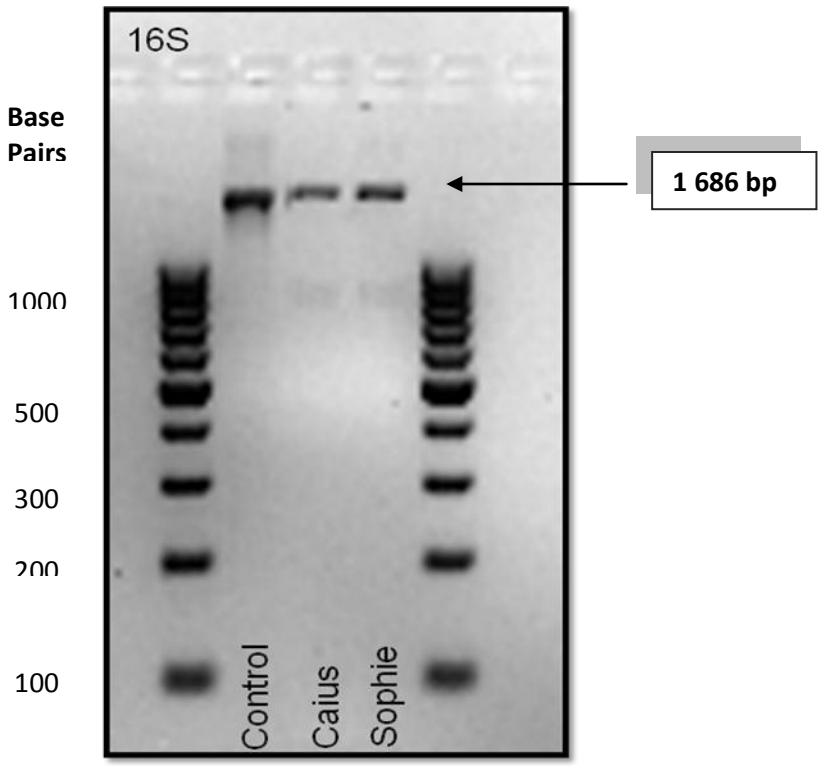


Figure 5.1: Agarose gel image showing bands for 16S rRNA gene for anthrax control samples.

The 16S rRNA PCR using DNA from control blood smear samples (blood spikes with $34F_2$ Sterne vaccine strain and 3 anthrax-positive blood smears obtained from the Onderstepoort Veterinary Hospital, South Africa) did not produce a visible PCR product. A secondary PCR was done and the Sterne blood smear control produced a ~1680 bp PCR product. None of the other samples produced a PCR product using secondary PCR. It was assumed that the *B. anthracis* DNA was too low and therefore the 3 control samples from Onderstepoort Veterinary Hospital (Cauis, Open-6-Giemsa and Sophie) were subjected to GenomiPhi® amplification to increase the genomic DNA. A primary and secondary PCR assay using the 16S rRNA primers was repeated on this DNA and provided visible bands when run on agarose (Figure 5.1). However, neither primary nor secondary 16S rRNA PCR performed on the field sample blood smear DNA and GenomiPhi DNA from Zimbabwe was able to produce a visible PCR product (Data not shown).

The 31 MLVA markers (Chapter 2: Table 2.1) were used to amplify each VNTR locus using GenomiPhi DNA of the control and field samples. The Sterne control sample, as well as Cauis and Sophie control samples had strong amplification signals for all 31 VNTR loci. For unknown reasons the Open-6-Giemsa sample did not amplify any of the 31 VNTR loci despite optimization attempts. The field samples amplified VNTR loci of 300 bp and smaller (CG3, vrrB1, vrr B2, pXO1, pXO2, Bams 44, VNTR 12 and VNTR19 as indicated in Table 2.2). Figure 5.2 shows the results of VNTR locus CG3 that amplified a 158 bp product containing 2 bp tandem repeat units for control and field samples. Less than than 8 VNTR markers for each sample tested were amplified and these loci were not consistently amplified in all the samples.

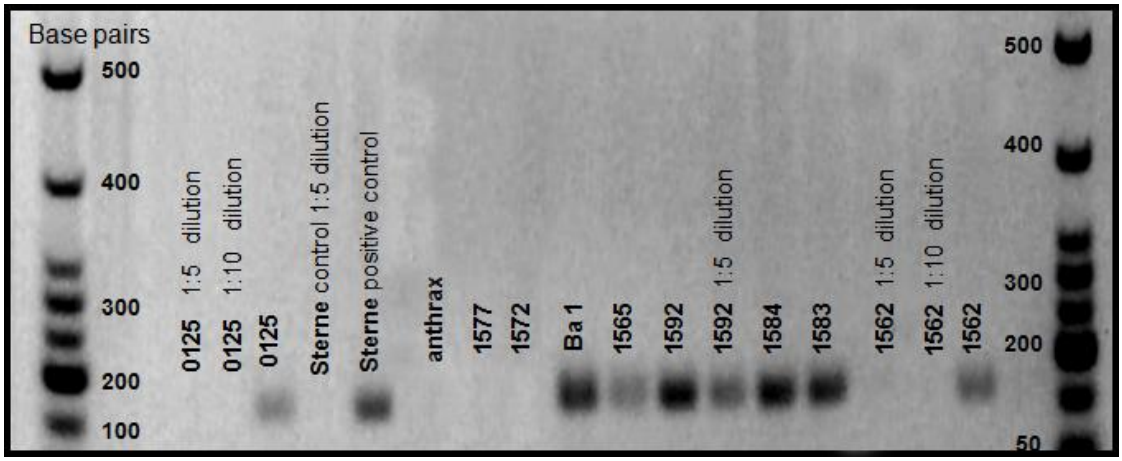


Figure 5.2: PCR products of CG3 VNTR locus amplifying 158 bp product that contains a 2 bp tandem repeat unit.

No VNTR locus with a fragment size greater than 300bp could be detected on automated DNA analyzer (capillary) despite using undiluted PCR products (data not shown). A variety of different size peaks were detected using VNTR loci amplified from GenomiPhi DNA , which made assigning a definitive peak (indicating a specific size) for a specific locus very difficult.

5.4 Discussion

It has been shown from previous studies that fixed blood smears are a good source of DNA for further study (Poljak *et al.,* 1995; Yokota *et al.,* 1995; Alger *et al.,* 1996; Amar *et al.,* 2001; Motazedian *et al.,* 2001; Amar *et al.,* 2002). This makes archived slides a "gold-mine" for retrospective studies and for the safe transportation of biological material. In this study we found that the *B. anthracis* DNA present on a blood smear in most cases needs to be amplified using Genomiphi (target DNA is insuffiecient or too low to produce visual PCR product) and that long term storage of blood smear slides results in degradation of DNA restricting the amplification to PCR products of less than 300bp. Therefore the MLVA31 panel cannot be used as only 8 of the 31 VNTR loci have PCR product sizes less than 300 bp (Table 2.1).

Diagnosis of anthrax in many countries in Africa is predominantly done with microscopy and samples in most cases are not sent to a central veterinary laboratory (reference laboratory) where suspected anthrax cases are confirmed by culturing. Therefore, often in African countries with limited resources, the only sample of anthrax cases is a fixed and stained blood smear. MLVA markers can be used to distinguish strains from different outbreaks, trace an outbreak strain back to its possible origin and/or to track the routes of transmission of an outbreak strain within and between animal populations. Therefore, we investigated the use of MLVA for DNA extracted from anthrax blood smears on microscope slides.

In this study we could use DNA from suspected anthrax, fixed and stained blood smear slides to amplify a large 16S rRNA fragment of *B. anthracis*. Amplification was only observed after secondary PCR and/or amplication of genome using GenomiPhi. Various factors that limit the application of MLVA fingerprinting using gel electrophoresis include (i) low DNA yield/target from the microscope slide, (ii) conditions of storage, and (iii) period of storage. The use of capillary electrophoresis that could detect lower amounts of PCR products was not sensitive enough to overcome low DNA template. Further analysis on blood smears such

as the strain typing method, MLVA, is not possible as only small VNTR loci can be amplified (smaller than 300 bp), small loci were not consistently amplified in all the samples making MLVA results insufficient to characterize strain. Most of the studies cited in this article showed successful amplification of fragments only up to 320 bp except Yap & McGee, (1991) and Alger *et al.* (1996), which demonstrated amplification of fragments up to 1 kb.

There are several determinants to be considered in the evaluation of the stained smears as a viable source of DNA for amplification using PCR in molecular diagnostics. These determinants appear to be: specimen quality; staining protocol; method employed for DNA isolation and the desired downstream applications (Poljak *et al.,* 1995; Yokota *et al.,* 1995; Alger *et al.,* 1996; Vince *et al.,* 1998; Farrar and Reboli, 2006).

Specimen quality certainly seems to play a large role in the success of the smears as a DNA source. In a study, the specimens were visually inspected for cell damage (which has implications for DNA degradation) to determine its worth for PCR (Vince *et al.,* 1998). The manner in which the specimens are stored and the length of storage are noteworthy. Though studies have demonstrated successful amplification in specimens stored in a metal box, in an air conditioned room for up to a decade (Alger *et al.,* 1996). Others have shown DNA stability for only up to 8 weeks at room temperature, un-mounted and with residual immersion oil, where this time period also included transport of the specimens by post (Amar *et al.,* 2002).

Alger *et al.* (1996) investigated the response during amplification based on the thickness of the smear. These authors found that the thicker smear shows a stronger response during amplification. This observation is common where the target organism is a parasite and the greater the red blood cell count, the greater the parasitaemia. It is a logical conclusion as thicker smears result in a higher DNA yield of the target organism. Similarly in our study, the amplification signal was more intense from thicker blood smear slides.

The method of fixing and staining also plays a role in the quality of DNA obtained from a slide. This is especially true for *B. anthracis,* where a longer heat fixing is recommended to inactivate the spore (Farrar and Reboli, 2006). Only negligible differences were observed in stained and un-stained smears, taken from the same patient (Poljak *et al.,* 1995). However, this was on freshly stained slides. Vince et al. (1998) could obtain amplification from sample

stored for 3 months and subsequently 9 months after staining. Furthermore, Yokota *et al.,* 1995 demonstrated that samples were only viable for up to 4 years after Giemsa staining.

Different DNA extraction protocols were used in the various studies. Crude boiling methods seem to produce inferior DNA yields to those that included Proteinase K (Motazedian *et al.,* 2001). A study was under-taken by Vince *et al.* (1998) to quantify the role of DNA isolation on amplification signals of stained smears. One of the extraction methods tested by Vince *et al.* (1998) was a modified QiAmp© protocol and another protocol included the addition of mineral oil to the sample to determine its effect. The Qiagen extraction method proved to have the best results for PCR with 100% amplification on all the samples. The isolation method which included the mineral oil had amplification for less than half of the samples tested. This seems to indicate that the mineral oil either prevents good quality DNA isolation or inhibits the PCR itself. This could be one explanation for the poor results obtained from the blood smears from Zimbabwe compared to the control sample specimens. The storage of the specimen and extraction method of anthrax blood smear needs to be further investigated for use in MLVA typing.

5.5 Conclusion

Freshly stained blood smears are a ready source of DNA for molecular assays. Archival blood smears can be used as a source of DNA for molecular studies, but the viability of the DNA is dependent on the treatment of the slide from the point of collection to the fixing and staining and thereafter the conditions of storage. The aforementioned conditions need to be further investigated for the successful application of the MLVA assay in the typing of *B. anthracis* from archival blood smears.

REFERENCES

Alger, J., Acosta, M. C., Lozano, C., Velasquez, C., & Labrada, L. A. (1996). Stained smears as a source of DNA. *Memorias do Instituto Oswaldo Cruz, 91*(5), 589-591.

Amar, C., Pedraza-Dı´az, S., & McLauchlin, J. (2001). *Cryptosporidium* genotyping from DNA extraction from a faecal smear on microscope slides stained by conventional techniques. *Journal of Clinical,* (39), 401-403.

Amar, C. F., Chalmers, R. M., Elwin, K., Tynan, P., & McLauchlin, J. (2002). Blinded evaluation of DNA extraction and genotyping of stained *Cryptosporidium* on glass slides. *Letters in Applied Microbiology, 35*(6), 486-488.

Farrar, W., & Reboli, A. (2006). In Dworkin M., Falkow S., Rosenberg E., Schleifer K. and Stackebrandt E.(Eds.), The genus *Bacillus—Medical* Springer New York.

Motazedian, H., Karamian, M., Noyes, H. A., & Ardehali, S. (2001). DNA extraction and amplification of *Leishmania* from archived, Giemsa-stained slides, for the diagnosis of cutaneous *Leishmaniasis* by PCR. *Annals of Tropical Medicine & Parasitology, Vol. 96, No. 1, 31–34*

OIE (2008). *Manual of diagnostic tests and vaccines for terrestrial animals*. Paris, world organisation for animal health (OIE).

Poljak, M., Barlic, J., Seme, K., Avsic-Zupanc, T., & Zore, A. (1995). Isolation of DNA from archival *Papanicolaou* stained cytological smears using a simple salting-out procedure. *Clinical Molecular Pathology,* (48: M55-M56)

Sacchi, C. T., Whitney, A. M., Mayer, L. W., Morey, R., Steigerwalt, A., Boras, A., *et al.* (2002). Sequencing of 16S rRNA gene: A rapid tool for identification of *Bacillus anthracis*. *Emerging Infectious Diseases, 8*(10), 1117-1123.

Turnbull, P. C. B. (Ed.). (1998). *WHO, EMC/ZDI/98 Guidelines for the surveillance and control of anthrax in humans and animals (fourth edition)*.

Vince, A., Poljak, M., & Seme, K. (1998). DNA extraction from archival Giemsa-stained bone marrow slides: Comparison of six rapid methods. *British Journal of Haematology, 101*, 349-351.

Yap, E. P. & McGee, J.O. (1991). Slide PCR: DNA amplification from cell samples on microscopic glass slides. *Nucleic Acids Research, 19*(15), 429

Yokota,M.; Tatsumi,N.; Tsuda,I.; Yano,I. (1995). DNA extraction and amplification from Giemsa-stained blood smears. Journal of Clinincal Laboratory Analysis, 9, 6, 387-391, UNITED STATES

Thesis Summary

Molecular characterisation of southern African *Bacillus anthracis* strains by Ayesha Hassim

Promoter: Henriette van Heerden

Co-promoter: Jenny Rossouw

Department: Veterinary Tropical Diseases, University of Pretoria

Degree: MSc (Veterinary Science)

Summary

With important diseases including zoonotic diseases, it is necessary to find accurate and reliable techniques in the diagnosis of the causal agent. *Bacillus anthracis* the causal agent of anthrax has received a great deal of attention due to its negative association with biological warfare. Microbiological tests have routinely been used to confirm diagnoses of *B. anthracis* in suspected anthrax cases and to distinguish it from *B. cereus* and *B. thuringiensis* that also belong to the *B. cereus* group, along with *B. anthracis.* Multiple locus variable number tandem repeat (VNTR) analysis (MLVA) is the current, rapid, molecular assay of choice in typing *B. anthracis* strains. It relies on commonly practiced PCR based methods to target regions which differ in tandem repeat unit. Various MLVA panels, of which the first consisted of 8 VNTR markers, followed by the MLVA15 and MLVA25 panels have been used to differentiate anthrax strains and to evaluate the diversity of *B. anthracis* from different geographical areas. In this study, we investigated the use of 31 VNTR markers (combination of MLVA15 and 25 panels) to type *B. anthracis* isolates from southern Africa using both the capillary and agarose electrophoresis methods to determine the comparative value of each. The samples included *B. anthracis sensu lato* isolates from southern Africa (n=112), a clinical *B. cereus* isolate and $34F_2$ Sterne vaccine strain. This study indicated that the resolution using agarose gel electrophoresis does not allow the accurate separation of 6 VNTR loci with tandem repeat consisting of 6 bp or less, but that the remaining 25 VNTR loci are sufficient to type *B. anthracis* strains for the purpose of epidemiological study. Agarose electrophoresis is also the most

cost effective and appropriate technique for the average African / developing country laboratory. Despite the fact that the 31 MLVA panel using capillary electrophoresis is not cost effective, it is a rapid and accurate method for *B. anthracis* typing. A comparison of the discriminative power of the four MLVA systems, using 8, 15, 25, and 31-markers clearly showed the superiority of the 31-marker MLVA. However cluster analysis of 113 *B. cereus/ B. antrhacis sensu lato* isolates from southern Africa indicated that the MLVA25 and MLVA31 panels are very similar and the latter only differentiated an additional 7 genotypes. As MLVA is known to reveal the genetic relationships within *B. anthracis*, we used the MLVA31 panel to also investigate whether it will differentiate isolates amongst the *B. cereus* group. The study revealed that MLVA alone may not be sufficient in resolving isolates from the *B. cereus* group, but is an effective tool in determining genetic and geographical distance and identifies isolates with anomalies that can be definitively identified with further study. Lastly, as blood smears are the most common method to diagnose anthrax and often the only available sample, 14 stained blood smear slides were evaluated as a source of DNA for the fingerprinting of anthrax using MLVA. This preliminary study indicated that typing of 31VNTR loci was only successful from freshly made stained blood smear slides. With stored, stained blood smears slide the quantity and quality of DNA from the slide was insufficient and had to be amplified using GenomiPhi and only amplified VNTR loci with fragment size smaller than 300 bp. This study identified factors that influenced fingerprinting / typing to be primarily the small amount of target DNA (anthrax spores) from a blood smear slide and the conditions of collection and storage. Results from this thesis, highlighted the use of MLVA for typing of *B. anthracis* and also identified areas that need further investigation.

