DEVELOPMENT OF MOLECULAR EPIDEMIOLOGICAL TOOLS FOR BETTER UNDERSTANDING AND CONTROL OF BURULI ULCER

BY

ISAAC FRIMPONG ABOAGYE

(10097405)

THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA, LEGON IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE AWARD OF PHD ZOOLOGY DEGREE

JULY, 2013

DECLARATION

This thesis is the result of research work undertaken by Isaac Frimpong Aboagye in the Departments of Animal Biology and Conservation Science and Bacteriology, Noguchi Memorial Institute for Medical Research, all of University of Ghana, Legon under the supervision of Professor Dominic Edoh and Dr. Anthony Ablordey.

Isaac Frimpong Aboagye

(Candidate)

Professor Dominic A. Edoh

(Principal Supervisor)

Dr. Anthony S. Ablordey

(Co-Supervisor)

DEDICATION

This thesis is dedicated first and foremost to God and to all stakeholders whose collective effort in the control of Buruli ulcer is yielding positive results.

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge with profound gratitude the assistance provided by all those who, in diverse ways, contributed to the success of this thesis.

I first and foremost wish to express my sincere appreciation to my supervisors: Dominic Edoh, a professor at the Department of Animal Biology and Conservation Science, University of Ghana and Anthony Ablordey, Senior Research Fellow, Noguchi Memorial Institute for Medical Research (NMIMR), University of Ghana for their constructive criticisms and the valuable suggestions and comments made during the study. I am also grateful to Timothy Stinear, a professor at the University of Melbourne, Australia for the vital input he made during the data analysis. It is gratifying to note that they were of tremendous help to me throughout the study.

Special appreciation and gratitude go to the sponsors: Flemish Inter-University Council, Belgium for sponsoring the sequencing and VNTR studies and the European Foundation Initiative for Neglected Tropical Diseases, International Association of National Public Health Institutes and Health and Labour Sciences Research Grant from the Ministry of Health, Labour and Welfare, Japan for sponsoring the LAMP study. I am also grateful to Nana Ama Amissah and Evans Ahorttor for their assistance during laboratory work.

Finally, I wish to thank my family and all those who helped me in diverse ways, but whose names have not been mentioned.

TABLE OF CONTENTS

LIST OF TABLES

PAGE

LIST OF FIGURES

LIST OF PLATES

PAGE

LIST OF ABBREVIATIONS

AFB	Acid-fast bacilli
ATP	Adenosine triphosphate
APH	Agogo Presbyterian Hospital
AFLP	Amplified fragment length polymorphism
AAND	Asante Akim North District
BU	Buruli ulcer
BUD	Buruli ulcer disease
BCG	Bacillus Calmette-Guerin
CDS	Coding sequences
CCD	Charge-coupling device
CFU	Colony-forming unit
CISs	Control ion spheres
DNA	Deoxyribonucleic acid
ddNTPs	Dideoxynucleotides
DDH	DNA-DNA hybridization
ECM	Extracellular matrix
FNA	Fine-needle aspiration
GBUI	Global Buruli Ulcer Initiative
ISE	Insertion sequence elements
IS*2404* PCR	Insertion sequence *2404* polymerase chain reaction
ISFET	Ion-sensitive field-effect transistors

ISPs	Ion sphere particles
LIF	Laser-induced fluorescence
LAMP	Loop-mediated isothermal amplification
MLSA	Multi-locus sequence analysis
MIRU	Mycobacterial interspersed repetitive unit
MU	*Mycobacterium ulcerans*
MPM	Mycolactone-producing mycobacteria
MURDs	*M. ulcerans* regions of difference
NGS	Next-generation sequencing
NMIMR	Noguchi Memorial Institute for Medical Research
NTM	Nontuberculous mycobacteria
PGM	Personal Genome Machine
PBS	Phosphate buffered saline
pw-LAMP	Pocket warmer loop-mediated isothermal amplification
PKS	Polyketide synthases
RFLP	Restriction fragment length polymorphism
rRNA	Ribosomal ribonucleic acid
SNP	Single nucleotide polymorphism
TB	Tuberculosis
VNTR	Variable number of tandem repeat
WHO	World Health Organisation
ZN	Ziehl-Neelsen

ABSTRACT

The limited ability of molecular typing methods to differentiate *Mycobacterium ulcerans* strains from the same geographic areas, particularly at the local level, hampers our understanding of Buruli ulcer (BU) epidemiology and calls for the search for molecular epidemiological tools with higher discriminatory power. Moreover, the current available laboratory tests for BU diagnosis have various limitations, ranging from low diagnostic sensitivities to financial and technical constraints associated with their use, that drive the search for simpler, cost-effective, rapid and reliable point of care diagnostic assays to facilitate effective control of Buruli ulcer especially in resource-poor countries. This study used two genome-wide approaches to index genetic diversity in *M. ulcerans* (MU). In a resequencing approach, whole genome sequence of *M. ulcerans* isolates were compared with a reference sequence for the identification of single nucleotide polymorphic sites. The study characterized whole genome single nucleotide polymorphisms (SNPs) of MU isolates using Ion Torrent and Illumina sequencing technologies and assessed the discriminatory power of variable number of tandem repeat (VNTR), based on newly-described loci, in typing *M. ulcerans* from Buruli ulcer endemic areas. It also assessed the diagnostic value of the pocket warmer loop-mediated isothermal amplification (pw-LAMP) in comparison with IS*2404* PCR and the conventional LAMP for *M. ulcerans* detection in clinical samples. Whole genome sequencing of MU isolates of different endemic areas in Ghana and Benin discovered a total of 299 SNP positions in comparison with the reference isolate, Mu_Agy99, with genomic regions of high SNP variability. Phylogenetic analysis based on SNP numbers inferred by neighbor-joining method produced five distinct *M. ulcerans* clusters with high confidence in clustal branches: two for the isolates from the Asante Akim North District and one each for the isolates from other endemic areas studied. Out of the 25 newly-described

VNTR loci tested, DNA amplification was observed in 23 (92%) of them with 15 (60%) VNTR loci amplification in 14 (82%) of the *M. ulcerans* isolates. The DNA band sizes observed for all loci were estimated to range between 300 bp and 750 bp, but all isolates had similar estimated band sizes for each locus. The results for the pw-LAMP, conventional LAMP and IS*2404* PCR were comparable, yielding 90.5% sensitivity for the pw-LAMP and 100% sensitivity for both the conventional LAMP and IS*2404* PCR when purified DNA extracts were used with no significant difference (χ^2 [1] = 0.40, p = 0.580). Specificities recorded for all three assays were 100%. Further genotyping study based on identified VNTR loci in combination with sequencing may provide useful information on the diversity of *M. ulcerans* and the epidemiology of Buruli ulcer. The pw-LAMP method for *M. ulcerans* DNA amplification is as efficient as that for IS*2404* PCR. The genomic regions of high SNP variability identified in this study may be useful targets for genotyping to differentiate *M. ulcerans* isolates with unknown SNP characteristics. Analysis of SNPs based on DNA sequencing of whole genomes using Ion Torrent and Illumina sequencing technologies is useful for MU strain discrimination at local level.

CHAPTER ONE

1.0 Introduction

1.1 General background to the study

Buruli ulcer (BU) is a severe necrotizing skin infection caused by *Mycobacterium ulcerans*. The discovery of BU dates back to 1897 when Sir Albert Cook, a British physician, first described skin ulcers in Uganda consistent with BU. However, the first definitive description of *Mycobacterium ulcerans* was published in 1948 (MacCallum *et al*., 1948). In that report, lesions in different stages of BU were described in few Australian victims. The disease has since been reported from several different regions of the world including eastern and western Africa.

The last few decades have witnessed considerable increase in the number of cases of BU reported in several parts of the world, particularly in sub-Saharan Africa where the largest numbers of patients with Buruli ulcer have been detected (Hayman, 1991; Portaels, 1995). Although many countries of sub-Saharan Africa are considered endemic for Buruli ulcer, the largest number of patients has been reported from riverine areas in distinct regions of Benin, Côte d'Ivoire and Ghana, where the number of detected cases has alarmingly increased in recent years (van der Werf, *et al*., 2005). In Côte d'Ivoire, approximately 24, 000 cases were recorded between 1978 and 2006 with Benin and Ghana recording nearly 7,000 cases (1989 – 2006) and more than 11, 000 cases (since 1993) respectively (WHOMC, 2007).

The increasing geographical distribution, the severe and devastating consequences of Buruli ulcer as well as the knowledge gap in several aspects of the disease prompted the World Health Organization to establish the Global Buruli Ulcer Initiative (GBUI) in February 1998

to coordinate Buruli ulcer control and research efforts. The GBUI is dedicated to raising awareness about the disease, improve access to early diagnosis, treatment and promotion of research to develop better tools for the treatment and prevention of Buruli ulcer. In this regard, the GBUI has made significant progress since its inception and has been acknowledged with satisfaction by the World Health Assembly (WHO, 2004) and the Cotonou Declaration on Buruli ulcer (WHO, 2009).

In spite of the progress made, BU is reported in more than 33 countries in Africa, the Americas, Asia and the Western Pacific, mainly in tropical and subtropical regions (Johnson *et al*., 2005; WHO, 2008b). It commonly affects poor people in remote rural areas with limited access to health care (WHO, 2008a). Although the disease can affect all age groups, children under the age of 15 years are predominantly affected (WHO, 2008a). It is the concern of all affected countries and other stakeholders that with the increasing number of cases and the associated health complications, the disease could hinder efforts to improve the economic and social development of the communities most affected. It is, therefore, imperative to encourage a lot more research in the grey areas of the disease to help address this emerging neglected tropical disease that is increasingly assuming public health importance.

1.2 Specific background to the study

Following the establishment of the GBUI, a lot of progress has been made in research on several aspects of Buruli ulcer management and control. These activities have generally contributed immensely to our understanding of the disease. However, the existing knowledge gap in key aspects of the disease such as chemotherapy for all stages of the disease, the

mystery surrounding the natural reservoir of *Mycobacterium ulcerans* (MU), mode of transmission and pathogenesis of the disease has hampered the global effort in mounting effective control strategies.

The diagnosis of BU by PCR is based on the detection of IS*2404*, an insertion sequence with more than 200 copies on the chromosome of *M. ulcerans* (Stinear, 2007). The IS*2404* PCR has also been useful for the detection of MU deoxyribonucleic acid (DNA) in various biotic and abiotic environmental aquatic samples, raising specific questions about the reservoir status of MU in these samples and their role in the transmission of BU to humans and other mammals. Genotyping methods particularly VNTR-based ones hold considerable promise for throwing light on the evolution and diversity of *Mycobacterium ulcerans* and can impact positively on our understanding of the reservoirs and transmission chains of *M. ulcerans*.

Earlier studies on highly monomorphic and clonal species such as *Mycobacterium tuberculosis* (Frothingham and Meeker-O'Connell, 1998; Mazars *et al*., 2001; Roring *et al*., 2002; Skuce *et al*., 2002; Supply *et al*., 2000), *Bacillus anthracis* (Keim *et al*., 2000) and *Yersinia pestis* (Klevytska *et al*., 2001) using variable number of tandem repeat (VNTR) genotyping methods have been very useful in differentiating and studying genetic relationships of strains of these species. VNTR-based genotyping methods have also made important contributions to the differentiation of *M. ulcerans* isolates (Stragier *et al*., 2006; Ablordey *et al*., 2005a, b; Hilty *et al*., 2006) and appear to give more information on their diversity compared to other standard genotyping methods.