Appendix

Appendix Table 1: Information of *Bacillus anthracis sensu lato* as well as *B. cereus* isolates including year, location and indication of source of isolation with each sample's translated co-ordinates.

*** Isolates used for 25 MLVA typing at** Institute of Genetics and Microbiology, University Paris-South Orsay, France

**** All isolates identified as *B. anthracis sensu lato* except for *B. cereus*.**

*****Isolates which have the 4 digit KNP military grid code (e.g. 1349C) were translated using Appendix Figures 9 & 10**

Isolate	Sample number **	Year or date	Location ***	Source of sample	Latitude	Longitude
Ba#000	*B. cereus*	Sample Details Unknown			unknown	
Ba#001*	N1	1975	Downstream from Gaisenga, KNP	unknown	-23.627	30.130
Ba#002*	G-12	09/1975	Pafuri, KNP	soil	-22.447	31.308
Ba#003*	B41	20/08/1990	KNP	Buffalo Bone	unknown	
Ba#004*	B31	23/08/90	1349C KNP	Buffalo Bone	-23.617	31.375
Ba#005*	B26	23/08/90	1540B KNP	Kudu Bone	-23.967	31.458
Ba#006*	B40	25/08/1990	KNP	Hippo Bone	unknown	
Ba#007*	C13	05/1975	Hapi Pan, Pafuri, KNP	Soil	-22.447	31.308
Ba#008*	C14	05/1975	Hapi Pan, Pafuri, KNP	Soil	-22.447	31.308
Ba#009*	H31	1975	Hapi Pan, Pafuri, KNP	Soil	-22.447	31.308
Ba#010*	B78	27/08/90	0948B KNP	Kudu Bone	-23.633	31.208
Ba#011*	N4	1975	Downstream from Gaisenga,KNP	Soil	-23.627	30.130
Ba#012*	N31	1975	Downstream from Gaisenga,KNP	Soil	-23.627	30.130
Ba#013*	B34	23/08/90	1248B KNP	Kudu Bone	-23.633	31.333
Ba#014*	B36	24/08/90	1545D KNP	Buffalo Bone	-23.783	31.458
Ba#015*	B37	24/08/90	1545A KNP	Buffalo Bone	-23.767	31.438
Ba#016*	B86	27/08/90	1544D KNP	Civet Bone	-23.817	31.458
Ba#017*	A11	02/1975	Pafuri, KNP	Soil	-22.447	31.308
Ba#018*	H36	11/1975	Hapi Pan, Pafuri, KNP	Soil	-22.447	31.308
Ba#019*	5	Sample Details Unknown			unknown	
Ba#020*	W23	1975	Lindada Sweni, KNP		unknown	
Ba#021*	Buffel 2	2006	Pafuri, KNP	Buffalo	-22.447	31.308
Ba#022*	A3	02/1975	Pafuri, KNP	Soil	-22.447	31.308
Ba#023*	B7	03/1975	Pafuri, KNP	Soil	-22.447	31.308
Ba#024*	B28	03/1975	Pafuri, KNP	Soil	-22.447	31.308
Ba#025*	B77	28/08/1990	KNP	Kudu Bone	unknown	
Ba#026*	B82	27/08/90	1041B KNP	Kudu Bone	-23.933	31.250
Ba#027*	W1	17/08/1990	Marumbeni crib, KNP	Water	-24.733	31.635
Ba#028*	W18	1975	Lindada Sweni, KNP	Bone	unknown	
Ba#029*	A2	Sample Details Unknown			unknown	
Ba#030*	A5	Sample Details Unknown			unknown	
Ba#031*	A8	02/1975	Pafuri, KNP	Soil	-22.447	31.308
Ba#032*	A6	02/1975	Pafuri, KNP	Soil	-22.447	31.308
Ba#033*	A16	03/1975	Pafuri, KNP	Soil	-22.447	31.308
Ba#034*	A19	02/1975	Pafuri, KNP	Soil	-22.447	31.308
Ba#035*	B42	unknown	KNP	Buffalo Bone	unknown	

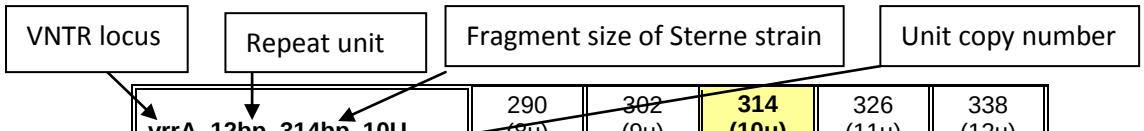
Table 1: Continued...

Isolate	Sample number **	Year or date	Location ***	Source of sample	Latitude	Longitude
Ba#036*	B43	18/08/90	1543C KNP	Buffalo Bone	-23.867	31.438
Ba#037*	B72	27/08/90	0944D KNP	Buffalo Bone	-23.817	31.208
Ba#038*	B73	28/08/90	0944D KNP	Buffalo Bone	-23.817	31.208
Ba#039*	B52	24/08/90	1245A KNP	Buffalo Bone	-23.767	31.333
Ba#040*	KB9	1999	KNP	Unknown	unknown	
Ba#041*	KB10	Sample details unknown			unknown	
Ba#042*	KB12	1999	KNP	Unknown	unknown	
Ba#043*	HP2	08/1970	Hapi Pan, Pafuri, KNP	soil	-22.447	31.308
Ba#044*	R41	unknown	Rondebosch	Unknown	-33.960	18.471
Ba#045*	R411	unknown	Rondebosch	Unknown	-33.960	18.471
Ba#046*	Z14	1999/2000	Luangwa Valley, Zambia	Unknown	-13.167	31.500
Ba#047*	Z15	1999/2000	Luangwa Valley, Zambia	Unknown	-13.167	31.500
Ba#048*	Z21	1975	Pafuri, KNP	kudu bone	-22.447	31.308
Ba#049	W21	1975	Downstream from Gaisenga, KNP	Soil	unknown	
Ba#050	N4	1975	Downstream from Gaisenga, KNP	Soil	unknown	
Ba#051	N7	1975	Downstream from Gaisenga, KNP	Soil	unknown	
Ba#052	W03	17/08/90	1443D KNP	Water	-23.867	31.417
Ba#053	W04	17/08/90	1441B KNP	Water	-23.933	31.417
Ba#054	W05	Sample details unknown			unknown	
Ba#055	W06	17/08/90	1441B KNP	Water	-23.933	31.417
Ba#056	W07	Sample details unknown			unknown	
Ba#057	W08	Sample details unknown			unknown	
Ba#058	W09	18/08/90	1544C KNP	Water	-23.817	31.438
Ba#059	W37	Sample details unknown			unknown	
Ba#060	W68	10/08/90	1052B KNP	Water	-23.467	31.250
Ba#061	302	14/09/90	0940A KNP	Impala Skin	-23.967	31.208
Ba#062	353	20/09/90	1451C KNP	Nyala Tissue	-23.533	31.396
Ba#063	359	20/09/90	1839C KNP	Kudu Skin	-24.033	31.563
Ba#064	361	20/09/90	1249B KNP	Kudu Skin	-23.600	31.333
Ba#065	363	21/09/90	1155C KNP	Impala Skin	-23.367	31.292
Ba#066	366	21/09/90	1641A KNP	Buffalo Skin	-23.933	31.479
Ba#067	370	21/09/90	1645B KNP	Buffalo Tissue	-23.767	31.500
Ba#068	374	21/09/90	1440C KNP	Kudu Tissue	-23.983	31.396
Ba#069	377	21/09/90	1540D KNP	Buffalo Skin	-23.983	31.396
Ba#070	382	21/09/90	1251A KNP	Kudu Skin	-23.517	31.333
Ba#071	391	21/09/90	1641B KNP	Kudu Skin	-23.411	31.500
Ba#072	392	21/09/90	1540C KNP	Buffalo Tissue	-23.967	31.438

Table 1: Continued...

Isolate *	Sample number **	Year or date	***Location	Source of sample	Latitude	Longitude
Ba#073	419	22/09/90	1448C KNP	Buffalo Skin	-23.650	31.396
Ba#074	25	Sample details unknown			Unknown	
Ba#075	29	Sample details unknown			Unknown	
Ba#076	65	30/08/90	1042C KNP	Impala Bone	-23.900	31.229
Ba#077	76	28/08/90	1645C KNP	Kudu Bone	-23.783	31.479
Ba#078	77	28/08/90	1645C KNP	Kudu Bone	-23.783	31.479
Ba#079	179	07/09/90	1345C KNP	Buffalo Bone	-24.067	31.479
Ba#080	79	28/08/90	1338B KNP	Impala Bone	-24.067	31.375
Ba#081	SBPRU0199	Sample details unknown			Unknown	
Ba#082	SBPRU0200	Sample details unknown			Unknown	
Ba#083	89	Sample details unknown			Unknown	
Ba#084	SBPRU0202	Sample details unknown			Unknown	
Ba#085	J 83/93	18/10/93	Masaka, KNP	Hippo	Unknown	
Ba#086	J 62/93	5/10/93	Olifants, KNP	Kudu	Unknown	
Ba#087	K 11/93	9/1/93	Kempania, KNP	Blue Wildebeest	Unknown	
Ba#088	kudu 93	Sample details unknown			Unknown	
Ba#089	V16	07-12-92	Northen Cape	Soil	Unknown	
Ba#090	V17	07-12-93	Northen Cape	Soil	Unknown	
Ba#091	V20	07-12-94	Northen Cape	Soil	Unknown	
Ba#092	3275-2D	2009	Kimberly	Soil	-28.734	24.760
Ba#093	2110	1997	Herbert	Buck	-29.049	24.010
Ba#094	6103-6B	1998	Botswana	Elephant	-28.483	24.677
Ba#095	6461-SP	1998	Sendelingsdrift	Caprise	-28.483	24.677
Ba#096	3080-5A	2009	Kimberly	Bovine	-28.734	24.760
Ba#097	6461-SP2	1998	Sendelingsdrift	Caprise	-28.483	24.677
Ba#098	5838	1998	Kimberly	Rooi hartebees	-28.734	24.760
Ba#099	3080-1B	2009	Kimberly	Bovine	-28.734	24.760
Ba#100	2991-2B	2009	Kimberly	Ovine	-28.734	24.760
Ba#101	7639	1995	unknown	Bovine	Unknown	
Ba#102	4980	1995	Vaalbos National Park	Kudu	-28.483	24.677
Ba#103	6103-6D	1998	Botswana	Elephant	-22.344	24.680
Ba#104	8334	1999	Lesotho	Giraffe	Unknown	
Ba#105	6461 SP1	1998	Sendelingsdrift	Caprise	-28.483	24.677
Ba#106	2991-1B	2009	Kimberly	Ovine	-28.734	24.760
Ba#107	3122-2B	2008	Kimberly	Gemsbok	-28.734	24.760
Ba#108	3008-1A	2009	Kimberly	Soil	-28.734	24.760
Ba#109	3080-3B	2009	Kimberly	Bovine	-28.734	24.760
Ba#110	3079-1C	2009	Kimberly	Gemsbok	-28.734	24.760
Ba#111	3132-1B	2009	Kimberly	Kudu	-28.734	24.760
Ba#112	6057	1998	Kimberly	Wildebees	-28.734	24.760
Ba#113	34F2	Sterne	vaccine strain	Reference	Not applicable	

Appendix Table 2: Comparitive table used to convert PCR product sizes to copy numbers for the MLVA 25 panel. Information includes the MLVA marker (VNTR locus) with the repeat unit length, PCR product size of Sterne reference strain and the amount of copy numbers (tandem repeat size) for the Sterne strain with listed PCR product sizes with corresponding copy number unit indicated in brackets obtained with 38 *B. anthracis* isolates.



vrrA_12bp_314bp_10U	290 (8u)	~~302~~ (9u)	**314 (10u)**	326 (11u)	338 (12u)							
vrrB1_9bp_229bp_16U	184 (11u)	193 (12u	220 (15u)	**229 (16u)**	256 (19u)	292 (23u)						
vrrB2_9bp_153bp_6U	135 (4u)	144 (5u)	**153 (6u)**	162 (7u)	171 (8u)	198 (11)						
vrrC1_9bp_580bp_53U	364 (29u)	504 (44u)	517 (46u)	535 (48u)	**580 (53u)**	607 (56u)	616 (57u)	688 (65u)				
vrrC2_18bp_532bp_17U	**532 (17u)**	568 (19u)	604 (21u)									
CG3_5bp_158bp_2U	153 (1u)	**158 (2u)**										
pXO1_aat_3bp_126bp_7U	120 (5u)	123 (6u)	**126 (7u)**	129 (8u)	132 (9u)	135 (10u)	138 (11u)	144 (13u)				
pXO2_at_2bp_141bp_10U	133 (6u)	135 (7u)	137 (8u)	139 (9u)	**141 (10u)**	143 (11u)	145 (12u)	147 (13u)	149 (14u)	155 (17u)		
BAMS1_21bp_485bp_16U	296 (7u)	380 (11u)	401 (12u)	422 (13u)	443 (14u)	**485 (16u)**	611 (22u)					
BAMS3_15bp_549bp_26U	474 (21u)	519 (24u)	**549 (26u)**	564 (27u)	579 (28u)	594 (29u)	609 (30u)	654 (33u)				
BAMS5_39bp_307bp_5U	229 (3u)	**307 (5u)**	329 (5,5)	346 (6u)	385 (7u)	424 (8u)						
BAMS13_9bp_814bp_70U	337 (17u)	346 (18u)	364 (20u)	373 (21u)	391 (23u)	400 (24u)	427 (27u)	436 (28u)	454 (30u)	463 (31u)	481 (33u)	571 (43u)
	607 (47u)	679 (55u)	778 (66u)	**814 (70u)**	823 (71u)	832 (72u)	868 (76u)					
BAMS15_9bp_418bp_24U	409 (23u)	**418 (24u)**	445 (27u)	535 (37u)	571 (41u)	580 (42u)	589 (43u)	598 (44u)	607 (45u)	616 (46u)	634 (48u)	643 (49u)
BAMS21_45bp_676bp_10U	541 (7u)	631 (9u)	**676 (10u)**	721 (11u)								

Table 2: Continued …

BAMS22_36bp_735bp_16U	519 (10u)	555 (11u)	627 (13u)	663 (14u)	699 (15u)	**735 (16u)**	933 (21,5u)	1041(24,5u)				
BAMS23_42bp_651bp_11U	399 (5u)	567 (9u)	609 (10u)	630 (10,5u)	**651 (11u)**	672 (11,5u)	693 (12u)					
BAMS24_42bp_595bp_11U	469 (8u)	511 (9u)	553 (10u)	574 (10,5u)	**595 (11u)**	637 (12u)						
BAMS25_15bp_391bp_13U	376 (12u)	**391 (13u)**										
BAMS28_24bp_493bp_14U	373 (9u)	397 (10u)	469 (13u)	**493 (14u)**	505 (14,5u)							
BAMS30_9bp_727bp_57U	268 (6u)	367 (17u)	493 (31u)	682 (52u)	691 (53u)	700 (54u)	**727 (57u)**	754 (60u)	826 (67u)	835 (69u)	853 (71u)	862 (72u)
	889 (75u)	916 (78u)	925 (79u)									
BAMS31_9bp_772bp_64U	331 (15u)	547 (39u)	637 (49u)	646 (50u)	691 (55u)	700 (56u)	709 (57u)	727 (59u)	745 (61u)	763 (63u)	**772 (64u)**	781 (65u)
	826 (70u)	952 (84u)	1087 (99u)									
BAMS34_39bp_503bp_11U	230 (4u)	386 (8u)	425 (9u)	**503 (11u)**	581 (13u)							
BAMS44_39bp_417bp_8U	339 (6u)	378 (7u)	**417 (8u)**	573 (12u)								
BAMS51_45bp_493bp_9U	358 (6u)	403 (7u)	448 (8u)	**493 (9u)**	516 (10u)	538 (11u)						
BAMS53_12bp_346bp_8U	322 (7u)	**346 (8u)**										

Appendix Table 3: Comparitive table used to convert PCR product sizes to copy numbers for 6 VNTR loci. Information includes the MLVA marker (VNTR locus) with the repeat unit length, PCR product size of Sterne reference strain and the amount of copy numbers (tandem repeat size) for the Sterne strain with listed PCR product size with copy number unit indicated in brackets obtained with *B. anthracis* isolates.

VNTR12_2bp_115bp_6U	105 (1U)	107 (2U)	109 (3U)	111 (4U)	113 (5U)	115 (6U)	117 (7U)	119 (8U)	121 (9U)	123 (10U)	125(11U)			
VNTR16_8bp_275bp_8U	219 (1U)	227 (2U)	235 (3U)	243 (4U)	251 (5U)	259 (6U)	267 (7U)	275 (8U)	283 (9U)	291 (10U)	299 (11U)	307 (12U)		
VNTR17_8bp_382bp_4U	358 (1U)	366 (2U)	374 (3U)	382 (4U)	390 (5U)	398 (6U)	406 (7U)	414 (8U)	422 (9U)	430 (10U)	438 (11U)			
VNTR19_3bp_96bp_4U	87 (1U)	90 (2U)	93 (3U)	96 (4U)	99 (5U)	102 (6U)	105 (7U)	108 (8U)	111 (9U)	114 (10U)	117 (11U)	120 (12U)	123 (13U)	126 (14U)
VNTR23_12bp_196bp_4U	160 (1U)	172 (2)	184 (3U)	196 (4U)	208 (5U)	220 (6U)	232 (7U)	244 (8U)	256 (9U)					
VNTR35_6bp_116bp_4U	98 (1U)	104 (2U)	110 (3U)	116 (4U)	122 (5U)	128 (6U)	134 (7U)	140 (8U)	146 (9U)	152 (10U)	158 (11U)			

Appendix Table 4: Comparative table used to convert amplicon sizes to copy units for capillary electrophoresis. This table was obtained from University of Hohenheim, Institute of Environmental and Animal Hygiene, Stuttgart, Germany (indicated as Beyerlab in table) and demonstrates the difference in amplicon reads with different electrophoresis machines. The reference strain used was the Ames strain of *B. anthracis.*

	VNTR12 locus PCR product size in bp			Copy number of VNTR 12 locus	VNTR12 locus PCR product size in bp				VNTR 17 locus PCR product size in bp			copy number of VNTR 17 locus
	VNTR12 ABI 377	VNTR12 ABI 3100	VNTR12 Beyerlab		VNTR16 ABI 377	VNTR16 ABI 3100	VNTR16 Beyerlab	Copy number of VNTR16 locus	VNTR17 ABI 377	VNTR17 ABI 3100	VNTR17 Beyerlab	
								-			-	-
					143	137	141				-	-
							-	-			-	-
					249		-	-			-	-
					257		-	-			-	-
	111		112	4	265	259	-	-	371	366	-	-
	113	107	114	5	273	266-267	271	7	379	374	379	3
	115	109	116	6	281	274-275	279	8	387	382	387	4
	117	111.5			289	283	287	9	395	390	395	5
	119	113.5			296	290	295	10	403	399	-	-
					305	298	303	11	411		-	-
					313	307	-	-	419		-	-
					321		-	-	427		-	-
					329		-	-	435		-	-
					337	331.5	-	-	445	441.5	-	-
Repeat unit size	2bp	2bp	2bp		8bp	8bp	8bp	-	8bp	8bp	8bp	-

Appendix Table 5: Copy numbers for *B. anthracis sensu lato* isolates at the 31 VNTR loci for capillary electrophoresis (blue) and agarose electrophoresis (green)

Sample ID	Bams1	Bams1	Bams3	Bams3	Bams5	Bams5	Bams13	Bams13	Bams15	Bams15	Bams21	Bams21	Bams22	Bams22	Bams23	Bams23
Ba#001	11	10	18	18	6	6	70	70	31	24	9	12	15	15	10	10
Ba#002	12	11	27	27	6	6	33	33	48	41	7	10	13	13	10	10
Ba#003	11	10	18	18	6	6	29	29	48	41	9	12	15	15	10	10
Ba#004	11	10	18	18	6	6	29	29	48	41	9	12	15	15	10	10
Ba#005	11	10	18	18	6	6	33	33	48	41	9	12	15	15	10	10
Ba#006	11	10	18	18	6	6	70	70	31	24	9	12	15	15	10	10
Ba#007	12	11	27	27	6	6	32	32	48	41	7	10	13	13	10	10
Ba#008	12	11	27	27	6	6	29	29	51	44	7	10	14	14	10	10
Ba#009	11	10	18	18	6	6	29	29	51	44	9	12	15	15	10	10
Ba#010	11	10	18	18	6	6	33	33	50	43	9	12	15	15	10	10
Ba#011	12	11	18	18	6	6	33	33	50	43	9	12	15	15	11	11
Ba#012	12	11	18	18	6	6	33	33	50	43	9	12	15	15	11	11
Ba#013	12	11	18	18	6	6	33	33	50	43	9	12	15	15	11	11
Ba#014	12	11	18	18	6	6	33	33	50	43	9	12	15	15	11	11
Ba#015	12	11	18	18	6	6	33	33	50	43	9	12	15	15	11	11
Ba#016	12	11	18	18	6	6	31	31	50	43	9	12	15	15	10	10
Ba#017	13	12	27	27	6	6	27	27	53	46	6	9	13	13	10	10
Ba#018	13	12	27	27	6	6	31	27	53	46	6	9	13	13	10	10
Ba#019	12	11	18	18	6	6	27	31	50	43	7	10	15	15	10	10
Ba#020	13	12	27	27	6	6	33	31	52	43	7	10	15	15	10	10
Ba#021	11	10	18	18	6	6	28	33	50	43	9	12	15	15	10	10
Ba#022	11	10	26	26	6	6	28	28	52	45	7	10	12	12	10	10
Ba#023	11	10	27	27	6	6	33	28	52	45	7	10	12	12	10	10
Ba#024	10	9	17	17	6	6	37	33	50	43	9	12	14	14	10	10
Ba#025	10	9	17	17	6	6	33	37	50	43	9	12	15	15	10	10
Ba#026	10	9	17	17	6	6	32	33	50	43	8	11	18	18	10	10
Ba#027	10	9	17	17	6	6	33	32	50	43	9	12	16	16	10	10
Ba#028	10	9	17	17	6	6	30	33	51	44	9	12	17	17	10	10
Ba#029	10	9	30	30	6	6	27	30	55	48	9	11	5	5	10	10
Ba#030	10	9	26	26	6	6	27	27	54	47	8	11	6	6	10	10
Ba#031	10	9	26	26	6	6	30	27	54	47	8	11	5	5	10	10
Ba#032	11	10	24	24	6	6	28	30	55	48	8	11	6	6	10	10

Table 5 continued...

Sample ID	Bams1	Bams1	Bams3	Bams3	Bams5	Bams5	Bams13	Bams13	Bams15	Bams15	Bams21	Bams21	Bams22	Bams22	Bams23	Bams23
Ba#033	11	10	28	28	6	6	24	28	55	48	8	11	5	5	10	10
Ba#034	11	10	27	27	6	6	28	24	54	47	8	11	4	4	10	10
Ba#035	10	9	17	17	6	6	30	28	48	41	10	13	7	7	11	11
Ba#036	10	9	16	16	6	6	32	30	55	45	10	13	7	7	11	11
Ba#037	10	9	17	17	6	6	32	32	55	45	10	13	9	9	11	11
Ba#038	10	9	19	19	6	6	33	32	54	47	10	13	12	12	11	11
Ba#039	12	11	18	18	6	6	28	33	50	43	10	12	15	15	10	10
Ba#040	13	12	27	27	6	6	33	28	52	45	9	10	13	13	10	10
Ba#041	12	11	27	27	6	6	33	33	50	43	7	10	13	13	11	11
Ba#042	12	11	27	27	6	6	33	33	50	43	7	12	15	15	11	11
Ba#043	12	11	18	18	6	6	33	33	50	43	10	12	15	15	10	10
Ba#044	14	13	27	27	7	7	70	71	53	46	10	10	16	16	11	11
Ba#045	14	13	27	27	7	7	70	72	53	46	7	10	16	16	11	11
Ba#046	12	11	18	18	6	6	33	33	51	44	9	12	15	15	10	10
Ba#047	12	11	18	18	6	6	33	33	51	44	9	12	15	15	10	10
Ba#048	12	11	18	18	6	6	28	28	52	45	8	10	13	13	10	10
Ba#049	10	9	29	29	6	6	31	31	52	45	10	13	18	18	11	11
Ba#050	10	9	19	19	6	6	36	36	50	43	10	13	15	15	11	11
Ba#051	10	9	18	18	6	6	30	30	50	43	9	12	15	15	11	11
Ba#052	10	9	16	16	6	6	28	28	51	44	10	13	16	16	11	11
Ba#053	10	9	18	18	6	6	29	29	52	45	10	13	15	15	11	11
Ba#054	10	9	17	17	6	6	28	28	53	46	10	13	14	14	11	11
Ba#055	10	9	17	17	6	6	30	30	53	46	10	13	16	16	11	11
Ba#056	10	9	17	17	6	6	29	29	54	47	10	13	14	14	11	11
Ba#057	10	9	28	28	6	6	33	33	56	49	11	14	14	14	11	11
Ba#058	10	9	28	28	6	6	33	33	58	51	11	14	15	15	11	11
Ba#059	16	15	11	11	11	11	70	13	54	47	8	11	5	5	15	15
Ba#060	10	9	17	17	6	6	31	31	52	45	10	13	14	14	10	10
Ba#061	10	9	17	17	6	6	30	30	52	45	10	13	14	14	11	11
Ba#062	10	9	18	18	6	6	33	33	51	44	10	13	15	15	10	10
Ba#063	10	9	17	17	7	7	32	32	51	44	10	13	16	16	10	10
Ba#064	10	9	21	21	6	6	31	31	51	44	10	13	16	16	10	10
Ba#065	10	9	18	18	6	6	36	36	50	43	10	13	15	15	10	10

Table 5 continued …

Sample ID	Bams1	Bams1	Bams3	Bams3	Bams5	Bams5	Bams13	Bams13	Bams15	Bams15	Bams21	Bams21	Bams22	Bams22	Bams23	Bams23
Ba#066	10	9	16	16	6	6	36	36	50	43	10	13	15	15	10	10
Ba#067	11	10	17	17	6	6	33	33	50	43	10	13	15	15	10	10
Ba#068	10	9	17	17	6	6	30	30	50	43	10	13	16	16	10	10
Ba#069	10	9	18	18	6	6	34	34	50	43	10	13	14	14	10	10
Ba#070	10	9	18	18	6	6	34	34	50	43	10	13	14	14	10	10
Ba#071	10	9	18	18	6	6	31	31	50	43	10	13	14	14	11	11
Ba#072	11	10	18	18	6	6	33	33	50	43	10	13	14	14	11	11
Ba#073	10	9	18	18	6	6	31	31	51	44	10	13	15	15	10	10
Ba#074	11	10	30	30	7	7	27	27	53	46	8	11	13	13	10	10
Ba#075	12	11	29	29	6	6	26	26	53	46	8	11	13	13	10	10
Ba#076	12	11	28	28	6	6	27	27	54	47	10	13	14	14	10	10
Ba#077	11	10	18	18	6	6	33	33	52	45	10	13	16	16	11	11
Ba#078	12	11	26	26	6	6	27	27	54	47	8	11	14	14	10	10
Ba#079	11	10	17	17	6	6	33	33	52	45	10	13	15	15	10	10
Ba#080	11	10	26	26	6	6	28	28	54	47	8	11	13	13	10	10
Ba#081	11	10	27	27	6	6	28	28	54	47	8	11	15	15	10	10
Ba#082	11	10	27	27	6	6	28	28	52	45	8	11	15	15	10	10
Ba#083	12	11	27	27	6	6	32	32	54	47	8	11	13	13	10	10
Ba#084	10	9	17	17	6	6	35	35	52	45	10	13	15	15	11	11
Ba#085	10	9	18	18	6	6	28	28	53	46	10	13	15	15	11	11
Ba#086	10	9	17	17	6	6	31	31	52	45	11	14	15	15	11	11
Ba#087	10	9	16	16	6	6	28	28	51	44	10	13	14	14	11	11
Ba#088	17	16	16	16	6	6	29	29	51	44	10	13	14	14	11	11
Ba#089	17	16	16	16	5	5	65	65	54	47	8	11	15	15	11	11
Ba#090	10	9	28	28	6	6	65	66	54	47	8	11	16	16	11	11
Ba#091	10	9	16	16	6	6	28	28	51	44	10	13	16	16	10	10
Ba#092	17	16	16	16	6	6	65	64	54	47	7	10	17	17	11	11
Ba#093	10	9	29	29	6	6	65	64	54	47	7	10	16	16	11	11
Ba#094	12	11	17	17	6	6	29	29	52	45	7	10	16	16	11	11
Ba#095	17	16	25	25	6	6	29	29	51	44	10	13	14	14	11	11
Ba#096	17	16	27	27	7	7	70	73	54	47	9	12	15	15	12	12
Ba#097	12	11	22	22	6	6	28	28	52	45	9	12	16	16	11	11
Ba#098	16	15	29	29	6	6	65	67	55	48	9	12	15	15	10	10

Table 5 continued …

Sample ID	Bams1	Bams1	Bams3	Bams3	Bams5	Bams5	Bams13	Bams13	Bams15	Bams15	Bams21	Bams21	Bams22	Bams22	Bams23	Bams23
Ba#099	16	15	28	28	7	7	70	78	55	48	8	11	16	16	11	11
Ba#100	16	15	29	29	6	6	70	69	55	48	8	11	16	16	11	11
Ba#101	12	11	25	25	6	6	24	24	55	48	8	11	13	13	11	11
Ba#102	17	16	25	25	6	6	65	66	55	48	8	11	16	16	10	10
Ba#103	11	10	16	16	6	6	29	29	54	47	10	13	14	14	9	9
Ba#104	17	16	15	15	3	3	51	51	54	47	8	11	16	16	11	11
Ba#105	12	11	22	22	6	6	29	29	54	47	8	11	16	16	11	11
Ba#106	16	15	30	30	6	6	65	64	54	47	8	11	17	17	11	11
Ba#107	16	15	29	29	6	6	~~55~~	69	55	48	7	10	17	17	11	11
Ba#108	16	15	28	28	6	6	58	55	55	48	8	11	16	16	11	11
Ba#109	16	15	27	27	6	6	65	57	54	47	8	11	16	16	10	10
Ba#110	16	15	29	29	6	6	65	60	56	49	8	11	16	16	11	11
Ba#111	16	15	31	31	6	6	65	68	55	48	8	11	16	16	11	11
Ba#112	16	15	28	28	5	5	65	61	55	48	8	11	16	16	11	11
Ba#113	16	16	26	26	5	7	70	75	24	45	10	13	16	16	11	11