Sequencing technology also has useful application in studying genetic diversity in *M. ulcerans*. Single nucleotide polymorphism (SNP) analysis, based on DNA sequencing of

whole genomes and selected SNP loci, has greatly improved our understanding of the diversity and evolution of pathogenic mycobacteria such as *M. ulcerans*. The knowledge of the evolution of *M. ulcerans* isolate (Agy99) from *M. marinum* M strain (Stinear *et al*., 2000; Yip *et al*., 2007), confirmation of the existence of five *M. ulcerans* haplotypes globally and two major *M. ulcerans* lineages (Kaser *et al*., 2007) and the identification of mechanisms that explain genomic differences in haplotypes of the classical and ancestral lineages (Kaser & Pluschke, 2008) were based on comparative genome sequencing techniques. Genetic discrimination of *M. ulcerans* clinical isolates at the local level (Roltgen *et al*., 2010) using sequencing techniques has also been demonstrated.

By exploring further the discriminatory power of VNTR genotyping method and characterizing whole genome SNPs, the genetic diversity of *M. ulcerans* isolates would be better understood. Such information can give insight into reservoirs and transmission chains of *M. ulcerans* when the spreading patterns of variants in space and time are studied.

1.3 Problem and Justification of the Study

Buruli ulcer (BU) is an indolent, necrotizing disease of the skin, subcutaneous tissue and occasionally bones (Portaels and Meyers, 2006) and is caused by *Mycobacterium ulcerans*. The disease begins as a painless nodule or papule that, if left untreated, can lead to extensive ulceration that could cover 15% of the body (George *et al*., 1999). It typically affects impoverished inhabitants, primarily children, of remote areas where medical services are unavailable or too expensive (Guedenon *et al*., 1995).

The prolonged periods of hospitalization, high costs of treatment and the devastating nature of complications of Buruli ulcer (Asiedu and Etuaful, 1998) have enormous adverse socioeconomic implications. Although the disease has a very low mortality rate, it can lead to profound morbidity in victims. The associated morbidity and high cost of treatment pose a serious challenge to a struggling rural economy and its health system.

In children, the prolonged morbidity and hospitalization could lead to serious disruption of school and, in some cases, school drop outs. In more severe cases involving complications such as amputations and contracture deformities, victims of Buruli ulcer may not be able to work and eventually become a burden on society. The effect of the disease on women can have more serious consequences given the important role women play in families and rural communities. Buruli ulcer has the potential of disrupting the family life of affected women, and their income-generating activities, health and welfare of their children may also be negatively affected.

The disease is complicated by the absence of an effective vaccine and chemotherapy for all stages of the disease as well as the poor understanding of the environmental reservoir and the mode(s) of transmission of *M. ulcerans*, making effective prevention and control of BU particularly difficult. As a result, the number of cases has been rising especially in West Africa, with a prevalence exceeding that of tuberculosis (TB) in certain areas (Debacker *et al*., 2004a). A national survey in Ghana in the year 1999 identified over 6,000 cases (i.e. 20.7 per 100 000), making BU the second most prevalent mycobacterial disease after TB (Amofah *et al.*, 2002).

VNTR-based genotyping methods have demonstrated to be useful in the genetic discrimination of *M. ulcerans* isolates (Stragier *et al*., 2006; Ablordey *et al*., 2005a, b; Hilty *et al*., 2006) and have the potential of differentiating MU isolates and providing better understanding of the genetic diversity of isolates of the same or different geographic regions. Again, the increasing availability of high-throughput sequencing technologies is expected to improve the discovery of *M. ulcerans* variants on account of their superiority to Sanger sequencing due to the high volume of sequence data produced (Hall, 2007; Church, 2006) in a relatively short time (Karl *et al*., 2009) at a much cheaper cost (Quail *et al*., 2012; Liu *et al*., 2012). Such information can give insight into reservoirs and transmission chains of *M. ulcerans* when the spreading patterns of variants in space and time are studied. Buruli ulcer recurrence rates have been observed to vary from 6% to >20% (Debacker *et al*., 2005). However, the degree to which the recurrence is attributable to dissemination of the pathogen from previous lesions or exogenous reinfection is unknown. The different genomic profiles of *M. ulcerans* isolates can also be useful in distinguishing recurrences from new infections.

This study seeks to genetically differentiate *M. ulcerans* isolates within an endemic area using VNTR-based genotyping method. Whole genome sequencing of *M. ulcerans* isolates will also be carried out to assess their diversity based on SNP analysis. The provision of new knowledge on the diversity of *M. ulcerans* isolates would be useful in tracking the transmission pattern of *M. ulcerans* and distinguishing recurrences from new infections. It will not only contribute to the body of knowledge on BU, but also be beneficial to BU endemic communities, the international community, among other stakeholders.

1.4 Hypothesis of the study

- Variable number of tandem repeat (VNTR) genotyping method can genetically differentiate *M. ulcerans* isolates within Buruli ulcer endemic area.

1.5 Aims and objectives of the study

This study aims at further assessment of the discriminatory power of VNTR in typing *M. ulcerans* isolates from Buruli ulcer endemic area. It also aims at assessing the diversity of *M. ulcerans* isolates based on whole genome sequencing and SNP analysis. The specific objectives of the study are as followed:

- ❑ To determine DNA amplification of newly-described VNTR loci in *M. ulcerans* (MU) and to assess their importance in typing MU for the purpose of genetic differentiation.
- ❑ To determine the genetic diversity of *M. ulcerans* isolates based on DNA banding patterns of the respective amplicons using VNTR typing method.
- ❑ To characterize single nucleotide polymorphisms (SNPs) in *Mycobacterium ulcerans* isolates.
- ❑ To assess the diversity of *Mycobacterium ulcerans* isolates based on SNP variation.
- ❑ To determine the distribution of MU variants in the Asante Akim North District Buruli ulcer endemic area.

1.6 Introduction of loop-mediated isothermal amplification (LAMP) study

The LAMP is a DNA amplification method completed in an hour or less under isothermal conditions (Wang *et al*., 2008). DNA amplification is carried out with high specificity, efficiency and rapidity using a set of four specially designed primers and a DNA polymerase with high strand displacement activity (Notomi *et al*., 2000; Nagamine, *et al*., 2002). Current available diagnostic laboratory tests include microscopic examination, culture, IS*2404* PCR of swab samples and tissue specimens, and histopathological analysis. However, the respective limitations of their diagnostic efficiency and the financial and technical constraints associated with their use reduce their diagnostic value. The result is that patients do not get much benefit because of the generally low diagnostic sensitivity for microscopy and culture (Herbinger *et al*., 2009) and the limited application of PCR and histopathology particularly in resource-poor countries because of high cost and technical implications.

IS*2404* PCR is considered the most sensitive method for laboratory confirmation of the disease (Beissner *et al*., 2010; Herbinger *et al*., 2009; Stienstra *et al*., 2003). It also has very high specificity (Ross *et al*., 1997; Rondini *et al*., 2003; Siegmund *et al*., 2005) and fairly good positivity ratios (Siegmund *et al*., 2007; Mensah-Quainoo *et al*., 2008). It is, therefore, considered to be the most reliable technique for the detection of *M. ulcerans* in human diagnostic samples. Despite these advantages, PCR detection is rarely available as a routine diagnostic tool especially in resource-poor countries where there are inadequate laboratory facilities and trained personnel as well as financial constraints.

The loop-mediated isothermal amplification (LAMP) assay is noted for its high specificity, efficiency and rapidity in DNA amplification (Notomi *et al*., 2000; Nagamine, *et al*., 2002). Gene amplification by the LAMP method is known to be superior to that by PCR (Goto *et al*.

2009). It is more sensitive (Kuboki *et al*., 2003), specific (Wang *et al*., 2008) and has the advantage in specificity, selectivity and rapidity over other nucleic acid amplification methods (Mori *et al*., 2001). Its simplicity and cost-effectiveness (Kuboki *et al*., 2003), overall efficiency in DNA amplification and the potential to be useful in field conditions (Thekisoe *et al.*, 2009) highlight the potential clinical importance of LAMP as a diagnostic tool.

1.7 Problem and justification of the LAMP study

Laboratory confirmation of clinically suspected cases became crucial for the clinical management of BU with the introduction of antibiotic treatment in 2004. The current available laboratory tests for BU diagnosis include microscopic examination, culture, IS*2404* PCR of swab samples and tissue specimens and histopathological analysis. These tests have various limitations, ranging from diagnostic efficiency to financial and technical constraints, associated with their use. In spite of its rapidity and affordability for BU diagnosis in many endemic areas, microscopic detection of acid-fast bacilli in swab and tissue smears is known to have low diagnostic sensitivity, detecting 29%–78% of clinically suspected BU cases (Herbinger *et al*., 2009; Bretzel *et al*., 2011). The detection rate of culture is comparable to that of microscopy, ranging from 34% to 79% (Herbinger *et al*., 2009) and takes an average of 9–12 weeks to yield positive results. Culture, therefore, cannot be used for rapid laboratory confirmation of BU.

Various studies comparing different diagnostic tests applied on different specimens suggest a wide range of diagnostic sensitivities, varying from as low as < 40% (swab smear microscopy) to > 95% (PCR and histopathological analysis of tissue specimens) (Guimaraes-Peres *et al*., 1999; N'Guessan *et al*., 2001; WHO 2001; Whitney *et al*., 2002). However, PCR

and histopathological analysis are rarely available in countries where Buruli ulcer disease is endemic. Histopathologic analysis is restricted to external reference laboratories and is unavailable in peripheral health centres, district or regional hospitals.

The application of PCR techniques in tropical regions presents technical difficulties. Transportation and storage of reagents require cold chains that are often interrupted by frequent power outages. Moreover, conventional PCR assays require careful handling of reaction components by skilled laboratory workers as well as strict measures to avoid contamination. These limitations drive the search for a simpler, cost-effective and efficient diagnostic tool for BU in resource-poor countries.

This study seeks to develop a field-based diagnostic method for *M. ulcerans* using disposable pocket warmer based on the LAMP method. As described by Hatano and colleagues (2010), the procedure for the pocket warmer LAMP (pw-LAMP) involves the placing of the reaction tubes on a pocket warmer, which is then folded and surrounded by a paper towel. This set up is put into a Styrofoam box with a thermometer and incubated for 90 minutes during which enough heat will be generated for the LAMP reaction without the need for electricity.

The pw-LAMP is a new, highly mobile and sensitive method that has been used to amplify DNA from *B. anthracis* (Hatano *et al*., 2010). The most significant advantage of pw-LAMP is its high mobility. It does not require any heavy equipment such as heat blockers or thermal cyclers, as it does not need any electric power. It is inexpensive, costing less than 100 Japanese Yen each (Hatano *et al*., 2010) equivalent to 3.36 Ganaian Cedis. Such low cost is a big advantage especially in developing countries where financial constraint is a major factor

in the management of BU. The pw-LAMP also has the potential to provide detection tool for BU in endemic areas that do not have functional or reliable electricity infrastructures.