Sample ID	Bams24	Bams24	Bams25	Bams25	VNTR23	VNTR23	VNTR12	VNTR12	VNTR19	VNTR19	VNTR35	VNTR35	VNTR16	VNTR16	VNTR17	VNTR17
Ba#001	11	11	13	13	3	3	4	3	5	5	3	3	8	8	5	5
Ba#002	11	11	13	13	3	3	2	2	3	3	2	2	8	8	5	5
Ba#003	11	11	13	13	3	3	4	3	3	3	2	2	11	11	3	3
Ba#004	11	11	13	13	3	3	5	4	3	3	2	2	8	8	5	5
Ba#005	11	11	13	13	3	3	5	4	3	3	2	2	8	8	4	4
Ba#006	11	11	13	13	3	3	5	4	5	5	2	2	8	8	4	4
Ba#007	11	11	13	13	4	4	5	5	5	5	3	3	11	11	5	4
Ba#008	11	11	13	13	4	4	5	3	5	5	3	3	8	8	5	5
Ba#009	11	11	13	13	4	4	5	4	5	5	3	3	8	8	5	5
Ba#010	11	11	13	13	4	4	5	4	3	3	3	3	8	8	5	5
Ba#011	11	11	13	13	3	3	5	6	3	3	3	3	8	8	4	4
Ba#012	11	11	13	13	4	4	5	6	3	3	3	3	8	8	4	4
Ba#013	11	11	13	13	3	3	5	7	4	4	3	3	8	8	4	4
Ba#014	11	11	13	13	4	4	5	8	5	5	3	3	11	11	4	4

Table 5 continued …

Sample ID	Bams24	Bams24	Bams25	Bams25	VNTR23	VNTR23	VNTR12	VNTR12	VNTR19	VNTR19	VNTR35	VNTR35	VNTR16	VNTR16	VNTR17	VNTR17
Ba#015	11	11	13	13	4	4	5	7	5	5	2	2	8	8	5	4
Ba#016	11	11	13	13	4	4	5	8	3	3	3	3	8	8	3	3
Ba#017	11	11	13	13	6	6	5	6	4	4	2	2	8	8	5	5
Ba#018	11	11	13	13	6	6	5	6	5	4	2	2	8	8	5	5
Ba#019	11	11	13	13	5	5	5	6	5	5	2	2	8	8	4	4
Ba#020	11	11	13	13	3	3	5	6	6	5	3	2	9	8	4	4
Ba#021	11	11	13	13	4	4	5	9	7	7	6	6	9	9	3	3
Ba#022	12	12	13	13	3	3	5	8	6	6	2	2	12	12	3	3
Ba#023	12	12	13	13	5	5	5	8	7	7	4	4	12	12	4	4
Ba#024	12	12	13	13	4	4	5	12	12	9	3	3	8	8	4	4
Ba#025	11	11	13	13	5	5	5	10	6	6	4	4	9	9	4	4
Ba#026	11	11	13	13	4	4	5	9	7	7	5	5	9	9	4	4
Ba#027	11	11	13	13	5	5	5	10	7	7	3	3	9	9	4	4
Ba#028	11	11	13	13	4	4	5	9	4	4	6	6	9	9	5	5
Ba#029	11	11	13	13	3	3	5	7	4	4	4	4	10	10	5	5
Ba#030	11	11	13	13	3	3	5	6	4	4	2	2	12	12	4	4
Ba#031	12	12	13	13	3	3	5	5	3	3	2	2	12	12	4	4
Ba#032	11	11	13	13	3	3	5	5	3	3	4	4	9	9	4	4
Ba#033	12	12	13	13	3	3	5	5	3	3	4	4	12	12	4	4
Ba#034	12	12	13	13	3	3	5	5	3	3	4	4	9	9	4	4
Ba#035	11	11	13	13	3	3	5	5	3	3	4	4	11	11	5	5
Ba#036	11	11	13	13	3	3	5	6	3	3	2	2	11	11	5	5
Ba#037	11	11	13	13	3	3	5	5	4	4	2	2	12	12	5	5
Ba#038	12	12	13	13	3	3	5	5	5	5	2	2	12	12	5	5
Ba#039	11	11	13	13	5	5	5	6	5	5	3	3	11	11	5	5
Ba#040	11	11	13	13	5	5	5	5	5	5	3	3	9	9	5	5
Ba#041	11	11	13	13	6	6	5	5	3	3	3	3	9	9	5	5
Ba#042	11	11	13	13	4	4	5	5	5	5	2	2	8	8	5	5
Ba#043	11	11	13	13	4	4	5	6	4	4	2	2	8	8	5	5
Ba#044	11	11	13	13	5	5	5	7	5	5	2	2	8	8	5	5
Ba#045	11	11	13	13	5	5	5	7	5	5	3	3	8	8	5	5
Ba#046	11	11	13	13	5	5	5	6	4	4	2	2	11	11	5	5
Ba#047	11	11	13	13	6	6	5	10	6	6	2	2	10	10	4	4

Table 5 continued …

Sample ID	Bams24	Bams24	Bams25	Bams25	VNTR23	VNTR23	VNTR12	VNTR12	VNTR19	VNTR19	VNTR35	VNTR35	VNTR16	VNTR16	VNTR17	VNTR17
Ba#048	11	11	13	13	6	6	5	10	6	6	2	2	9	9	4	4
Ba#049	12	12	13	13	4	4	6	9	3	5	3	3	8	8	5	5
Ba#050	11	11	13	13	4	4	6	9	2	5	3	3	8	8	5	5
Ba#051	12	12	13	13	4	4	6	10	3	4	2	2	8	8	5	5
Ba#052	12	12	13	13	4	4	6	9	2	4	2	2	9	9	4	4
Ba#053	12	12	13	13	4	4	5	10	4	5	3	3	9	9	4	4
Ba#054	11	11	13	13	4	4	6	10	2	5	3	3	10	10	4	4
Ba#055	11	11	13	13	4	4	6	9	2	5	3	3	22	22	4	4
Ba#056	11	11	13	13	4	4	6	9	3	5	3	3	22	22	4	4
Ba#057	11	11	13	13	4	4	6	9	2	5	3	3	10	10	4	4
Ba#058	11	11	13	13	4	4	6	10	3	5	3	3	9	9	5	5
Ba#059	11	11	11	11	5	5	6	5	1	9	3	4	22	3	4	12
Ba#060	11	11	13	13	4	4	6	9	3	5	3	3	9	9	5	5
Ba#061	12	12	13	13	4	4	6	9	2	6	3	3	8	8	6	6
Ba#062	11	11	13	13	4	4	6	9	2	6	3	3	9	9	5	5
Ba#063	11	11	13	13	4	4	6	9	2	6	3	3	9	9	5	5
Ba#064	11	11	13	13	4	4	6	9	2	6	3	3	9	9	6	6
Ba#065	11	11	13	13	4	4	6	7	2	6	3	3	8	8	6	6
Ba#066	11	11	13	13	4	4	5	9	2	5	3	3	9	9	6	6
Ba#067	11	11	13	13	4	4	6	9	2	6	3	3	9	9	5	5
Ba#068	11	11	13	13	4	4	6	9	2	5	3	3	8	8	5	5
Ba#069	12	12	13	13	4	4	6	9	2	5	3	3	9	9	6	6
Ba#070	11	11	13	13	4	4	6	9	2	5	3	3	10	10	5	5
Ba#071	12	12	13	13	4	4	7	9	2	5	3	3	8	10	5	5
Ba#072	11	11	13	13	4	4	6	9	2	5	3	3	9	9	5	5
Ba#073	12	12	13	13	4	4	6	10	2	5	3	3	10	10	4	4
Ba#074	12	12	13	13	2	2	5	8	3	5	3	3	12	12	4	4
Ba#075	11	11	13	13	3	3	5	9	2	5	3	3	12	12	4	4
Ba#076	11	11	13	13	3	3	5	9	2	5	2	2	12	12	4	4
Ba#077	11	11	13	13	4	4	6	11	2	5	2	2	9	9	5	5
Ba#078	11	11	13	13	3	3	5	9	2	5	2	2	12	12	4	4
Ba#079	11	11	13	13	4	4	5	10	2	5	2	2	10	10	5	5
Ba#080	11	11	13	13	3	3	5	9	2	5	2	2	12	12	4	4

Table 5 continued …

Sample ID	Bams24	Bams24	Bams25	Bams25	VNTR23	VNTR23	VNTR12	VNTR12	VNTR19	VNTR19	VNTR35	VNTR35	VNTR16	VNTR16	VNTR17	VNTR17
Ba#081	11	11	13	13	3	3	5	10	2	5	2	2	12	12	4	4
Ba#082	11	11	13	13	3	3	4	10	2	5	3	3	12	12	4	4
Ba#083	11	11	13	13	3	3	5	10	2	5	3	3	14	14	5	4
Ba#084	11	11	13	13	4	4	5	10	1	5	2	2	8	8	5	5
Ba#085	11	11	13	13	4	4	6	9	2	5	5	5	8	8	5	5
Ba#086	12	12	13	13	4	4	6	10	3	5	2	2	8	8	5	5
Ba#087	11	11	12	12	4	4	5	10	2	5	2	2	9	9	4	4
Ba#088	11	11	12	12	4	4	5	10	2	6	3	3	9	9	4	4
Ba#089	11	11	12	12	6	6	6	10	1	5	3	3	9	9	3	3
Ba#090	11	11	12	12	6	6	6	10	1	5	3	3	10	10	4	4
Ba#091	11	11	12	12	3	3	6	9	2	6	4	4	10	10	4	4
Ba#092	11	11	13	13	5	5	6	9	1	5	5	5	8	8	3	3
Ba#093	11	11	13	13	5	5	5	10	1	5	4	4	8	8	3	3
Ba#094	11	11	13	13	4	4	6	10	2	5	4	4	10	10	5	5
Ba#095	11	11	13	13	4	4	6	10	3	5	3	3	9	9	5	5
Ba#096	11	11	13	13	5	5	6	9	1	4	3	3	9	9	5	5
Ba#097	11	11	13	13	3	3	5	8	1	5	6	6	7	7	6	6
Ba#098	10	10	13	13	4	4	6	8	2	3	4	4	6	6	4	4
Ba#099	11	11	13	13	5	5	6	8	1	3	3	3	12	12	4	4
Ba#100	11	11	13	13	3	3	6	8	1	3	6	6	12	12	4	4
Ba#101	11	11	13	13	6	6	4	7	2	4	4	4	11	11	4	4
Ba#102	10	10	13	13	3	5	5	7	1	4	3	3	7	7	4	4
Ba#103	11	11	13	13	5	4	7	7	3	4	3	3	8	8	5	5
Ba#104	11	11	14	14	4	5	6	8	2	3	3	3	8	8	5	5
Ba#105	11	11	13	13	5	4	6	7	1	4	4	4	8	8	5	5
Ba#106	11	11	14	14	4	5	6	7	1	3	4	4	8	8	4	4
Ba#107	11	11	13	13	5	5	5	9	1	3	4	6	7	7	4	4
Ba#108	11	11	13	13	5	5	6	9	1	3	6	6	8	8	4	4
Ba#109	11	11	13	13	5	5	6	10	1	3	3	3	8	8	5	5
Ba#110	11	11	13	13	5	5	5	11	1	4	3	3	7	7	4	4
Ba#111	11	11	13	13	5	4	6	13	1	5	3	2	8	8	4	4
Ba#112	11	11	13	13	4	4	5	13	1	5	2	2	8	8	4	4
Ba#113	11	11	13	13	4	4	6	9	4	4	4	4	8	8	4	4

Table 5 continued …

Sample ID	VrrA	VrrA	VrrB1	VrrB1	VrrB2	VrrB2	VrrC1	VrrC1	VrrC2	VrrC2	CG3	CG3	pXO1	pXO1	pXO2	pXO2	Bams28	Bams28
Ba#001	13	12	16	16	7	7	56	56	17	17	1	2	4	6	5	7	14	14
Ba#002	10	9	16	16	7	7	53	53	17	17	1	2	2	6	10	10	14	14
Ba#003	13	12	16	16	7	7	57	57	17	17	3	2	3	6	4	7	14	14
Ba#004	13	12	16	16	7	7	56	56	17	17	2	2	3	6	5	7	14	14
Ba#005	13	12	19	19	7	7	56	56	17	17	2	2	4	6	4	7	14	14
Ba#006	13	12	19	19	8	8	56	56	17	17	2	2	4	6	10	10	14	14
Ba#007	10	9	19	19	8	8	54	54	17	17	1	2	4	6	12	11	14	14
Ba#008	13	12	16	16	7	7	53	53	17	17	1	2	4	6	12	11	14	14
Ba#009	13	12	16	16	7	7	58	58	17	17	1	2	4	6	10	8	13	13
Ba#010	13	12	16	16	7	7	57	57	17	17	1	2	4	6	10	8	14	14
Ba#011	13	12	16	16	6	6	57	57	17	17	2	2	2	6	10	8	14	14
Ba#012	13	12	16	16	6	6	57	57	17	17	1	2	4	6	10	8	14	14
Ba#013	13	12	16	16	6	6	57	57	17	17	2	2	4	6	10	8	14	14
Ba#014	13	12	16	16	6	6	57	57	17	17	2	2	2	6	10	8	14	14
Ba#015	13	12	16	16	6	6	57	57	17	17	1	2	2	6	10	7	14	14
Ba#016	13	12	16	16	6	6	57	57	17	17	1	2	4	6	10	8	14	14
Ba#017	10	9	25	25	7	7	53	53	17	17	1	2	4	6	12	10	13	13
Ba#018	10	9	25	25	7	7	53	53	17	17	2	2	3	6	12	10	13	13
Ba#019	13	12	29	29	6	6	57	57	17	17	1	2	2	6	4	6	14	14
Ba#020	13	12	29	29	6	6	57	57	17	17	2	2	4	6	12	6	14	14
Ba#021	13	12	16	16	7	7	52	52	18	18	1	2	3	6	4	5	14	14
Ba#022	10	9	23	23	8	8	52	52	18	18	1	2	4	7	4	5	13	13
Ba#023	10	9	19	19	8	8	52	52	18	18	2	3	4	7	5	6	13	13
Ba#024	13	12	19	19	8	8	52	52	17	17	2	3	2	7	5	6	14	14
Ba#025	14	13	16	16	6	6	55	55	7	7	2	3	2	7	5	6	14	14
Ba#026	14	13	16	16	11	11	55	55	17	17	1	3	2	7	10	7	14	14
Ba#027	13	12	16	16	6	6	55	55	16	16	2	3	4	7	6	6	14	14
Ba#028	13	12	16	16	6	6	58	58	17	17	1	3	4	7	1	6	14	14
Ba#029	9	8	18	18	8	8	55	55	18	18	2	2	4	7	1	1	13	13
Ba#030	9	8	18	18	8	8	55	55	18	18	1	3	4	6	2	1	13	13
Ba#031	9	8	19	19	8	8	55	55	18	18	1	3	4	5	2	1	13	13
Ba#032	9	8	19	19	7	7	55	55	19	19	1	3	4	2	1	1	13	13

Table 5 continued …

Sample ID	VrrA	VrrA	VrrB1	VrrB1	VrrB2	VrrB2	VrrC1	VrrC1	VrrC2	VrrC2	CG3	CG3	pXO1	pXO1	pXO2	pXO2	Bams28	Bams28
Ba#033	9	8	18	18	7	7	55	55	18	18	1	2	4	5	2	1	13	13
Ba#034	9	8	19	19	7	7	55	55	18	18	1	2	2	5	3	2	14	14
Ba#035	9	8	15	15	6	6	60	60	18	18	1	2	2	4	3	2	15	15
Ba#036	9	8	16	16	6	6	59	59	18	18	1	2	2	7	2	2	15	15
Ba#037	9	8	16	16	6	6	59	59	18	18	1	2	2	7	3	2	15	15
Ba#038	9	8	16	16	7	7	60	60	18	18	2	3	4	7	3	2	15	15
Ba#039	13	12	16	16	7	7	57	57	17	17	2	2	4	7	6	7.5	14	14
Ba#040	9	9	19	19	7	7	54	54	17	17	2	2	4	7	10	11	13	13
Ba#041	13	12	16	16	7	7	57	57	17	17	2	2	4	7	9	10	13	13
Ba#042	13	12	16	16	7	7	57	57	17	17	1	2	8	7	6	8	14	14
Ba#043	13	12	16	16	7	7	57	57	17	17	1	2	3	6	6	8	14	14
Ba#044	11	10	16	16	7	7	57	57	17	17	1	2	2	13	12	9	14	14
Ba#045	11	10	16	16	7	7	57	57	17	17	1	2	2	11	12	10	14	14
Ba#046	11	10	16	16	6	6	57	57	17	17	2	2	4	8	5	7	14	14
Ba#047	10	9	16	16	6	6	57	57	17	17	2	2	3	8	10	10	14	14
Ba#048	10	9	16	16	6	6	53	53	17	17	1	2	4	7	10	10	13	13
Ba#049	12	11	16	16	6	6	60	60	16	16	1	2	4	6	4	5	14	14
Ba#050	12	11	16	16	6	6	60	60	17	17	1	2	4	6	4	6	15	15
Ba#051	12	11	16	16	6	6	60	60	16	16	1	2	4	6	3	5	14	14
Ba#052	12	11	16	16	7	7	61	61	17	17	3	2	2	6	4	5	13	13
Ba#053	12	11	16	16	7	7	61	61	17	17	2	1	4	6	4	5	13	13
Ba#054	13	12	16	16	6	6	59	59	17	17	2	2	4	6	4	5	13	13
Ba#055	13	12	16	16	6	6	59	59	16	16	2	2	4	7	4	5	13	13
Ba#056	13	12	16	16	6	6	60	60	17	17	2	2	4	7	3	5	13	13
Ba#057	13	12	16	16	6	6	60	60	17	17	1	2	4	7	4	6	13	13
Ba#058	13	12	16	16	6	6	62	62	17	17	2	2	4	7	4	6	14	14
Ba#059	12	12	11	11	16	16	66	66	17	17	1	2	8	7	4	6	14	14
Ba#060	12	11	16	16	6	6	60	60	19	19	1	2	4	7	5	6	14	14
Ba#061	12	11	17	17	6	6	61	61	17	17	1	2	4	6	4	5	14	14
Ba#062	13	12	17	17	6	6	63	63	18	18	3	2	2	6	4	5	13	13
Ba#063	13	12	17	17	6	6	62	62	18	18	1	2	4	6	5	4	14	14
Ba#064	13	12	17	17	7	7	58	58	18	18	1	2	4	6	4	4	13	13

Table 5 continued …

Sample ID	VrrA	VrrA	VrrB1	VrrB1	VrrB2	VrrB2	VrrC1	VrrC1	VrrC2	VrrC2	CG3	CG3	pXO1	pXO1	pXO2	pXO2	Bams28	Bams28
Ba#065	13	12	17	17	7	7	58	58	18	18	2	2	4	6	4	4	13	13
Ba#066	13	12	17	17	6	6	61	61	18	18	2	2	4	6	4	4	13	13
Ba#067	13	12	17	17	6	6	60	60	18	18	1	2	4	6	4	5	13	13
Ba#068	13	12	17	17	6	6	61	61	18	18	2	2	4	7	4	5	13	13
Ba#069	13	12	17	17	7	7	60	60	18	18	2	3	4	6	4	5	13	13
Ba#070	13	12	17	17	6	6	61	61	18	18	2	3	4	6	7	5	13	13
Ba#071	13	12	17	17	6	6	60	60	17	17	2	2	4	7	4	5	13	13
Ba#072	13	12	17	17	6	6	59	59	18	18	2	3	4	6	4	5	13	13
Ba#073	10	9	16	16	6	6	56	56	17	17	1	3	4	7	4	5	14	14
Ba#074	10	9	20	20	8	8	53	53	18	18	2	4	4	7	4	8	14	14
Ba#075	10	9	20	20	8	8	53	53	18	18	2	2	4	6	6	7	16	16
Ba#076	10	9	20	20	8	8	57	57	17	17	2	2	4	7	7	9	13	13
Ba#077	13	12	17	17	7	7	57	57	16	16	2	2	4	7	3	6	14	14
Ba#078	10	9	20	20	7	7	53	53	17	17	2	2	4	7	5	9	12	12
Ba#079	13	12	17	17	7	7	57	57	17	17	1	1	4	6	4	6	15	15
Ba#080	10	9	20	20	8	8	53	53	17	17	1	2	4	7	8	9	13	13
Ba#081	10	9	20	20	8	8	52	52	17	17	2	2	4	7	7	9	13	13
Ba#082	10	9	20	20	8	8	52	52	17	17	2	2	4	7	7	9	15	15
Ba#083	10	9	19	19	8	8	52	52	17	17	2	2	4	6	7	9	13	13
Ba#084	13	12	16	16	7	7	56	56	17	17	1	2	4	6	3	6	16	16
Ba#085	12	11	16	16	7	7	56	56	17	17	2	3	4	6	4	6	14	14
Ba#086	14	13	16	16	6	6	56	56	17	17	2	2	4	6	3	5	13	13
Ba#087	14	13	16	16	6	6	57	57	17	17	2	3	5	6	4	4	13	13
Ba#088	14	13	16	16	8	8	56	56	17	17	1	2	4	5	6	6	14	14
Ba#089	12	11	16	16	6	6	56	56	17	17	2	3	8	5	4	12	14	14
Ba#090	12	11	17	17	6	6	57	57	17	17	2	3	8	9	4	5	14	14
Ba#091	14	13	16	16	6	6	56	56	17	17	1	3	4	8	4	6	14	14
Ba#092	14	13	17	17	6	6	55	55	18	18	2	1	12	-2	-2	6	15	15
Ba#093	12	11	17	17	6	6	57	57	17	17	1	2	8	4	10	5	14	14
Ba#094	13	12	17	17	7	7	56	56	18	18	2	2	4	4	9	-2	14	14
Ba#095	12	11	17	17	7	7	59	59	18	18	2	2	5	12	3	8	12	12
Ba#096	12	11	16	16	7	7	59	59	18	18	1	2	7	8	3	6	15	15

Table 5 continued …

Sample ID	VrrA	VrrA	VrrB1	VrrB1	VrrB2	VrrB2	VrrC1	VrrC1	VrrC2	VrrC2	CG3	CG3	pXO1	pXO1	pXO2	pXO2	Bams28	Bams28
Ba#097	11	10	16	16	6	6	58	58	21	21	1	3	5	5	9	11	13	13
Ba#098	10	9	16	16	7	7	60	60	17	17	2	2	8	8	3	11	15	15
Ba#099	11	10	17	17	7	7	54	54	17	17	1	2	7	8	3	12	15	15
Ba#100	11	10	17	17	6	6	58	58	18	18	1	2	7	8	4	12	17	17
Ba#101	10	9	19	19	8	8	53	53	17	17	2	2	5	6	6	13	14	14
Ba#102	11	10	16	16	7	7	57	57	17	17	2	3	8	9	10	13	14	14
Ba#103	12	11	16	16	6	6	56	56	17	17	1	3	4	5	4	7	15	15
Ba#104	9	8	16	16	6	6	57	57	18	18	4	4	-2	-2	-2	-2	14	14
Ba#105	11	10	17	17	7	7	57	57	21	21	2	3	5	5	9	10	14	14
Ba#106	11	10	17	17	6	6	58	58	17	17	2	3	8	9	11	12	16	16
Ba#107	11	10	16	16	7	7	58	58	17	17	2	3	8	9	11	12	16	16
Ba#108	11	10	16	16	7	7	58	58	17	17	2	2	-2	-2	10	13	14	14
Ba#109	11	10	17	17	6	6	57	57	17	17	1	3	8	8	10	6	14	14
Ba#110	11	10	16	16	7	7	56	56	18	18	2	2	8	9	10	6	13	13
Ba#111	11	10	17	17	7	7	56	56	18	18	2	2	8	9	4	12	14	14
Ba#112	11	10	17	17	7	7	57	57	17	17	2	2	8	9	11	6	16	16
Ba#113	10	10	16	16	6	6	53	53	17	17	2	3	7	7	-2	-2	14	14

Sample ID	Bams30	Bams30	Bams31	Bams31	Bams34	Bams34	Bams44	Bams44	Bams51	Bams51	Bams53	Bams53
Ba#001	72	68	49	64	9	9	2	2	9	9	8	8
Ba#002	72	68	50	65	9	9	2	2	6	6	6	6
Ba#003	73	69	50	65	9	9	2	2	9	9	8	8
Ba#004	72	68	49	64	9	9	2	2	9	9	8	8
Ba#005	72	68	49	64	9	9	2	2	9	9	8	8
Ba#006	72	68	49	64	9	9	2	2	9	9	8	8
Ba#007	10	6	49	64	9	9	8	8	6	6	6	6
Ba#008	10	6	49	64	9	9	8	8	6	6	5	5
Ba#009	72	68	49	64	9	9	2	2	9	9	8	8
Ba#010	73	69	49	64	9	9	2	2	6	6	8	8
Ba#011	72	68	49	64	9	9	2	2	9	9	8	8
Ba#012	72	68	49	64	9	9	2	2	9	9	8	8

Table 5 continued …

Sample ID	Bams30	Bams30	Bams31	Bams31	Bams34	Bams34	Bams44	Bams44	Bams51	Bams51	Bams53	Bams53
Ba#013	73	69	49	64	9	9	2	2	9	9	8	8
Ba#014	72	68	49	64	9	9	2	2	9	9	8	8
Ba#015	72	68	49	64	9	9	2	2	9	9	8	8
Ba#016	72	68	49	64	9	9	2	2	9	9	8	8
Ba#017	10	6	50	65	9	9	8	8	9	9	6	6
Ba#018	10	6	50	65	9	9	8	8	6	6	6	6
Ba#019	72	68	49	64	9	9	2	2	9	9	8	8
Ba#020	10	68	50	64	9	9	2	2	9	9	6	8
Ba#021	72	68	48	63	9	9	2	2	9	9	8	1
Ba#022	10	6	50	65	7	7	9	9	6	6	1	1
Ba#023	10	6	49	64	6	6	8	8	6	6	1	1
Ba#024	73	69	49	64	9	9	3	3	9	9	1	1
Ba#025	70	67	48	63	9	9	2	2	8	8	1	1
Ba#026	72	68	47	62	9	9	2	2	8	8	6	1
Ba#027	72	68	49	64	9	9	3	3	8	8	1	1
Ba#028	71	67	49	64	9	9	2	2	8	8	1	1
Ba#029	10	6	49	64	8	8	8	8	6	6	7	7
Ba#030	10	6	52	67	9	9	8	8	6	6	7	7
Ba#031	10	6	51	66	8	8	8	8	6	6	7	7
Ba#032	9	5	51	66	8	8	8	8	6	6	7	7
Ba#033	10	6	51	66	9	9	8	8	6	6	7	7
Ba#034	9	5	52	67	9	9	8	8	6	6	7	7
Ba#035	76	72	50	65	9	9	2	2	9	9	7	7
Ba#036	75	71	51	66	9	9	2	2	9	9	10	9
Ba#037	73	69	51	66	9	9	2	2	9	9	9	9
Ba#038	75	71	50	66	9	9	2	2	9	9	10	9
Ba#039	72	68	50	65	9	9	2	2	9	9	9	8
Ba#040	10	6	42	66	9	9	8	8	6	6	8	6
Ba#041	10	6	42	66	9	9	2	2	9	9	6	8
Ba#042	73	69	50	65	9	9	2	2	9	9	8	8
Ba#043	73	69	50	65	9	9	2	2	9	9	8	8
Ba#044	61	57	42	57	9	9	8	8	9	9	8	8
Ba#045	61	57	42	57	9	9	8	8	9	9	8	8
Ba#046	73	69	50	65	9	9	2	2	9	9	8	8
Ba#047	73	69	50	65	9	9	2	2	9	9	8	8
Ba#048	10	6	50	65	9	9	2	2	9	9	8	6
Ba#049	75	71	48	63	9	9	2	2	9	9	6	6
Ba#050	77	73	48	63	8	8	2	2	9	9	10	10
Ba#051	75	72	45	60	9	9	2	2	9	9	10	10
Ba#052	77	73	44	59	8	8	2	2	9	9	10	10
Ba#053	77	73	45	60	9	9	2	2	9	9	10	10
Ba#054	77	73	44	59	9	9	2	2	9	9	10	10
Ba#055	76	72	47	62	9	9	2	2	9	9	10	10
Ba#056	76	72	44	59	9	9	2	2	9	9	10	10
Ba#057	76	72	46	61	8	8	2	2	9	9	10	10
Ba#058	77	73	47	62	8	8	2	2	9	9	10	10
Ba#059	53	49	45	60	8	8	8	8	9	9	10	10
Ba#060	75	72	48	63	8	8	8	8	9	9	11	11
Ba#061	80	76	49	64	8	8	2	2	9	9	10	10
Ba#062	75	71	50	65	8	8	2	2	9	9	10	10
Ba#063	76	72	50	65	8	8	2	2	9	9	11	11
Ba#064	76	72	53	68	8	8	2	2	9	9	11	11

Table 5 continued …

Sample ID	Bams30	Bams30	Bams31	Bams31	Bams34	Bams34	Bams44	Bams44	Bams51	Bams51	Bams53	Bams53
Ba#065	75	71	50	65	8	8	2	2	9	9	10	10
Ba#066	75	71	50	65	9	9	2	2	9	9	10	10
Ba#067	74	70	50	65	9	9	2	2	9	9	10	10
Ba#068	74	70	51	66	8	8	2	2	9	9	10	10
Ba#069	75	71	50	65	8	8	2	2	9	9	10	10
Ba#070	72	68	51	66	8	8	2	2	9	9	10	10
Ba#071	76	72	51	66	8	8	2	2	9	9	10	10
Ba#072	76	72	50	65	8	8	2	2	9	9	10	10
Ba#073	9	5	49	64	8	8	2	2	9	9	10	10
Ba#074	10	6	52	67	8	8	8	8	6	6	8	8
Ba#075	9	5	54	69	9	9	8	8	6	6	7	7
Ba#076	10	6	51	66	8	8	8	8	6	6	6	6
Ba#077	77	73	50	65	9	9	2	2	9	9	10	10
Ba#078	11	7	53	68	8	8	8	8	6	6	7	7
Ba#079	78	74	52	67	8	8	2	2	9	9	9	9
Ba#080	10	6	52	67	8	8	8	8	6	6	7	7
Ba#081	11	7	54	69	8	8	8	8	6	6	8	8
Ba#082	10	6	54	69	8	8	8	8	9	9	8	8
Ba#083	11	7	52	67	8	8	8	8	6	6	8	8
Ba#084	77	73	50	65	8	8	2	2	9	9	10	10
Ba#085	77	73	49	64	9	9	2	2	9	9	10	10
Ba#086	76	72	51	66	8	8	2	2	9	9	9	9
Ba#087	74	70	51	66	8	8	2	2	9	9	9	9
Ba#088	75	71	52	67	9	9	8	8	9	9	9	9
Ba#089	51	47	46	61	9	9	8	8	9	9	9	9
Ba#090	53	49	53	68	9	9	2	2	9	9	9	9
Ba#091	76	72	46	61	8	8	8	8	9	9	9	9
Ba#092	56	52	46	61	8	8	8	8	9	9	9	9
Ba#093	54	50	46	61	9	9	4	4	9	9	9	9
Ba#094	81	77	52	67	9	9	4	4	9	9	9	9
Ba#095	76	15	52	67	9	9	2	2	9	9	9	9
Ba#096	58	53	51	66	11	11	8	8	9	9	9	9
Ba#097	82	78	51	66	8	8	4	4	9	9	10	10
Ba#098	53	49	45	60	9	9	8	8	9	9	9	9
Ba#099	55	51	52	67	9	9	8	8	9	9	9	9
Ba#100	53	49	46	61	9	9	8	8	9	9	9	9
Ba#101	10	6	51	66	9	9	8	8	6	6	7	7
Ba#102	53	49	45	60	10	10	2	2	9	9	8	8
Ba#103	76	72	51	66	9	9	9	9	9	9	9	9
Ba#104	-2	-2	50	65	-2	-2	2	2	9	9	9	9
Ba#105	81	77	50	65	9	9	8	8	9	9	10	10
Ba#106	53	49	44	59	10	10	8	8	9	9	9	9
Ba#107	53	49	45	60	9	9	8	8	9	9	10	10
Ba#108	53	49	44	59	9	9	8	8	9	9	10	10
Ba#109	50	46	43	58	9	9	8	8	9	9	9	9
Ba#110	52	48	44	59	9	9	8	8	9	9	9	9
Ba#111	54	50	-2	59	9	9	8	8	9	9	9	9
Ba#112	54	50	44	59	9	9	8	8	9	9	10	10
Ba#113	57	45	64	59	11	11	8	8	9	9	8	10

Appendix Figure 1:

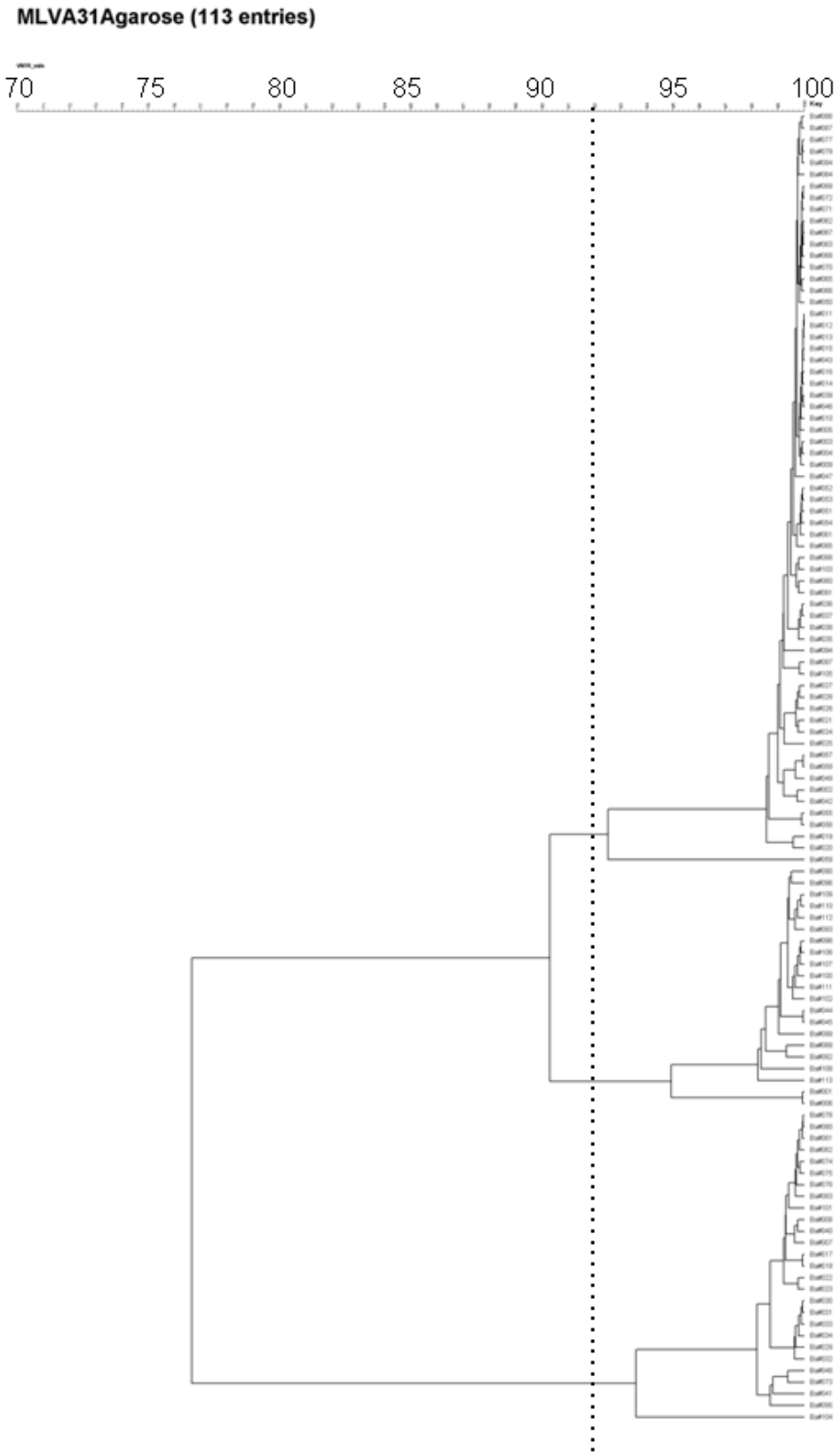


Figure 1: UPGMA dendrogram based on 31 VNTR loci generated using the agarose electrophoresis method demonstrating 92% similarity between isolates.

*See page 130 for more detail

Appendix Figure 2:

Figure 2: UPGMA dendogram based on 31 VNTR loci generated using the capillary electrophoresis method demonstrating 78% similarity between isolates.

*See page 131 for more detail

Appendix Figure 3:

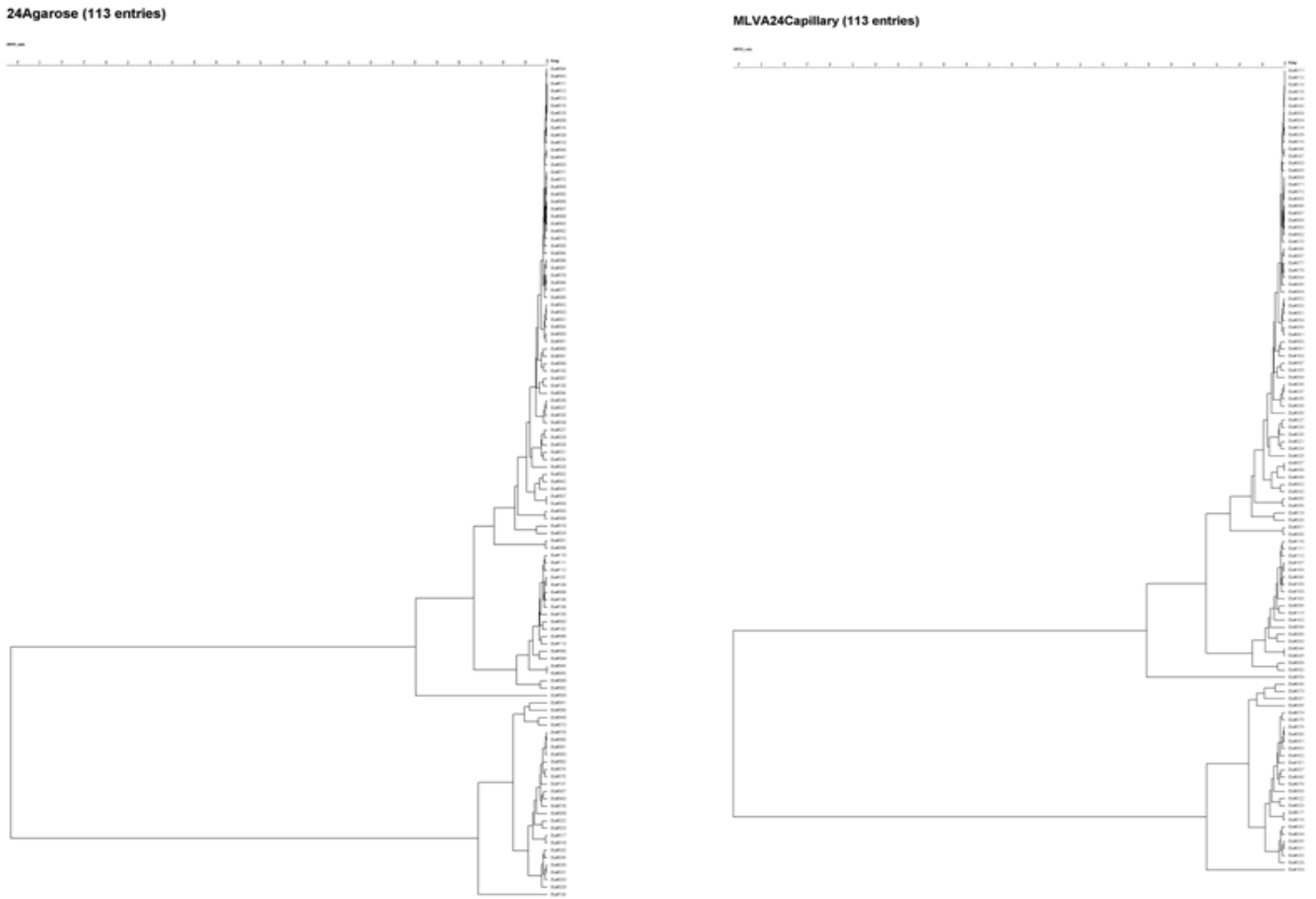


Figure 3: Dendogram based on 24 MLVA (which excludes VNTR's: pXO1; pXO2; CG3; VNTR 12; 19; 35 and Bams 13) and generated on each of the electrophoresis platforms using UPGMA clustering analysis.

Appendix Figure 4:

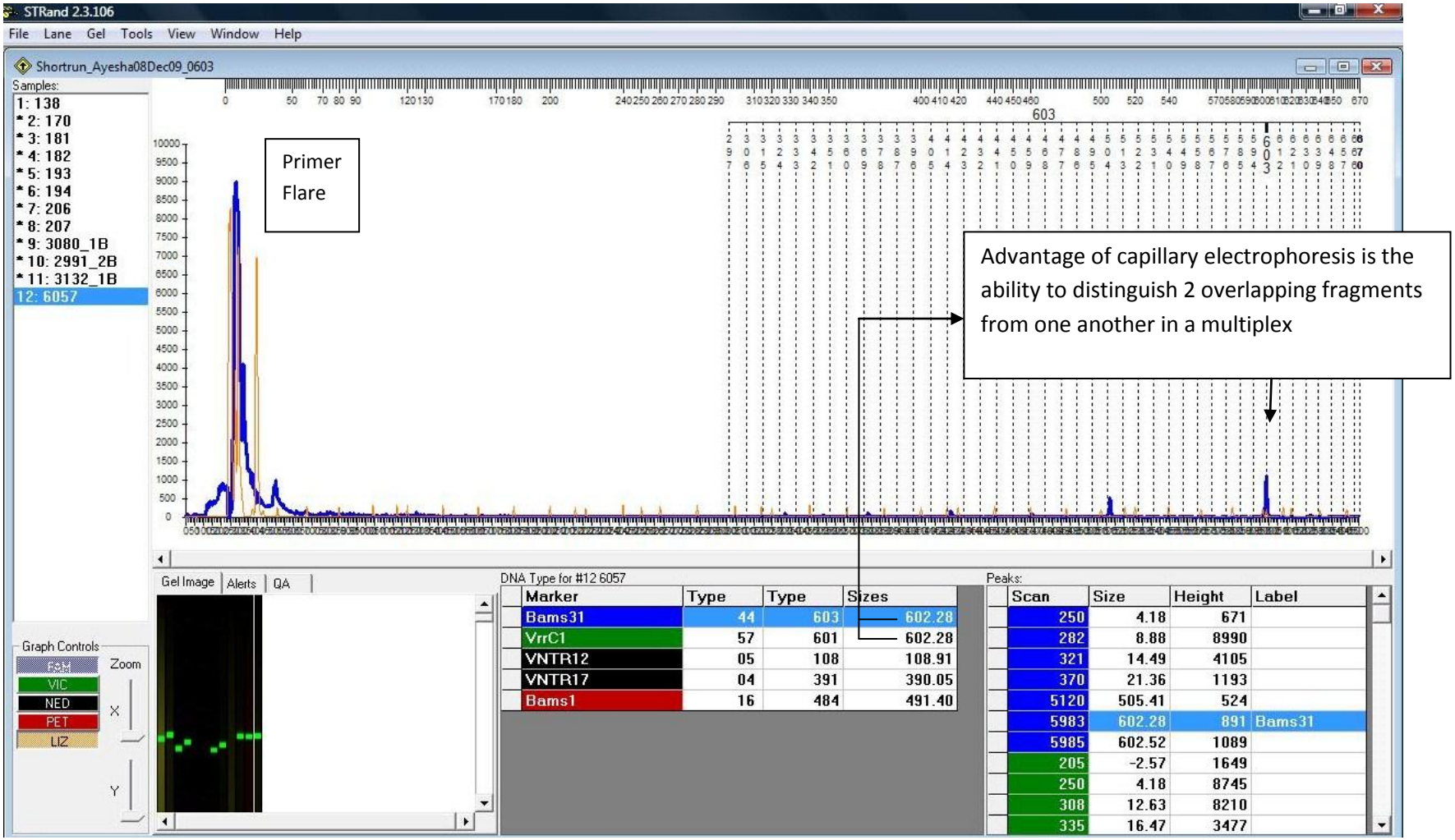


Figure 4: Example of an electropherogram demonstrating the advantages of the capillary electrophoresis technique using multiple loci in a single PCR reaction and the ease to distinguish amplicons with similar size of different markers in the multiplex enabled by different fluorophore labels.

Appendix Figure 5:

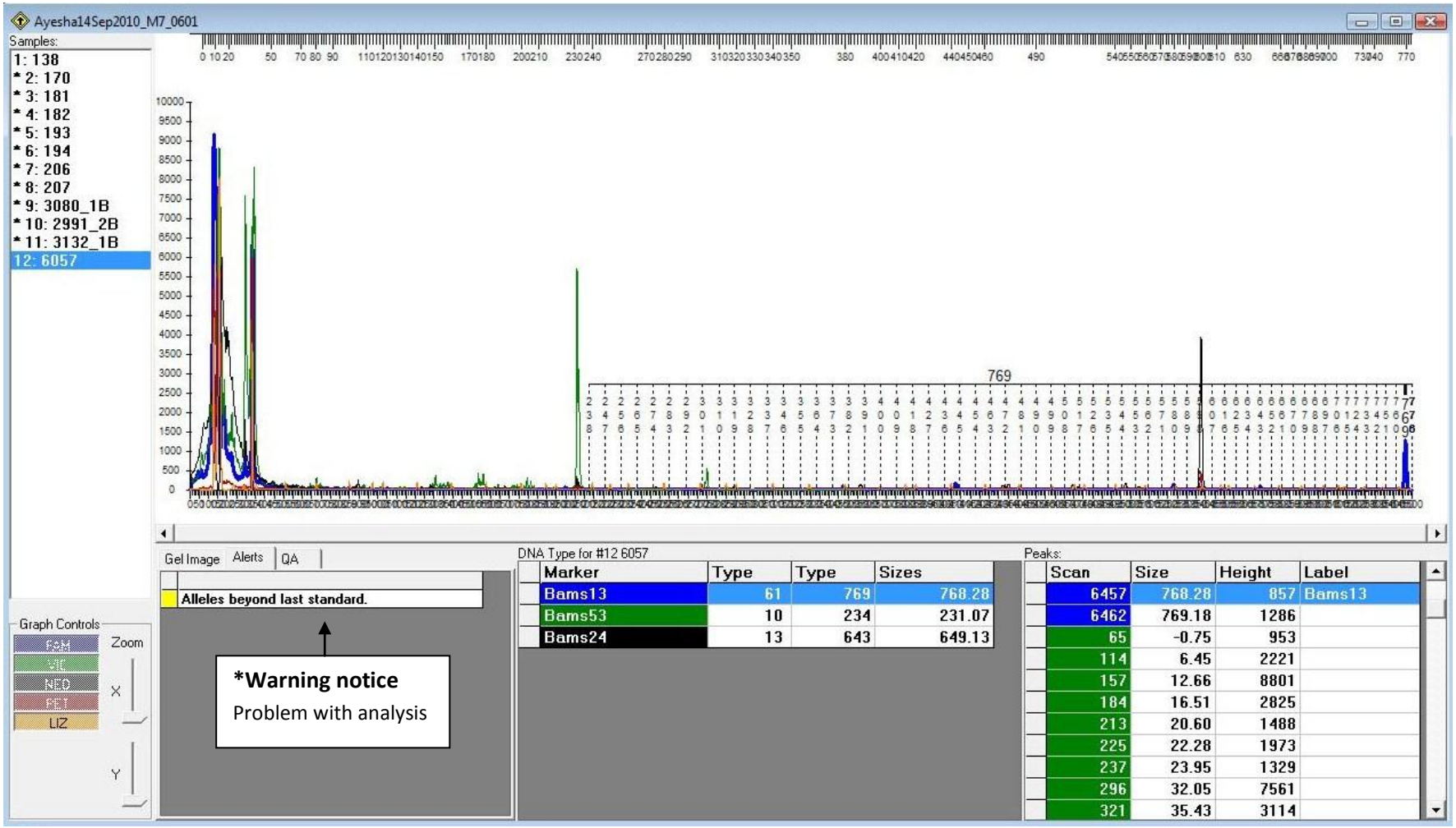


Figure 5: An example of limitations in the analysis of multiplex PCR for the sample Ba#112 (indicated as 6057) in the correct assigning of copy number for large amplicon sizes (when analysed on a 31 cm capillary) represented as an electropherogram using STRand® software. A copy number of 61 units was assigned using capillary electrophoresis, where as 65 units was assigned using agarose electrophoresis (Appendix Table 5). The “true peak” for Bams 13 cannot be read as it falls passed the last size standard indicated by the *warning notice.