The development of pw-LAMP method for *M. ulcerans* DNA amplification is expected to provide useful information about its sensitivity and potential for use for field-based diagnosis of Buruli ulcer in endemic and resource-poor countries.

1.8 Hypothesis of the LAMP study

The pw-LAMP method for *M. ulcerans* DNA amplification is as sensitive as that for IS*2404* PCR.

1.9 Aim and objectives of the LAMP Study

To develop a simple and rapid diagnostic test for Buruli ulcer under field conditions using modified conventional LAMP assay (pw-LAMP). The specific objectives are:

- To determine the sensitivities and specificities of IS*2404* PCR, conventional LAMP and pw-LAMP for *M. ulcerans* detection.

- To assess the sensitivity and specificity of IS*2404* PCR compared with those of pw-LAMP for *M. ulcerans* detection using crude and purified DNA templates.

- To assess the sensitivity and specificity of IS*2404* PCR compared with those of conventional LAMP for *M. ulcerans* detection using crude and purified DNA templates.

- To compare the positivities` in crude DNA specimens with those in purified DNA and assess their significance in *M. ulcerans* detection.

- To determine optimal temperature required for the detection of *M. ulcerans* DNA.

CHAPTER TWO

2.0 Literature Review

2.10 Taxonomic classification and characteristics of Mycobacteria

Mycobacteria are a group of bacteria classified by Lehmann and Neumann (1896) as follows;

Kingdom: Bacteria
Phylum: Actinobacteria
Class: Actinomycetes
Order: Actinomycetales
Suborder: Corynebacterineae
Family: Mycobacteriaceae
Genus: *Mycobacterium*

The classification is based on characteristics common to members of this group of bacteria. All *Mycobacterium* species share a characteristic cell wall consisting of hydrophobic mycolate layer and a peptidoglycan layer held together by a polysaccharide, arabinogalactan (Fig. 2.1).

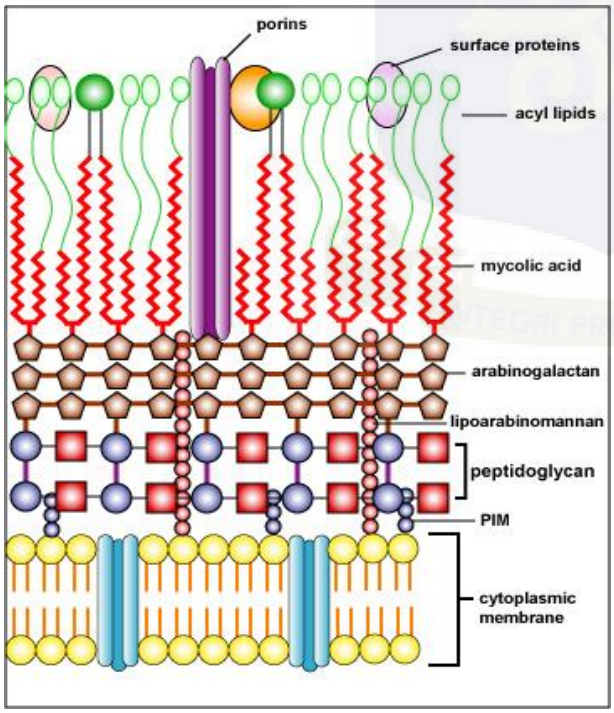


Fig.2.1: Structure of Mycobacterial cell wall (*student.ccbcmd.edu*)

The hydrophobic and waxy cell wall, which is rich in mycolic acids, makes a substantial contribution to the hardiness of this genus (Bhamidi, 2009). The use of the Latin prefix *"myco"* in this group of bacteria could be related to the waxy compounds that make up parts of the cell wall.

With the exception of *Mycobacterium marinum*, which is motile within macrophages, mycobacteria are aerobic and nonmotile bacteria that are acid-alcohol fast (Ryan and Ray, 2004), a characteristic that has to do with their ability to resist decolourization by acidified alcohol following staining with a basic fuschin dye. They are classified as acid-fast Gram-positive bacteria due to their lack of an outer cell membrane.

The minimal standards for including a species in the genus *Mycobacterium* are 1) acid-alcohol fastness 2) the presence of mycolic acids containing 60-90 carbons, which are cleaved to C22 to C26 fatty acid methyl esters by pyrolysis and 3) a guanine plus cytosine (G+C) content of DNA in the range of 61 to 71 mol % except *Mycobacterium leprae*, which has G+C content of 54 to 57 mol % (Levy-Frebault and Portaels, 1992). The differences in growth rates and pigmentation properties form the basis of the Runyon's classification. As stated by Shinnick and Good (1994), there are 71 recognized or proposed species of *Mycobacterium*, which can be divided into two main groups based on growth rate. The slowly growing species require > 7 days to form visible colonies on solid media while the rapidly growing species require < 7 days. The slowly growing species have been grouped into three based on pigmentation. On exposure to light, photochromogenic species may synthesize yellow to red carotenoid pigments, but nonphotochromogens are unable to develop pigmentation under any light conditions. Scotochromogenic species develop pigmentation under both light and dark conditions. For the purpose of diagnosis and treatment,

mycobacteria can be classified into several major groups: *M. tuberculosis* complex (*M. tuberculosis, M. bovis, M. africanum and M. microti), M. leprae* and Nontuberculous mycobacteria (NTM) that are all other mycobacteria including mycolactone-producing mycobacteria.

2.11 Mycolactone-producing Mycobacteria

All mycolactone-producing mycobacteria (MPM) share a lot of genetic and phenotypic characteristics. They have < 1% nucleotide variation, possess pMUM plasmids, contain the insertion sequence IS*2404* and make mycolactone (Pidot *et al*., 2010). As revealed by multi-locus sequence analysis (MLSA) and InDel analysis, they have > 98% 16S rRNA identity to *Mycobacterium ulcerans*, > 70% DNA-DNA hybridization (DDH) consistent with the findings of Vandamme *et al*. (1996). Clearly, all MPM share sufficient phenotypic and genotypic characteristics on the basis of which Pidot *et al*. (2010) described them as variants of the same species, *M. ulcerans*. They have, therefore, proposed reclassification of all MPM as *M. ulcerans*.

2.20 *Mycobacterium ulcerans* genome

The genome of *Mycobacterium ulcerans* is composed of 5,805,761 bp consisting of two circular replicons, a 5,631,606-bp chromosome (Fig. 2.2) and a 174,155-bp plasmid, pMUM001 (Fig. 2.3). The chromosome contains 4160 coding sequences (CDS), 771 pseudogenes, two bacteriophages, and 302 insertion sequence elements (ISE), including 209 complete or partial copies of IS*2404* and 83 copies of IS*2606* (Stinear *et al.*, 1999; 2007). Genome comparison with the completed 6,636,827-bp genome sequence of *M. marinum*, with its 5426 CDS revealed the existence of many deletions, accounting for 1064 kb, and

many DNA rearrangements in *M. ulcerans* (Stinear *et al.*, 2007). All these data indicate that *M. ulcerans* has recently evolved via lateral gene transfer and reductive evolution from the more rapid-growing environmental species *M. marinum* to become a niche-adapted specialist. Via plasmid acquisition, among other more subtle genomic changes, *M. ulcerans* has evolved adaptive advantages for new environments, adaptations that may also present a pathogenic phenotype.

ISEs

Figure 2.2: *Mycobacterium ulcerans* chromosome

The plasmid, which has an average G+C content of 62.5%, contains 81 protein-coding sequences (CDS) and encodes polyketide synthases and polyketide-modifying enzymes (Stinear *et al.*, 2004, 2005a, b). As reported by Stinear *et al*. (2005b), over half of this circular plasmid is devoted to six genes, three of which encode giant polyketide synthases (PKS) that produce mycolactone, an unusual cytotoxic lipid produced by *M. ulcerans*, and suggested to be responsible for the extensive tissue damage and immunosuppression (George *et al.*, 1999).

The genes *mlsA1* (50,973 bp) and *mlsA2* (7,233 bp) encode mycolactone core-producing polyketide synthethase (PKS) and mlsB (42,393 bp) encoding side chain enzymes.

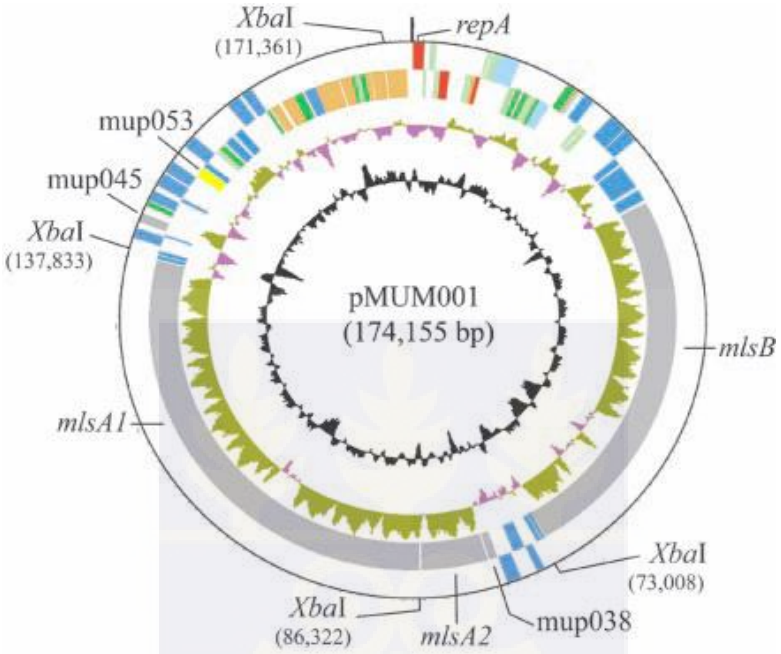


Figure 2.3: *Mycobacterium ulcerans* plasmid, pMUM001

2.30 History and global distribution of Buruli ulcer

The discovery of Buruli ulcer dates back to 1897, when Sir Albert Cook, a British physician, first described skin ulcers in Uganda consistent with Buruli ulcer (BU). However, the first definitive description of *Mycobacterium ulcerans* was published in 1948 (MacCallum *et al*., 1948), following the discovery of acid-fast bacilli in a biopsy from ulcerative lesions in a young Australian in 1940. In this report, lesions in different stages of Buruli ulcer were described in few Australian victims in a riverine area in Bairnsdale, Victoria. Between the

years 1923 and 1935, Kleinschmidt, a missionary physician in north-east Congo observed undermined skin lesions rich in acid-fast bacilli (Meyers *et al*., 1974a).

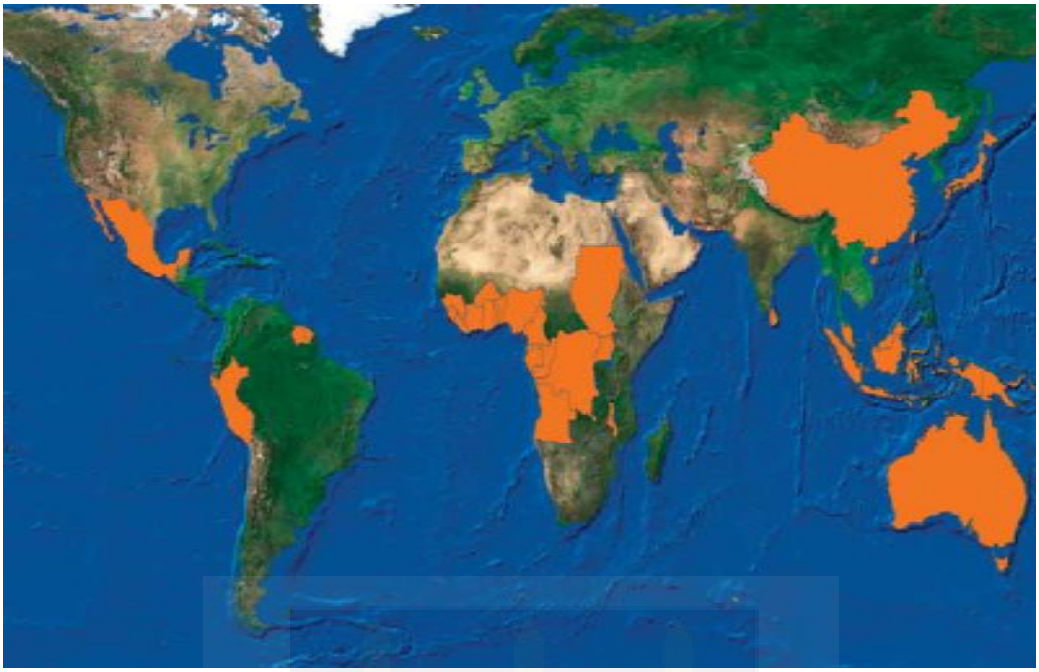
Prior to 1980, there were many reports on the disease in several African countries including Cameroon (Ravisse, 1977), Uganda (UBG, 1969-71) and Gabon (Burchard *et al*., 1986). One of the earliest reports on BU came from Zaire now called the Democratic Republic of the Congo, from the area south-west of Kinshasa (van Oye and Ballion, 1951) where it is still prevalent (Bafende, 2004). In the 1960s, many cases occurred in Buruli County (Clancey *et al*., 1961) (now called Nakasongola District) in Uganda, giving rise to the name 'Buruli ulcer' (Barker, 1971). All of these cases occurred in villages in close proximity to the River Nile.

In West Africa, cases were reported in Nigeria (Oluwasanmi *et al*., 1975) and Sierra Leone (Gibson, 1975). In Ghana, the first probable case of BU was reported in the Greater Accra Region in 1971; the presence of additional cases along the tributaries of the Densu River in the area was considered likely (Bayley, 1971). In 1989, cases were described in the Asante Akim North District of the Ashanti Region (van der Werf *et al*., 1989) followed by the description of a major endemic focus in Amansie West District in the same region (Amofah *et al*., 1993).

The last three decades (after 1980) have witnessed the discovery of new foci of Buruli ulcer in several West African countries of which Cote d'Ivoire, Benin and Ghana are included. Although many sub-Saharan African countries are considered endemic for the disease, the largest numbers of patients have been reported from riverine areas in distinct regions of Benin (Debacker *et al*., 2004a), Côte d'Ivoire (Marston *et al*., 1995) and Ghana (Amofah *et*

al., 2002; Raghunathan *et al*., 2005) where the number of detected cases has alarmingly increased in recent years. New foci of the disease have also been discovered in Togo (Meyers *et al*., 1996), Angola (Bär *et al*., 1998; Kibadi *et al*., 2008), Democratic Republic of Congo (Kibadi *et al*., 2008), among some other African countries.

The disease is focally endemic in Africa, the Americas, Australia, and Asia, where rural populations are the most affected (Janssens *et al*., 2005; Sizaire *et al*., 2006; WHO, 2008). However, relatively few cases have been reported in regions other than the African continent. Figure 2.4 shows the regions with reported cases of Buruli ulcer.



Regions with reported cases of Buruli ulcer

Fig. 2.4: Global geographical distribution of Buruli ulcer (WHO/GBUI, 2005)

West Africa: Côte d'Ivoire, Ghana, Benin, Burkina Faso, Nigeria, Sierra Leone, Togo, Liberia, Guinea

Central Africa: Angola, Cameroon, Congo, Equatorial Guinea, Democratic Republic of Congo, Sudan, Uganda, Gabon, Central African Republic

Other African countries: Kenya, Malawi,

Asia: China, Japan, Malaysia, Indonesia, Sri Lanka

Western pacific: Australia, Papua New Guinea, Kiribati

Americas: Brazil, French Guiana, Mexico, Peru, Suriname, Bolivia

2.40 Mode of *M. ulcerans* transmission and risk factors for Buruli ulcer

2.41 Mode of *Mycobacterium ulcerans* transmission

The diagnosis of BU by PCR is based on the detection of IS2404, an insertion sequence with more than 200 copies on the chromosome of *M. ulcerans* (Stinear, 2007). The IS2404 is restricted to *M. ulcerans* and other mycolactone-producing mycobacteria (MPM), which have been described by Pidot *et al*. (2010) as variants of *M. ulcerans*. Several studies have detected IS2404 in environmental aquatic samples, indicating that *M. ulcerans* (MU) is probably present in such samples, and supporting the concept that MU is an environmental pathogen. The IS*2404* has been detected in aquatic plants (Marsollier *et al*. 2004), environmental samples such as water and detritus from swamps in Australia (Ross, 1997; Roberts and Hirst, 1997; Stinear *et al*., 2000), and more recently, mosquitoes in Australia (Fyfe *et al*., 2007). Aquatic insects of the families Belostomatidae, Naucoridae and Hydrophilidae, crustaceans and molluscs (*Bulinus* sp. and *Planorbis* sp.), as well as fish including *Tilapia* sp. in western tropical Africa have all been found to have IS*2402* (Portaels *et al*., 1999; Marsollier *et al*., 2002; Marsollier *et al*., 2004; Portaels *et al*., 2001; Kotlowski *et al*., 2004).

On the basis of these detections, Portaels and colleagues (2001) hypothesized that mycobacteria (*M. ulcerans*) present in water and mud are mechanically concentrated by small water-filtering organisms like microphagous fish, mosquito larvae, small crustacea or molluscs or even some protozoa such as amoeba (Drancourt *et al*., 2002) that can be ingested subsequently by water bugs, which become reservoirs of the bacteria. They further proposed the hypothesis that human beings as well as domestic and wild animals could be contaminated or infected by biting insects such as water bugs. Previous findings (Marsollier *et al*., 2002) describing the transmission of *M. ulcerans* from water bugs of the family Naucoridae to mice as well as the localization of *M. ulcerans* within the salivary glands of

these bugs following experimental infection of the insects are consistent with the possibility of this mode of transmission to humans.

Aquatic bugs are cosmopolite insects that fly from one nearby swamp or pond to another and are aggressive predators of other species of aquatic arthropods, molluscs and young fish. These insects can inflict painful bites on humans in natural settings based on previous observation (Marsollier *et al*., 2007) that non-infected humans exposed to aquatic environments in BU endemic areas have higher titres of antibodies to salivary proteins of Naucoridae and Belostomatidae than BU patients in the same areas. However, there are suggestions that trauma seems essential for the introduction of *M. ulcerans* into the skin (Meyers *et al*., 1974b) following the development of lesions at the site of trauma to skin. The aetiological agent may, therefore, be introduced directly by bites of these insects or by trauma at skin sites contaminated by insect products containing *M ulcerans*.

Portaels and colleagues (2008) successfully cultivated MU from water strider (*Gerris* sp.) from Gerridae, another family of aquatic Hemiptera, which are also aggressive predators of other aquatic organisms such as insects and small fish. However, these insects may be only passive, incidental and transient reservoirs of *M. ulcerans* without an obvious role in the transmission of BU to humans and other mammals, as there are no reports of Gerridae biting humans.

Recent reports by some Australian BU victims that the appearance of BU at the site of what might have been a mosquito bite implicated mosquitoes in the transmission of *M. ulcerans* and therefore Buruli ulcer. This observation and the knowledge from field studies in Africa implicating aquatic insects in the mode of transmission prompted Johnson *et al*. (2007) to test

over 11,000 mosquitoes trapped from pools in Southeastern Australia. Of 11,504 mosquitoes tested, 48 pools were positive for IS*2404* present in *M. ulcerans* DNA. Interestingly, female *Aedes camptorhynchus* dominated (41 pools) with *Anopheles annulipes*, *Culex australicus* and *Ae. notoscriptus* testing positive in one pool each, suggesting that *M. ulcerans* contamination of mosquitoes is not species-specific. Because *M. ulcerans* is an environmental pathogen, PCR-positive mosquitoes may only be indicators of its presence in the environment rather than mosquitoes being a true productive reservoir and vector. The findings of this study do not demonstrate that mosquitoes are linked to or responsible for *M. ulcerans* transmission, but these possibilities require further investigation.

2.42 Risk factors for Buruli ulcer

Several sociodemographic and environmental factors have been identified to be important risk factors for BU. In a large case-control study by Debacker *et al*. (2006), the disease in endemic areas was found to be associated with age, unprotected water sources, and place of residence. The risk for BU was higher in children <15 years old in accordance with previous studies (WHO, 2000; Amofah *et al*., 1993; Marston *et al*., 1995; Debacker *et al*., 2004a; Debacker *et al*., 2004b), which indicated that children <15 years of age are at highest risk for acquiring BU and that in areas endemic for BU, exposure to unprotected water is a risk factor for the disease (WHO, 2000; Johnson *et al*., 2005). The lower risk of the disease observed in adults 15–49 years old, who are frequently exposed to wetlands, suggests acquired resistance to the disease and may be related to acquired specific immunity or to cross-immunity from other mycobacterioses (Smith, 1970). Housing near river, lake or dam, among other factors such as young children and women having daily water related activities have also been identified to be strongly associated with the disease (Ahoua *et al*., 2009). The differences

among sociodemographic variables with respect to the risk for BU may reflect differences in frequency and intensity of exposure to *M. ulcerans*.

Although no association was found between BU and BCG status in children <5 years old, the presence of BCG vaccination scar in persons >5 years old conveyed 2.5 risk for BU compared with those without BCG scar, confirming findings by (Floyd *et al*., 2000; Fine *et al*., 1989), that the disease has been found to develop in most BCG vaccinated persons. However, BCG vaccination partially protects against BU (UBG, 1969; Smith *et al*., 1997), but this protection appears to decrease after 6 months (UBG, 1969). The disparity in the results may reflect known variations of efficacy of BCG (Fine, 2001) or variation of BCG coverage for controls compared with BU patients (Wünsch-Filho *et al*., 1993).