Appendix Figure 6:

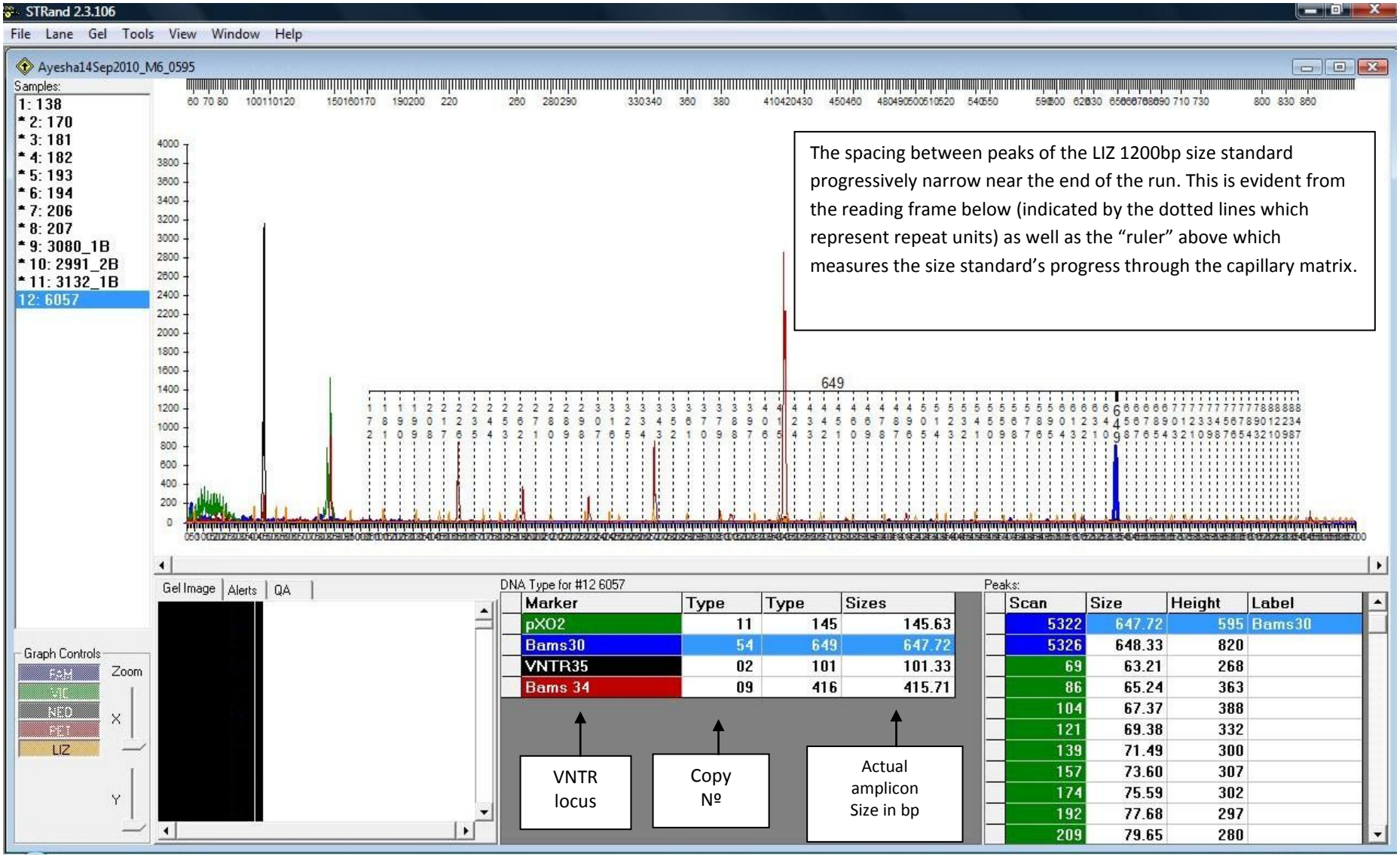
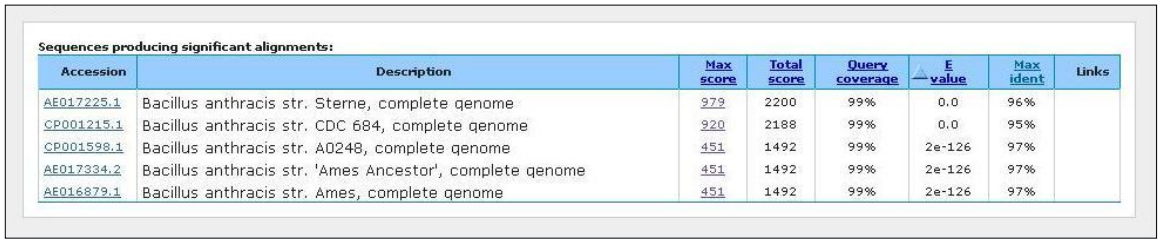


Figure 6: Electropherogram profile for multiplex PCR reaction containing pOX2, Bams30, VNTR35 and Bams34 loci. This profile also demonstrates the decreasing accuracy of copy assignments with increasing amplicon size.

Appendix Figure 7:



Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
AE017225.1	Bacillus anthracis str. Sterne, complete genome	979	2200	99%	0.0	96%	
CP001215.1	Bacillus anthracis str. CDC 684, complete genome	920	2188	99%	0.0	95%	
CP001598.1	Bacillus anthracis str. A0248, complete genome	451	1492	99%	2e-126	97%	
AE017334.2	Bacillus anthracis str. 'Ames Ancestor', complete genome	451	1492	99%	2e-126	97%	
AE016879.1	Bacillus anthracis str. Ames, complete genome	451	1492	99%	2e-126	97%	

>gb|AE017225.1| Bacillus anthracis str. Sterne, complete genome
Length=5228663

Sort alignments for this subject sequence by:
E value Score Percent identity
Query start position Subject start position

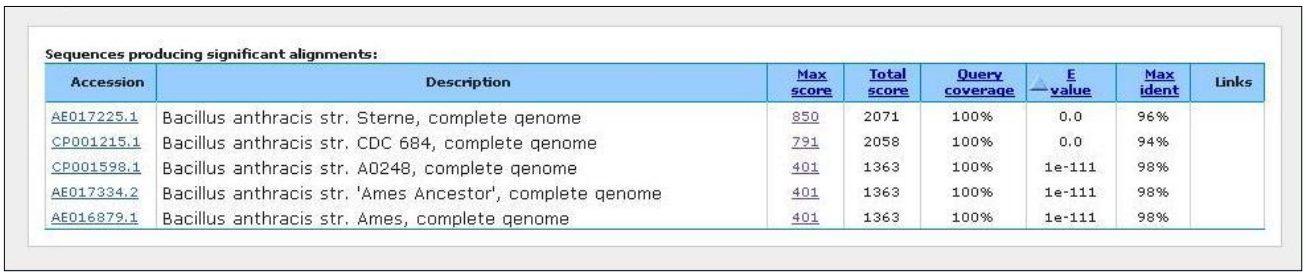
Features in this part of subject sequence:
conserved hypothetical protein, collagen triple helix rep...

Score = 979 bits (530), Expect = 0.0
Identities = 581/602 (97%), Gaps = 20/602 (3%)
Strand=Plus/Minus

```
Query  2        GTGTACANNG-TGATTCATGCTGTTTTAGTTGCGATGGAACAGTTCCTAAGCTGGGTCCA  60
                |||||||  | ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333484  GTGTACA-TGTTGATTCATGCTGTTTTAGTTGCGATGGAACAGTTCCTAAGCTGGGTCCA  4333426

Query  61       ACAGGTCCGACGGGAGCAACGGGAGCAACAGGAGCAACAGGAGTAACGGGAGTAACGGGA  120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333425  ACAGGTCCGACGGGAGCAACGGGAGCAACAGGAGCAACAGGAGTAACGGGAGTAACGGGA  4333366

Query  121      GTAACAGGAGCGACAGGAATAACGGGAGCGACAGGAATAACAGGAGCAACAGGAATAACG  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333365  GTAACAGGAGCGACAGGAATAACGGGAGCGACAGGAATAACAGGAGCAACAGGAATAACG  4333306

Query  181      GGAGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGAGCAACGGGTCCA  240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333305  GGAGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGAGCAACGGGTCCA  4333246

Query  241      ACGGGAGCAACAGGTCCGACGGGAGCGACAGGAATAACGGGAGCAACAGGAGCAACAGGA  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333245  ACGGGAGCAACAGGTCCGACGGGAGCGACAGGAATAACGGGAGCAACAGGAGCAACAGGA  4333186

Query  301      GCAACAGGAGCAACGGGAGTAACGGGAGTAACAGGAGCGACAGGAATAACAGGAGCAACA  360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333185  GCAACAGGAGCAACGGGAGTAACGGGAGTAACAGGAGCGACAGGAATAACAGGAGCAACA  4333126

Query  361      GGAATAACGGGAGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGTCCA  420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333125  GGAATAACGGGAGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGTCCA  4333066

Query  421      ACGGGAGCGACAGG--T--C-----C-A-A---C----GGGAGTCACAGGTACAAGTATT  462
                |||||||||||||  |  |     | | |   |    ||||||||||||||||||||||
Sbjct  4333065  ACGGGAGCGACAGGAATAACGGGAGCGACAGGTCCAACGGGAGTCACAGGTACAAGTATT  4333006

Query  463      ACGGCGACGTATGCATTTGCAAATAATACGTCAGGGACAGCGATATCGGTACTTCTTGGT  522
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333005  ACGGCGACGTATGCATTTGCAAATAATACGTCAGGGACAGCGATATCGGTACTTCTTGGT  4332946

Query  523      GGAACAAATGTCCCGCTTCCAAATAATCAAAATATCGGTCCGGGGATTACTGTATCTGGG  582
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4332945  GGAACAAATGTCCCGCTTCCAAATAATCAAAATATCGGTCCGGGGATTACTGTATCTGGG  4332886

Query  583      GG  584
                ||
Sbjct  4332885  GG  4332884
```

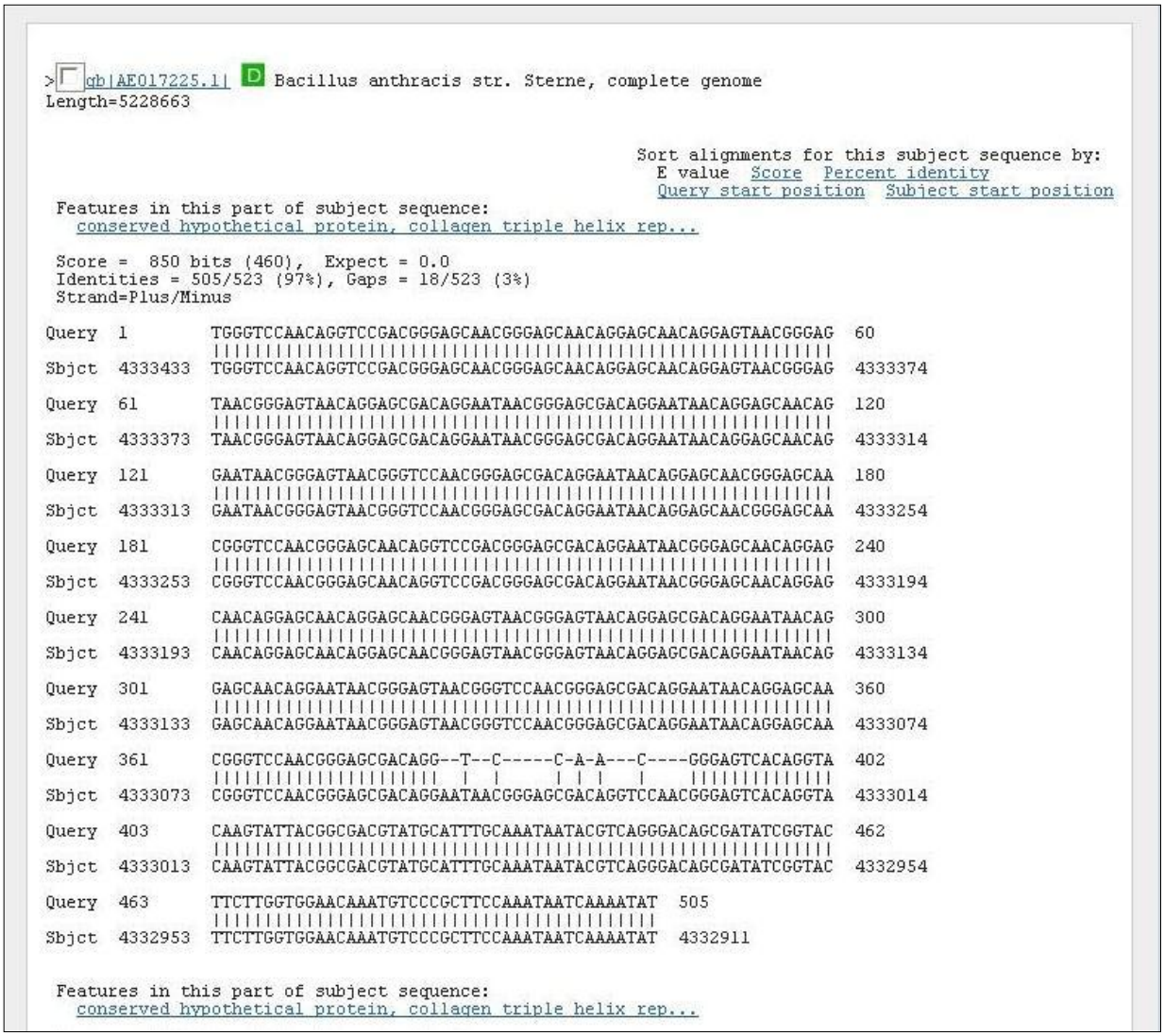
Figure 7: Blast search of of isolate Ba#94 Bams15 sequence (indicated as Query; PCR product of 606 bp) with Sterne sequence AE017225.1 (indicated as Sbjct). The repeat unit of Bams 15 in isolates Ba#94 translates to 45 copy units. The sequence was edited and blasted in NCBI nucleotide database.

Appendix Figure 8:



Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident	Links
AE017225.1	Bacillus anthracis str. Sterne, complete genome	850	2071	100%	0.0	96%	
CP001215.1	Bacillus anthracis str. CDC 684, complete genome	791	2058	100%	0.0	94%	
CP001598.1	Bacillus anthracis str. A0248, complete genome	401	1363	100%	1e-111	98%	
AE017334.2	Bacillus anthracis str. 'Ames Ancestor', complete genome	401	1363	100%	1e-111	98%	
AE016879.1	Bacillus anthracis str. Ames, complete genome	401	1363	100%	1e-111	98%	



>gb|AE017225.1| Bacillus anthracis str. Sterne, complete genome
Length=5228663

Sort alignments for this subject sequence by:
E value Score Percent identity
Query start position Subject start position

Features in this part of subject sequence:
conserved hypothetical protein, collagen triple helix rep...