2.50 Pathogenesis and clinical presentation of Buruli ulcer

2.51 Pathogenesis of Buruli ulcer

Following the introduction of *M. ulcerans* into the subcutaneous tissue, the organism proliferates extracellularly and produces a family of toxin molecules, mycolactones, which are required for the tissue destruction and local immunosuppression characteristic of Buruli ulcer (George *et al*., 1999). Biofilms for human bacterial pathogens such as *Pseudomonas aeruginosa*, *Haemophilus influenzae*, and *Vibrio cholerae* (Costerton *et al*., 1999) consist of discrete bacteria surrounded by an extracellular matrix (ECM) (Hunter and Beveridge, 2005; Werner *et al*., 2004), which typically shapes the bacterial network. MU adopts a biofilm-like structure *in vitro* and *in vivo*, displaying an abundant extracellular matrix (ECM) that harbours vesicles and serves as the reservoir of mycolactone (Marsollier *et al*., 2007). The matrix of *Staphylococcus* biofilms has been shown to interfere with macrophage phagocytic activity (Shiau and Wu, 1998) and to prevent antibodies from reaching the bacterial cell

surface (Ward *et al*., 1992). *Mycobacterium ulcerans* escaping from immune recognition could be due to the inaccessibility of the surface antigens to the host immune system.

Mycolactone, a polyketide-derived macrolide, is the major virulence determinant in *M. ulcerans* and appears to play a key role in the pathogenesis of Buruli ulcer. It enters cells by diffusion and accumulates in the cytosol (Snyder and Small, 2003) and has been shown to limit phagocytic activity (Adusumilli *et al*., 2005; Coutanceau *et al*., 2005). *In-vitro* experiments show that the major mycolactones, a mixture of *cis-trans* isomers (mycolactone A and B), produce apoptosis and necrosis in many human cell types (George *et al*., 2000; Dobos *et al.*, 2001). The resulting necrosis provides a favourable milieu for further proliferation of *M. ulcerans*. During the necrotic phase, there is very little or no cellular immune response. There is extensive tissue damage and immunosuppression suggested to be due to mycolactone in *in-vivo* studies involving a guinea pig model of infection (George *et al*., 1999).

The toxin may either be neutralized or the organism may cease to proliferate or to produce toxin. However, the mechanism by which they do that is unknown. Healing appears to begin when the host develops cell-mediated immunity. The inflammatory cells then destroy the etiological agent and the disease subsides with scarring. Bones may be affected by direct spread from the lesion or as a result of *M. ulcerans* bacteraemia.

2.52 Signs and symptoms of Buruli ulcer

In developing countries, socio-cultural beliefs and practices strongly influence the health-seeking behaviours of people affected by Buruli ulcer. In focus group discussions in Benin,

some people expressed the belief that the disease may be natural or induced by another person through sorcery (Aujoulat *et al*., 2003). Many patients and unaffected individuals in endemic areas attribute the disease to witchcraft and curses (Stienstra *et al*., 2002). The first recourse is often traditional treatment especially at the pre-ulcerative stage and later, the hospital when victims develop ulcerative lesions. Stigmatization, high cost of surgical treatment, fear of treatment effects, notably amputation, among other factors, cause delay in reporting to hospital (Stienstra *et al*., 2002). As a result, the most frequent lesion is an ulcer and with increasing patient delay, lesions tend to become larger with the classical undermined borders resulting in more severe functional limitations (Barogui *et al*., 2009).

The clinical forms of Buruli ulcer may be active or inactive. The inactive form is characterised by previous infection with depressed stellate (starshaped) scar with or without sequelae (complications). The active form is an ongoing infection and manifests as pre-ulcerative and ulcerative disease. The disease initially presents as painless nodule, papule, plaque, or oedema, which are non-ulcerative lesions. It often evolves into a painless ulcer with characteristically undermined edges with occasional involvement of the bones followed by debilitating complications such as contracture deformities if left untreated (George *et al*., 1999; WHO, 2000; Portaels and Meyers, 2006). Plates 2.1 to 2.5 (WHO BU photos) show some stages of the disease.

As defined by WHO (2004), a nodule is a lesion that extends from the skin into the subcutaneous tissue and is 1–2 cm in diameter. It is usually painless, but may be itchy, and the surrounding skin may be discoloured compared with adjacent areas. Nodules are commonly seen in Africa.

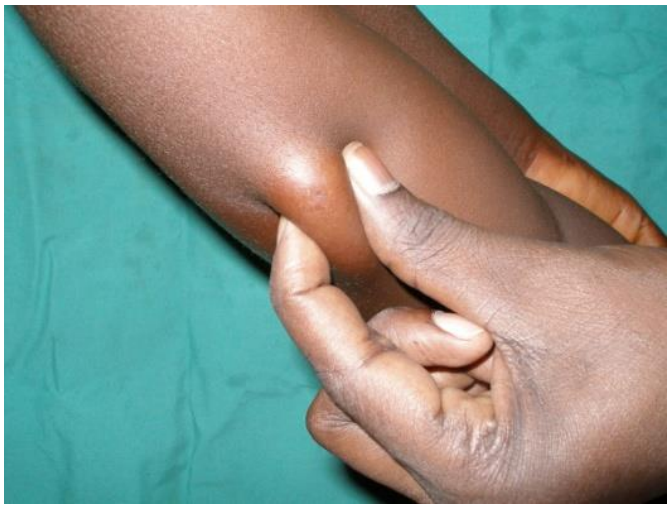


Plate 2.1: Typical BU nodule (Photo: K. Asiedu)

A papule is a painless, raised skin lesion, less than 1 cm in diameter. The surrounding skin is reddened and the papule is commonly seen in Australia.

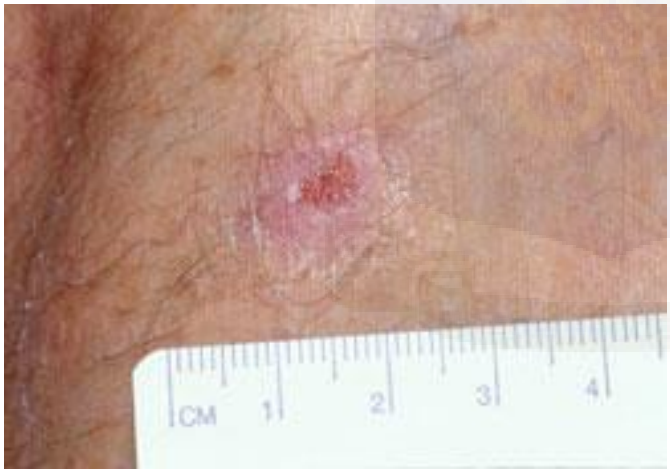


Plate 2.2: Typical BU papule (Photo: John Hayman)

Plague is a firm, painless, elevated and well-demarcated lesion more than 2 cm in diameter with irregular edges and often reddened or otherwise discoloured.

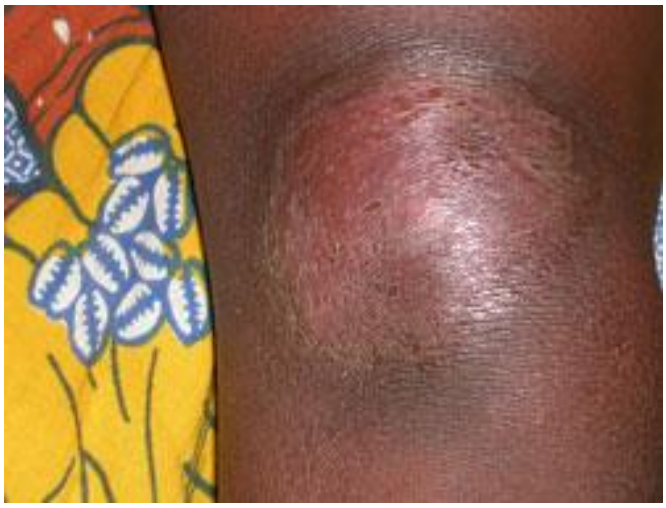


Plate 2.3: A large BU plague (Photo: A. Chauty, Benin)

The ulcerative lesion is usually painless, unless there is secondary bacterial infection, and characterised by necrotic centre, undermined edges and oedematous skin.

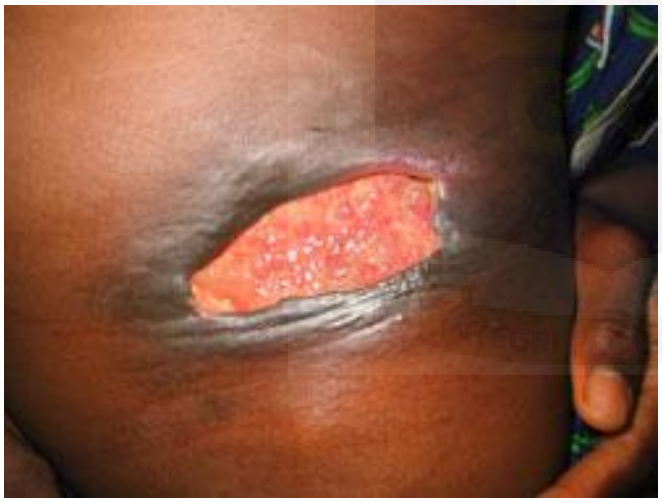


Plate 2.4: Small ulcer of BU (National BU control programme, Benin)

Contractures result from scarring caused by lesions over or close to joints after treatment.

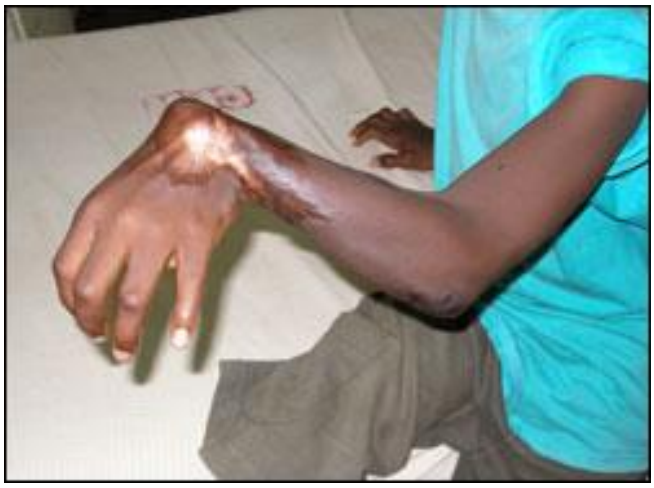


Plate 2.5: Contracture of the hand (Photo: K. Asiedu)

WHO BU Photos. Available at: *http://www.who.int/buruli/photos/en/index.html*

The disease progresses with no pain and fever probably due to the local immunosuppressive properties of mycolactone and may partly explain why those affected often do not seek prompt treatment. When patients delay in reporting to the hospital, lesions tend to become larger resulting in more severe functional limitations. Sometimes, bone is affected causing gross deformities. When lesions heal, scarring may cause restricted movement of limbs and other permanent disabilities in about a quarter of patients.

2.60 Diagnosis and treatment of Buruli ulcer

2.61 Clinical diagnosis

Clinical diagnosis is based on clinical findings by experienced health workers. In endemic areas, it is important to consider every nodule or ulcer as a suspected *M. ulcerans* infection until proven otherwise, as the disease is associated with nonspecific clinical manifestations and indolent course. A painless or minimally painful ulcer with undermined edge and a whitish-yellow necrotic base are characteristic clinical features for the diagnosis of BU. In

addition, residence or previous residence in an endemic area, persons under 15 years of age with about 85% of lesions on the limbs are important diagnostic clues.

2.62 Laboratory diagnosis

Although laboratory diagnosis of Buruli ulcer is usually based on microscopic examination, culture, IS*2404* PCR of swab samples and tissue specimens, and histopathological analysis, IS*2404* PCR is considered the most sensitive method for laboratory confirmation of the disease (Beissner *et al*., 2010). Swabs and/or tissue fragment from the necrotic base or undermined edge of ulcers where the bacteria are usually concentrated are used for the diagnosis of the disease.

2.621 Microscopy

Microscopic examination of smears from the necrotic base of ulcers stained by the Ziehl-Neelsen (ZN) method reveals clumps of extracellular acid-fast bacilli (AFB) in positive specimens (Plate 2.6). Biopsy specimens that include the necrotic base and the undermined edge of lesions with subcutaneous tissue are nearly always diagnostic. Although ZN microscopy is simple, less expensive to perform and very useful in resource-poor countries, it has low sensitivity and specificity. An assessment of the sensitivity of the above diagnostic tests in relation to clinical presentation of BU, type of diagnostic specimen and treatment history by Herbinger *et al*. (2009) revealed an overall sensitivity of 57% for microscopic examination. This sensitivity was comparable to that of culture (51%), but significantly lower than that of PCR (85%). Misdiagnosis has been observed with the use of microscopy, as among true Buruli ulcer cases, direct smear examination and culture are less frequently

positive when nodules are tested than when samples are collected from oedematous forms (WHO, 2014).

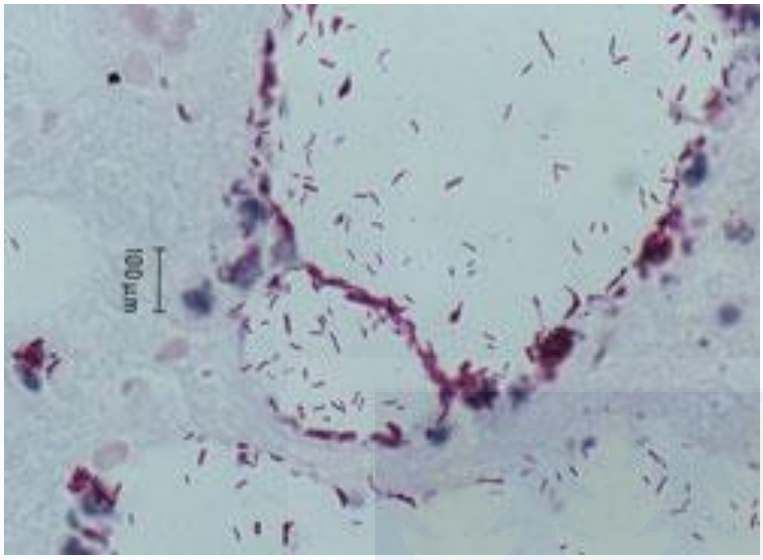


Plate 2.6: Ziehl-Neelsen stained smear from a Buruli ulcer showing red extracellular acid-fast bacilli (*Portaels, WHO photo library*).

2.622 Culture

The cultivation of *M. ulcerans* from clinical specimens can be carried out on Lowenstein-Jensen (L-J), Ogawa and Middlebrook media. However, for solid media, L-J is considered the most suitable medium for *M. ulcerans*. For the BACTEC system, Middlebrook 7H12B medium is recommended. *Mycobacterium ulcerans* grows under the same conditions as *M. tuberculosis* except that the optimal temperature is 29–33 ºC. In liquid media (e.g. in the BACTEC system) *M. ulcerans* may also show enhanced growth under reduced oxygen tension of 2.5–5% as observed by Palomino *et al*. (1998), suggesting that this organism has a preference for microaerobic environments. Optimal incubation temperature and oxygen concentration for the cultivation of MU are 33 ºC and 2.5% respectively. Positive primary cultures have colonies suggestive of *M. ulcerans* that appear yellowish, rough and with well-

demarcated edges. When BACTEC is positive, L-J is inoculated from the BACTEC medium using a single colony for the subculture. On L-J medium *M. ulcerans* forms yellowish-white colonies after 6 - 8 weeks incubation (Plate 2.7).

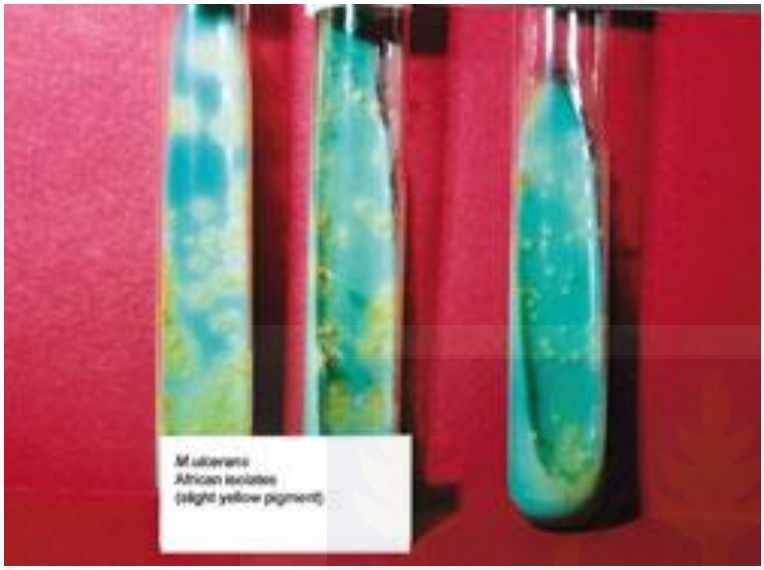


Plate 2.7: African *M. ulcerans* isolates cultivated on L-J medium showing a yellowish-white pigment (*Portaels, WHO photo library*).

2.623 Polymerase chain reaction (PCR) diagnosis of Buruli ulcer

The diagnosis of BU by PCR is based on the amplification of the insertion sequence IS*2404* present in the genome of *M. ulcerans* (Stinear *et al.*, 1999; 2007) using appropriate primers. Of the several published methods available, IS*2404* PCR is the best method for BU diagnosis and requires suitable clinical specimens such as dry swabs, fresh tissue or specimens kept in transport medium. With an analytical specificity of 100% (Ross *et al.*, 1997; Rondini *et al.*, 2003; Siegmund *et al.*, 2005), a sensitivity of 79%–85% (Guimaraes-Peres *et al.*, 1999; Stienstra *et al.*, 2003) and positivity ratios of 61%–72% (Siegmund et al., 2007; Mensah-Quainoo *et al.*, 2008), IS*2404* PCR is considered to be the most reliable technique for the detection of *M. ulcerans* in human diagnostic samples.

Although the disease can be definitively diagnosed within 24 hours of receipt of a clinical specimen by the laboratory, PCR detection is costly, requires careful handling of reaction components by skilled laboratory workers as well as strict measures to avoid contamination. Unfortunately, PCR detection of MU is rarely available in resource-poor and BU endemic countries where there are inadequate laboratory facilities and trained personnel and financial constraints.

2.624 Histopathology

This diagnostic method requires tissue biopsies. It depends on detection of characteristic histopathological changes in tissue samples, from the various stages such as non-ulcerated, ulcerated and healing lesions, and the presence of ghost cells, extracellular bacilli and vasculitis (Plate 2.8). Its sensitivity is about 90% (Guarner *et al*., 2003; Herbinger *et al*., 2009; Siegmund, *et al*., 2007; Siegmund, *et al*., 2005) and is useful also for differential diagnoses when the results of microscopy, culture and PCR are negative.

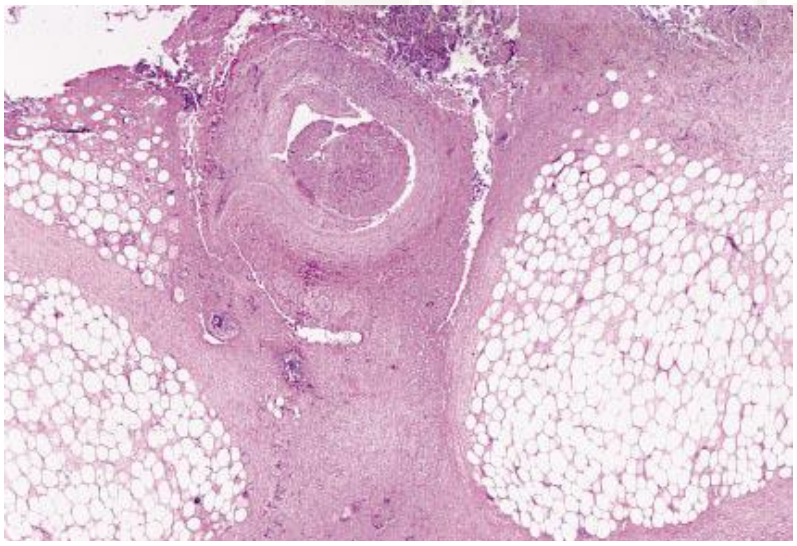


Plate 2.8: Ghost cells and vasculitis. (Photo: Wayne Meyers)

2.63 Treatment of Buruli ulcer

The standard treatment of BU is surgical and consists of two stages. In the first stage, there is excision of all dead tissue including a healthy tissue margin around the lesion, typically 3 - 4 cm (van der Werf *et al*., 1999). Generally, this procedure results in a large surgical lesion that needs to be skin grafted in the second stage (Aguiar and Stenou, 1997). The whole procedure requires long hospitalization with reported median duration of hospital admission of around three months (Asiedu and Etuaful, 1998; Ouattara *et al*., 2004) or less (Barogui *et al*., 2009). Although early detection and excision can prevent development of the large, disfiguring ulcers often associated with persistent deformity after healing (Evans *et al*., 2003), many patients present late (Aguiar *et al*., 1997), with extensive ulcers that require wide surgical excision followed by skin grafting, which is costly (Asiedu and Etuaful, 1998).

In addition, antimycobacterial treatment alone or in combination with surgery was introduced in 2004 (WHO, 2004; Buntine and Crofts, 2001) and have been shown to be useful in the treatment of BU. Rifampin, amikacin, and streptomycin appeared to be bactericidal (Dega *et al*., 2002) in a mouse model study involving the treatment of mouse footpad lesions. The size of mouse footpad lesions treated with rifampin and amikacin together for 12 weeks decreased progressively, with reduced mean colony-forming unit (CFU) counts of *M. ulcerans* and no relapses (Dega *et al*., 2002; Pattyn, 1965).

Moreover, Etuaful *et al*. (2005) studied lesion size and treatment of humans with early nonulcerative *M. ulcerans* disease and observed that treatment with rifampin and streptomycin for 4 weeks or more inhibited growth of *M. ulcerans* in human tissue. In this study no lesions became enlarged during antibiotic treatment and most of them became

smaller. It is hypothesized that with antibiotics, less extensive excision is required to obtain cure. With antibiotics alone, small lesions (diameter < 5 cm) successfully heal in about 81% of the cases (Chauty *et al*., 2007). Antibiotic treatment for BU, irrespective of the stage, requires a combination of daily injections with streptomycin (15 mg/kg body weight) and daily oral rifampicin (10 mg/kg body weight) for eight weeks (WHO, 2015). However, antibiotic treatment with streptomycin is not an option for pregnant women, who are treated by surgery (Barogui *et al*., 2009). A combination of rifampicin (10 mg/kg once daily) and clarithromycin (7.5 mg/kg twice daily) is also considered the safer option for pregnant patients (WHO, 2015).