Score = 850 bits (460), Expect = 0.0
Identities = 505/523 (97%), Gaps = 18/523 (3%)
Strand=Plus/Minus

```
Query  1        TGGGTCCAACAGGTCCGACGGGAGCAACGGGAGCAACAGGAGCAACAGGAGTAACGGGAG  60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333433  TGGGTCCAACAGGTCCGACGGGAGCAACGGGAGCAACAGGAGCAACAGGAGTAACGGGAG  4333374

Query  61       TAACGGGAGTAACAGGAGCGACAGGAATAACGGGAGCGACAGGAATAACAGGAGCAACAG  120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333373  TAACGGGAGTAACAGGAGCGACAGGAATAACGGGAGCGACAGGAATAACAGGAGCAACAG  4333314

Query  121      GAATAACGGGAGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGAGCAA  180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333313  GAATAACGGGAGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAACGGGAGCAA  4333254

Query  181      CGGGTCCAACGGGAGCAACAGGTCCGACGGGAGCGACAGGAATAACGGGAGCAACAGGAG  240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333253  CGGGTCCAACGGGAGCAACAGGTCCGACGGGAGCGACAGGAATAACGGGAGCAACAGGAG  4333194

Query  241      CAACAGGAGCAACAGGAGCAACGGGAGTAACGGGAGTAACAGGAGCGACAGGAATAACAG  300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333193  CAACAGGAGCAACAGGAGCAACGGGAGTAACGGGAGTAACAGGAGCGACAGGAATAACAG  4333134

Query  301      GAGCAACAGGAATAACGGGAGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAA  360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333133  GAGCAACAGGAATAACGGGAGTAACGGGTCCAACGGGAGCGACAGGAATAACAGGAGCAA  4333074

Query  361      CGGGTCCAACGGGAGCGACAGG--T--C-----C-A-A---C----GGGAGTCACAGGTA  402
                ||||||||||||||||||||||  |  |     | | |   |    ||||||||||||||
Sbjct  4333073  CGGGTCCAACGGGAGCGACAGGAATAACGGGAGCGACAGGTCCAACGGGAGTCACAGGTA  4333014

Query  403      CAAGTATTACGGCGACGTATGCATTTGCAAATAATACGTCAGGGACAGCGATATCGGTAC  462
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  4333013  CAAGTATTACGGCGACGTATGCATTTGCAAATAATACGTCAGGGACAGCGATATCGGTAC  4332954

Query  463      TTCTTGGTGGAACAAATGTCCCGCTTCCAAATAATCAAAATAT  505
                |||||||||||||||||||||||||||||||||||||||||||
Sbjct  4332953  TTCTTGGTGGAACAAATGTCCCGCTTCCAAATAATCAAAATAT  4332911
```

Features in this part of subject sequence:
conserved hypothetical protein, collagen triple helix rep...

Figure 8: Blast search of of isolate Ba#35 Bams15 sequence (indicated as Query; PCR product of 568 bp) with Sterne sequence AE017225.1 (indicated as Sbjct). The repeat unit of Bams 15 in isolates Ba#35 translates to 41 copy units. The sequence was edited and blasted in NCBI nucleotide database.

Appendix Figure 9: **A** **B**

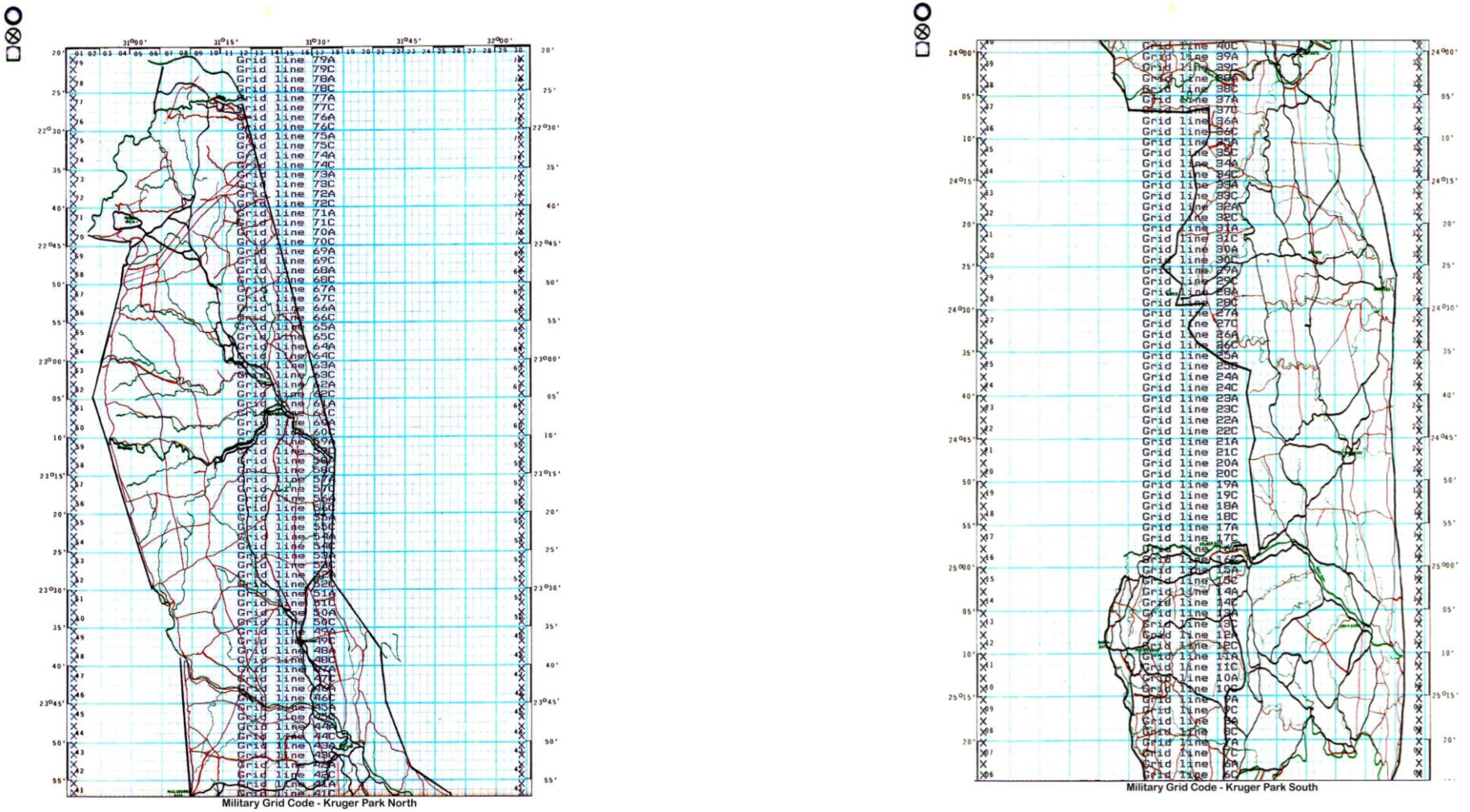
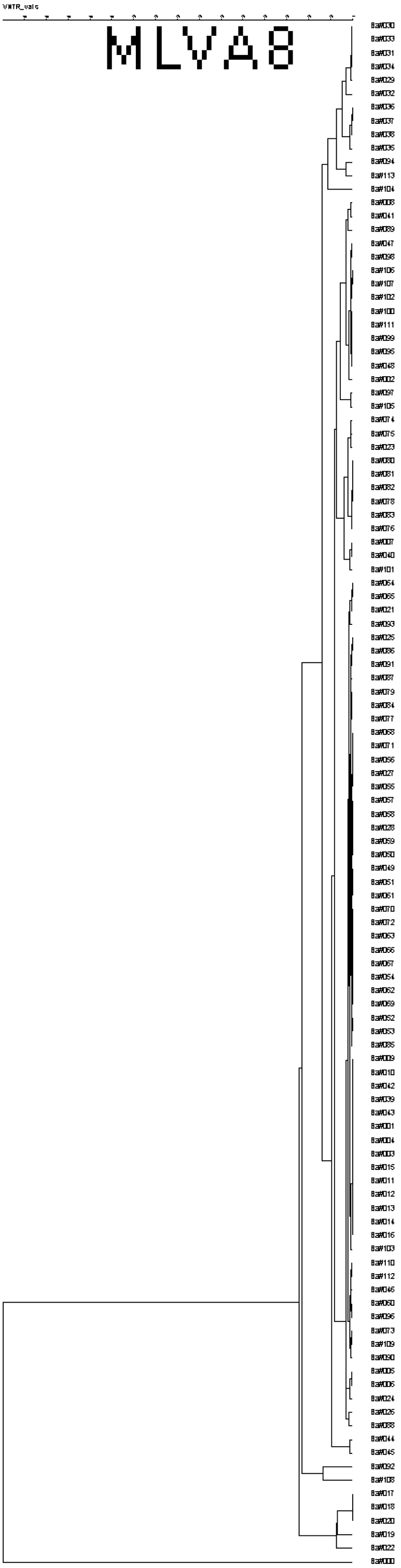
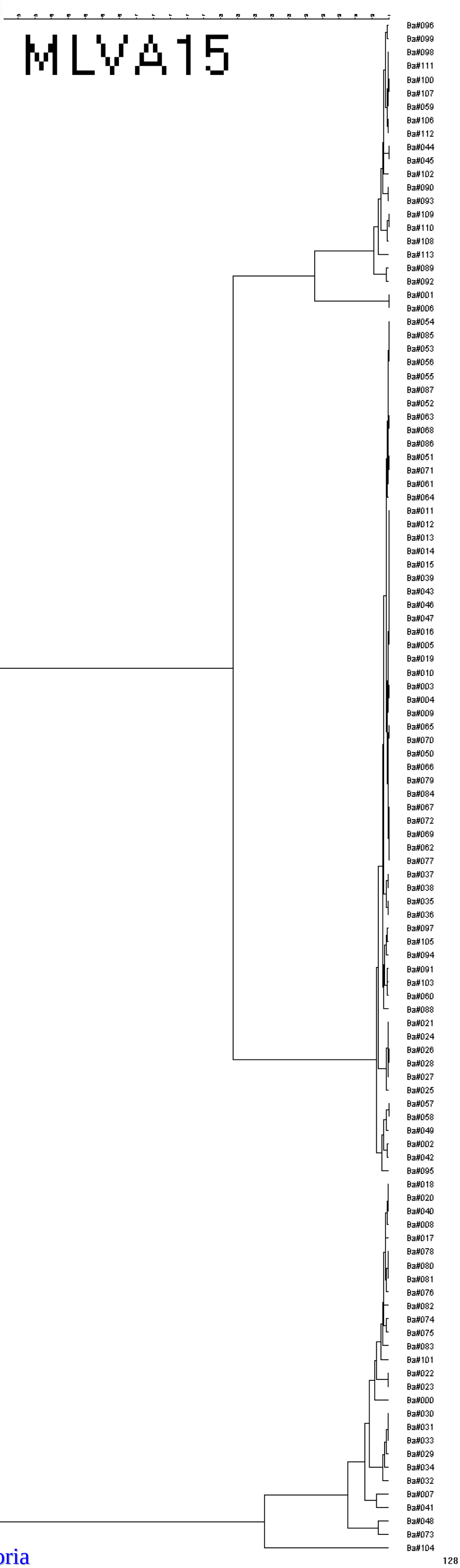


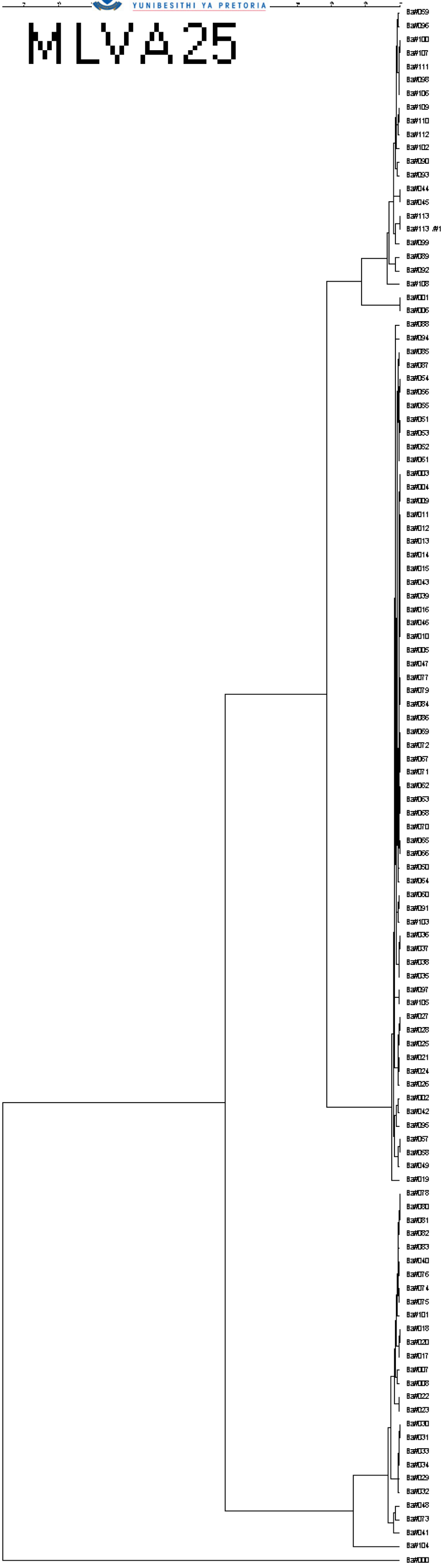
Figure 9: **A** = Map of northern region of the Kruger National Park (KNP) that was used to translate the 4 digit military grid code (see Table 1) into useable lattitude and longitude co-ordinates. **B** = Map of southern region of the Kruger National Park (KNP) that was used to translate the 4 digit military grid code (see Table 1) into useable lattitude and longitude co-ordinates.





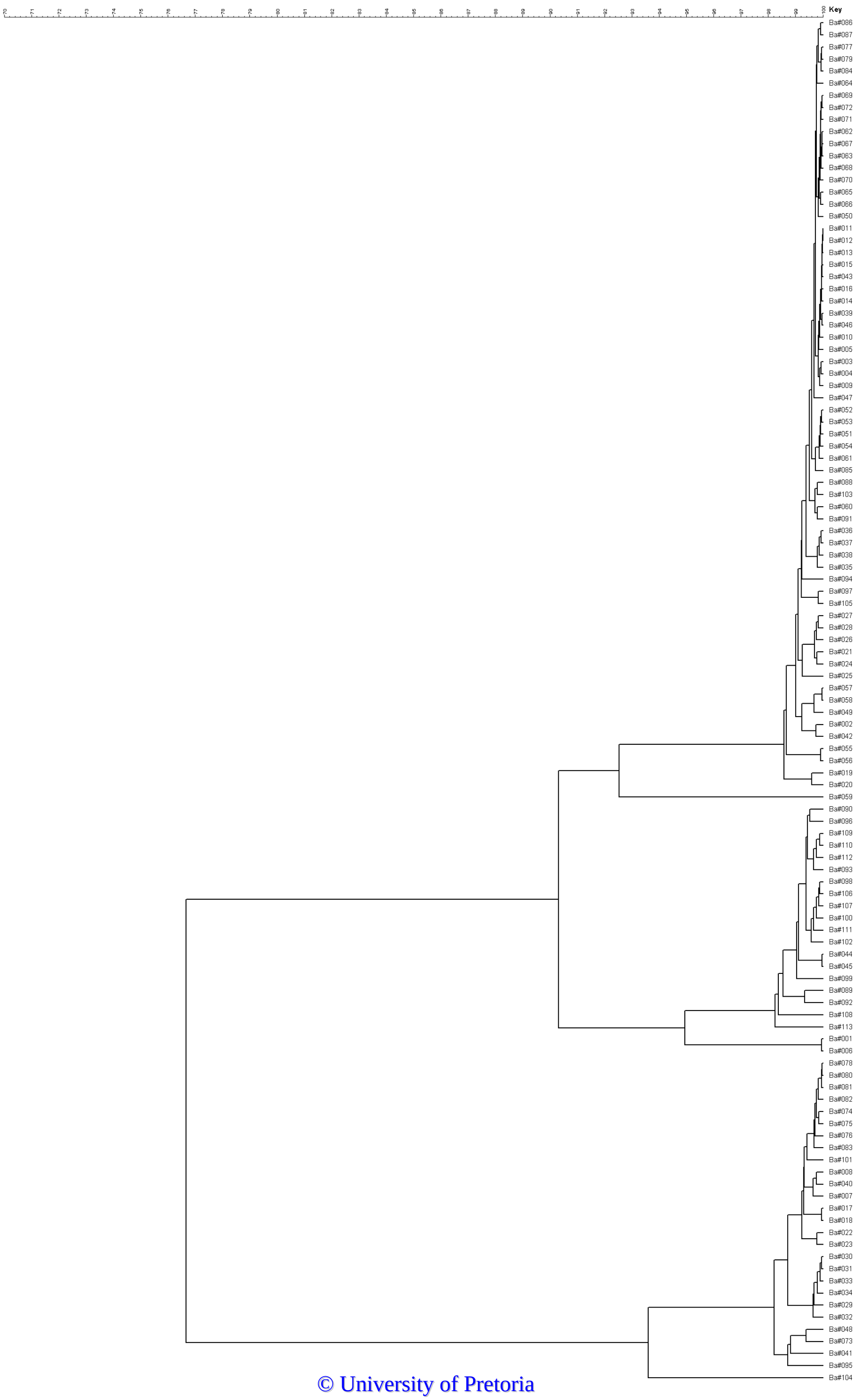
UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

MLVA25



MLVA31Agarose (113 entries)

VNTR_vals



MLVA 31Capillary (113 entries)