2.70 Genotyping methods and genetic diversity of *Mycobacterium ulcerans*

2.71 Non-VNTR-based typing methods

The genetic diversity in mycobacteria is driven by the activity of mobile DNA such as insertion sequence (IS) rather than nucleotide sequence drift (Sreevatsan *et al*., 1997) making IS elements potentially ideal targets for indexing rapidly evolving change in mycobacterial populations (Picardeau and Vincent, 1996; Yoder *et al*., 1999). *M. ulcerans* has at least two high-copy-number IS elements, IS*2404* and IS*2606*, that demonstrate RFLP between strains (Stinear *et al*., 1999). Based on PCR amplification of the region between adjacent copies of IS*2404* and IS*2606* (2426 PCR), Stinear and colleagues (2000) designed a simple genotyping method to investigate its discriminatory capability in *M. ulcerans.* They observed nine distinct genotypic profiles that correlated with the geographic source of the isolates based on the banding pattern, indicating ongoing transposition events or genome rearrangement associated with insertion sequences. The close genetic relationship among the five genotypes from Southeast Asian isolates suggests occasional genome rearrangement events to produce a

unique genotype in a particular region. They further observed the African genotype to be closely related to those of Southeast Asia, suggesting a close evolutionary link between *M. ulcerans* isolates from Africa and Southeast Asia.

Moreover, restriction fragment length polymorphism (RFLP) based on the IS*2404* probe has proven to be useful in differentiating *M. ulcerans* strains from different geographic areas, but not most strains from the same continent. The IS*2404* probe (Chemlal *et al*., 2001) resulted in six RFLP profiles and six geographic types based on the distribution of IS*2404* in the genome of the strains isolated in those regions, confirming that *M. ulcerans* can be divided into subgroups corresponding to different geographic variants of the same species. The isolates from Africa, Australia and Southeast Asia were observed to contain more IS*2404* copies in their genomes than the strains from Asia, South America and Mexico. The authors thought that the differential genomic distribution of IS*2404* in *M. ulcerans* may be explained by the ability of this repeat element to actively jump or duplicate in the genome. A comparison of the banding patterns further discovered only a few differences in the corresponding subtypes (I, II, and IV) with a close genotypic relationship consistent with the close evolutionary link of those strains suggested by Stinear and colleagues (2000).

The plasmid pTBN12, which hybridizes to areas of repeated sequence in many mycobacterial species, has been successfully used as a probe for typing and differentiating strains of *M. tuberculosis*, *M. bovis* and *M. kansasii* (Cousins *et al*., 1993; Ross *et al*., 1992a, b). Jackson and colleagues (1995) applied this technique to genetically distinguishing *M. ulcerans* strains from at least two different states in Australia and from two countries, Benin and Democratic Republic of Congo, in Africa based on the differences in RFLPs types obtained, supporting the discriminatory ability of the pTBN12 probe to differentiate strains of *M. ulcerans*.

However, its discriminatory ability is countered by the length of time (three to four days) taken to perform a Southern hybridization compared with the four to six hours taken by 2426 PCR and the ability of 2426 PCR to obtain a genotype directly from a clinical specimen (Stinear *et al*., 2000).

Previous genotypic study involving nucleotide sequence analysis of the 16S rRNA genes could only differentiate *M. ulcerans* strains at the continental level. In that study, Portaels and colleagues (1996) analysed the 3´-terminal region of the 16S rRNA gene sequence of *M. ulcerans* strains from Africa, the Americas and Australia, and observed three subgroups corresponding to the continent of origin. The observed invariable correlation between the geographic origins of the strains and polymorphisms in the 16S rRNA gene is generally in agreement with the proposed classification of *M. ulcerans* strains based on restriction fragment length polymorphism analysis (Jackson *et al*., 1995; Chemlal *et al*., 2001).

2.72 VNTR-based typing methods

Tandem repeats occur in DNA when a pattern of two or more nucleotides is repeated and the repetitions are directly adjacent to each other (NLM, 2009). Tandem repeat (TR) DNA sequences are important sources of polymorphism in the genome of many eukaryotes and prokaryotes. Genomic regions showing polymorphism due to different numbers of TR motifs in different strains or individuals are described as variable number tandem repeat (VNTR) loci. VNTR is a type of minisatellite whose repeat unit ranges from 9 bp to 80 bp. VNTR typing is, therefore, a PCR-based technique for identifying alleles of defined regions of DNA that contain variable numbers of copies of short sequence stretches.

VNTR genotyping has been applied to highly monomorphic and clonal species such as *Mycobacterium tuberculosis* (Frothingham & Meeker-O'Connell, 1998; Mazars *et al*., 2001; Roring *et al*., 2002; Skuce *et al*., 2002; Smittipat & Pallittapongarnpim, 2000; Supply *et al*., 2000), *Bacillus anthracis* (Keim *et al*., 2000) and *Yersinia pestis* (Klevytska *et al*., 2001) and has proven to be very useful in differentiating and studying genetic relationships of strains of these species. Ablordey and colleagues (2005), therefore, investigated the existence in the genome of *M. ulcerans* of VNTR loci and assessed their importance in differentiating strains of *M. ulcerans*. The nine VNTR loci identified in *M. ulcerans* revealed eight genotypes among the 23 *M. ulcerans* reference isolates tested and highlighted the extremely high clonal homogeneity within certain geographic regions.

In addition, VNTR analysis discriminated one Papua New Guinean genotype by a single-locus difference to the South East Asian genotype, emphasizing the very close evolutionary link between strains from Southeast Asia, Australia and Africa observed previously (Chemlal *et al*., 2001; Stinear *et al*., 2000). It further discriminated between the isolates from French Guyana and Surinam by two loci, contrary to RFLP results based on the IS*2404* probe (Chemlal *et al*., 2001) and amplified fragment length polymorphism (AFLP) results (Chemlal *et al*., 2001 b), identifying for the first time two closely related, but different *M. ulcerans* genotypes in South America. Clearly, the resolution power of this method is slightly higher compared with those of standard molecular typing methods with resolution mostly only at continental level and support the potential of a VNTR-based genotyping method for *M. ulcerans* strain discrimination within geographical regions.

Although allele-length polymorphism associated with VNTR loci allow for some inter- and intraspecies differentiation in *Mycobacterium marinum* and *M. ulcerans,* there is limited

intraspecies discrimination among *M. ulcerans* isolates within the same geographic region (Ablordey, *et al*. 2005b). Allele-length analysis of VNTR loci has also not been helpful in differentiating isolates from Africa, Southeast Asia or Asia neither has other typing methods such as mycobacterial interspersed repetitive unit-VNTR typing (Stragier *et al*., 2005), multilocus sequence typing (Stinear *et al*., 2000) and IS*2404*-restriction fragment length polymorphism typing (Chemlal *et al*., 2001). However, some gains in strain discrimination have been indicated when length polymorphism data complement sequence analysis (Amonsin *et al*., 2004; Frothingham, 1995).

Comparative nucleotide sequence analysis of polymorphic VNTR loci by Ablordey *et al*. (2005a) made important contribution to allelic diversity in some *M. ulcerans* genotypes. Although a general trend of tandem repeat (TR) sequence conservation in isolates from the same geographic region was noticed, sequence analysis revealed three Southeast Asian alleles and two alleles within the Asian genotype, supporting data from IS*2404*-Mtb2 PCR (Ablordey *et al*, 2005) and 2426 PCR (Stinear *et al*., 2000). The sequence variants in two loci permitted intraspecies resolution of Southeast Asian and Asian genotypes. However, intraspecies differentiation within the African genotype was not possible, as the isolates displayed complete sequence homology across all the VNTR loci. This lack of sequence variants further emphasizes the clonal homogeneity and recent evolutionary origin and distribution of the African genotype (Stinear *et al*., 2000).

The previously identified tandem repeat loci, VNTR (Ablordey *et al*., 2005b) and mycobacterial interspersed repetitive unit, MIRU (Stragier *et al.*, 2005) in the genome of *M. ulcerans* have made important contribution to the differentiation of *M. ulcerans* isolates. These typing methods are valuable for distinguishing *M. ulcerans* isolates from different

geographic regions and for distinguishing between *M. ulcerans* and the other MPM (Ablordey *et al*., 2005a, b, 2007; Stragier *et al*., 2007).

Stragier and colleagues (2006) reported the first evidence of genetic diversity in *M. ulcerans* samples from three African countries: Democratic Republic of Congo (DRC), Sudan and Uganda. In this study, a selection of MIRUs and VNTRs were used to analyze *M. ulcerans* extracts from tissue specimens and the results compared with those of a geographically diverse collection. *M. ulcerans* isolates from DRC and Uganda, and tissue extracts from patients from Sudan and Uganda showed distinct profiles different from the originally homogeneous African genotype. The results demonstrate heterogeneity among *M. ulcerans* from different African countries, categorising the three African profiles based on their origin as follows; Atlantic African profile, Central African Congo River basin profile and the East African Nile River basin profile.

The analysis of 72 African isolates of *M. ulcerans*, including 57 strains from Ghana by Hilty and colleagues (2006), demonstrated, for the first time, genetic diversity in *M. ulcerans* in an African country. In this study, the use of VNTR typing based on a newly identified polymorphic locus designated ST1 and MIRU 1 locus, described previously, revealed three VNTR allele combinations in the clinical isolates from Ghana. The results indicate the emergence and spread within Ghana of new genetic variants of *M. ulcerans*, thus supporting the potential of VNTR-based typing for genotyping *M. ulcerans*.

2.730 SNP typing and DNA sequencing technique application

2.731 Single nucleotide polymorphism typing and sequencing of *M. ulcerans*

Single nucleotide polymorphism (SNP) is a DNA sequence variation occurring when a single nucleotide in a genome or other shared sequence differs between members of a species or paired chromosome in an individual. SNP typing methods have been described as the most promising genotyping methods for genetic diversity in genetically monomorphic bacteria (Achtman, 2008). Several studies in *M. ulcerans* using SNP typing techniques have provided more detailed genomic information compared with conventional genetic fingerprinting methods. These studies provide strong basis for further investigation into MU distribution and transmission particularly at regional and local levels using SNP typing methods.

The ISE-SNP typing by Kaser and colleagues (2009b) achieved higher resolution compared with other DNA fingerprinting techniques. The identified ISE-SNP types correlated with the more limited mycobacterial interspersed repetitive unit-VNTR fine-typing and, in particular, enhanced the resolution within the Atlantic African genotype (Hilty *et al*., 2006; Stragier *et al*., 2006). It was observed that some ISE-SNP types of MU isolates in the classical lineage investigated appeared to be widespread across West African countries including Ghana, Ivory Coast, Togo and the Democratic Republic of Congo. However, others appeared more delimited as was the case in isolates from Benin and some isolates from the Democratic Republic of Congo. Among *M. ulcerans* isolates from Ghana, four ISE-SNP types were identified. These observations were based on only two copies of IS*2404* specific to the African/Australian haplotypes, suggesting that analysis of a larger number of ISE copies or of the entire genome of a collection of isolates has the potential of yielding larger number of SNPs to resolve the spatial and temporal dispersal of genetic *M. ulcerans* variants on the regional level.

The comparison of genome sequences involving two Ghanaian MU strains and a Japanese strain with the reference genome of the Ghanaian classical lineage isolate Agy99 by Qi and colleagues (2009) revealed 26,564 SNPs in the Japanese strain representing the ancestral lineage and 173 SNPs in the Ghanaian strains. Further analysis of MU strains from Ghana using 68 SNP loci based on the discovered SNPs differentiated 54 strains into 13 distinct SNP haplotypes. The identified SNPs among Ghanaian strains for the first time resolved the population structure and evolutionary relationship of an intra-continental population of *M. ulcerans*. The SNP typing of Ghanaian isolates was consistent with VNTR typing (Hilty *et al*., 2006). In addition, it allowed for further differentiation of MU isolates coming from the same BU endemic focus into clonal complexes, an indication for ongoing microevolution.

Moreover, Roltgen and colleagues (2010) compared SNP typing results of MU clinical isolates from two geographically separate BU endemic areas in Ghana and observed a total of 61 different alleles at 65 SNP loci. In each of the endemic areas, SNP variation was observed to be significantly smaller than the overall variation between the two endemic areas, indicating the dominance of two different clonal complexes in two separate Ghanaian BU endemic areas. The dominance of one clonal complex with local clustering of some variants revealed focal transmission of MU, demonstrating the potential of SNP typing throwing light on MU transmission. With the exception of the strain from Togo and one strain from Ivory Coast that had similar haplotypes to those of some Ghanaian strains, all other isolates from the other West African countries shared distinct SNP pattern when compared with the Ghanaian isolates.

2.732 DNA sequencing applications

Sequencing technology has wide application, spanning diverse research sectors including comparative genomics and evolution, forensics, epidemiology, and applied medicine for diagnostics and therapeutics. The identification of mycobacteria in the past few decades relied primarily on phenotypic, chemotaxonomic and serotaxonomic tests performed on cultures. Identification of mycobacterial isolates based on phenotypic characters was a time-consuming procedure, with many different features requiring complex and specialized tests necessary to identify mycobacteria at the species level (Wayne and Kubica, 1986). Although attempts to subdivide mycobacterial species on the basis of immunological methods, DNA homology and related analyses (McFadden *et al.*, 1987; Imaeda *et al.*, 1988; Wayne and Kubica, 1986) proved useful in determining taxonomic relationships, reliable identification of mycobacteria remained problematic. The development of sequencing strategies has made considerable impact in differentiating *Mycobacterium* spp.

Several gene targets including 16S rRNA gene (Rogall *et al*., 1990; El Amin *et al*., 2000), *hsp65* gene (Ringuet *et al*., 1999) and *recA* gene (Blackwood *et al*., 2000) have been reported to be useful for sequencing to distinguish *Mycobacterium* species. A polymerase chain reaction-based sequencing strategy by Rogall and colleagues (1990) demonstrated the importance of 16s rRNA sequence for rapid identification of mycobacterial isolates. Based on comparative 16s rRNA sequencing within a wide range of *Mycobacterium* species, they observed, among others, that *Mycobacterium gastri/M. kansasii, M. scrofulaceum* and *M. simiae* have different residues at specific positions, thus allowing them to be distinguished. This procedure further differentiated *Mycobacterium* spp. such as *M. malmoense, M. szulgai* and *M. flavescens*, which are difficult to identify by classical methods, but revealed 100% homology in the 16s rRNA structure for *M . tuberculosis, M. bovis* and *M. africanum*.

The *recA* gene, a highly conserved gene in bacteria (Brendel *et al*., 1997), was the first gene discovered to mediate homologous recombination (Clark and Margulies, 1965). In 2000, Blackwood and colleagues proposed sequencing of this gene to offer a complementary method to 16S rRNA gene sequencing for the accurate identification of the *Mycobacterium* species. In this study, the *recA* genes for 30 *Mycobacterium* species were amplified by PCR, sequenced and compared with the published *recA* sequences of *M. tuberculosis, M. smegmatis* and *M. leprae* available from GenBank. Using the first 915 bp of all species for sequence alignment and analysis of the *recA* genes, the authors discovered the presence of a large number of nucleotide substitutions among the species tested, with interspecies similarities ranging from 75.7% between *M. leprae* and *M. aurum* or *M. mucogenicum* to 96.2% between *M. gastri* and *M. kansasii*. Unlike the 16S rRNA gene, sequence similarities in the *recA* genes were significantly lower among *Mycobacterium* species, with variability occurring throughout the *recA* gene and demonstrating intraspecies variability ranging from 98.7 to 99.9% in 13 of the species (Blackwood *et al*., 2000). Exceptions to this were members of the *M. tuberculosis* complex, which were identical, in agreement with previous findings (Rogall *et al*., 1990).

Moreover, the discrimination of members of *M. avium* complex (Turenne *et al*., 2006) was based on the sequencing of the 3' region of the *hsp65* gene, which identified 14 different *hsp65* sequevars that were resolved into eight variants by examining the 441-bp region of the gene alone. This provided adequate variability in distinguishing among *M. avium, M. intracellulare*, and *M. chimaera*. However, within a species, *M. avium* in particular, the 3' region of *hsp65* provided much greater discrimination, differentiating subspecies and even *M. avium* subsp. *paratuberculosis* host-associated types.

2.733 DNA sequencing application in *Mycobacterium ulcerans*

Sequencing technology also has useful application in studying genomic diversity in *M. ulcerans*. Whole genome sequence comparisons of *M. ulcerans* isolate from Ghana (Agy99) with *M. marinum* M strain showed that Agy99 evolved from *M. marinum* M strain by a process of lateral gene transfer and reductive evolution (Stinear *et al*., 2000; Yip *et al*., 2007) with the acquisition of the virulence plasmid, pMUM001, being characteristic for *M. ulcerans* (Stinear *et al*., 2005a, b). *M. ulcerans* Agy99 genome had many examples of DNA deletions referred to as *M. ulcerans* regions of difference (MURDs) (Stinear *et al*., 2007), accounting for the loss of 1000 kb of DNA between *M. marinum* and *M. ulcerans*. The development of a plasmid-based DNA microarray based on the Ghanaian reference strain Agy99 by Rondini *et al*. (2007) facilitated the detection of large sequence polymorphisms and identified twelve regions of difference (RDs among 30 *M. ulcerans* strains of diverse geographic origin.

Using a combination of PCR, cloning, sequencing and primer walking, Kaser and colleagues (2007) carried out detailed RD sequence comparison that confirmed five haplotypes and revealed the existence of two major *M. ulcerans* lineages. By sequencing representatives of the five InDel haplotypes and aligning sequences to *M. marinum* M genome, they determined deletion sizes and their breakpoints, and identified sequence insertions, substitutions, dislocations, inversions and rearrangements and observed consistently that members of a given subgroup yielded identical results in all RDs analysed, confirming the occurrence of five haplotypes. More frequent deletions in the absences of substituting DNA such as insertion sequence element (ISE) were observed in genomes of the South American, Mexican and Asian haplotypes and were larger than in the African/Australian cluster. However, ISEs (IS*2404*, IS*2606* and IS*2404*/IS*2606* tandems) were frequently found in the African/Australian haplotypes, but not in the South American, Mexican and Asian

haplotypes. Moreover, a multitude of genomic rearrangements such as large DNA fragment dislocation, DNA fragment inversions and rearrangements were observed in the African/Australian cluster and not in any of the twelve RDs for the South American, Mexican and Asian haplotypes. Kaser and colleagues (2007), thus, concluded that *M. ulcerans* has evolved into five InDel haplotypes that separate into two distinct lineages: (i) the "classical" lineage including the most pathogenic genotypes – those that come from Africa, Australia and South East Asia, and (ii) an "ancestral" *M. ulcerans* lineage comprising strains from Asia (China/Japan), South America and Mexico. The ancestral lineage designation was based on its members being genetically closer to the progenitor *M. marinum* in both RD composition and DNA sequence identity. However, the classical lineage was observed to have undergone major genomic rearrangements and its designation was based on the inclusion of the sequenced African strain (Agy99) and most of the existing *M. ulcerans* clinical isolates.

Morevover, comparative genomics in *M. ulcerans* involving haplotypes of the classical and ancestral lineages as well as the clinical isolate *M. marinum* strain M by Kaser and Pluschke (2008) using direct sequencing and Real-time PCR techniques identified several mechanisms that explain the genomic differences in these groups of haplotypes. Analysis of 7% of the entire *M. ulcerans* genome associated with RDs1 to 15 revealed frameshift mutations resulting in pseudogenization, interspersing of ISEs into coding sequences (CDS) that led to their disruption and physical deletions of sizes between 2 and 53 kbp with replacement by ISEs. Both pseudogenization or functional disruption and physical deletion of the CDSs lead to gene silencing. The observed differential gene inactivation patterns associated with the *M. ulcerans* haplotypes identified candidate genes that may confer enhanced adaptation upon ablation of expression. They concluded that a number of gene conversions confined to the classical lineage may contribute to particular virulence of this group comprising isolates from

Africa and Australia.

As observed previously (Stinear *et al*., 2004), the acquisition of virulence plasmid followed by the expression of macrolide toxin, mycolactone, was an important step in the development of the ancestor of *M. ulcerans* to a mammalian pathogen. However, other mycolactone-producing mycobacteria closely related to *M. marinum* and *M. ulcerans* have been isolated from lesions in frogs and fish (Ranger *et al*., 2006; Ucko *et al*., 2002), but not from infected humans, an indication that additional factors contribute to the high virulence of the classical lineage of *M. ulcerans*. This study further indicates that in addition to the acquisition of virulence plasmid, loss of distinct anti-virulence genes, partly driven by ISE – in particular IS*2606* – expansion, might have equipped the classical lineage with a particular virulence and transmissibility. Identification of this spectrum of anti-virulence gene candidates expands our understanding of the pathogenicity of *M. ulcerans*.

Genetic discrimination of *M. ulcerans* clinical isolates at the local level (Roltgen *et al*., 2010) using sequencing techniques has also been demonstrated in a comparative study involving MU clinical isolates from two geographically separate BU endemic areas of Ghana. The validation of SNP loci using Sanger DNA sequencing revealed five different haplotypes, other than Agy99, within the Ga District endemic area strains and differences between Amansie West District strains and isolates from other African countries. A phylogeographic analysis of haplotype distribution by the authors, facilitated by locating the homes of patients from whom the strains were isolated, revealed a clustering of unique *M. ulcerans* haplotypes within the Densu river basin in the Ga District of Ghana, forming clusters of independent focal transmission. These studies clearly demonstrate that comparative genomics with DNA sequencing technology has greatly improved our understanding of the evolution and diversity

of pathogenic mycobacteria.

2.80 Characteristics of sequencing methods

2.81 Basic sequencing methods

The Maxam-Gilbert DNA sequencing method published in 1977 was based on chemical modification of DNA and subsequent cleavage at specific bases (Maxam & Gilbert, 1977). This method requires radioactive labeling at one 5' end of the DNA and purification of the DNA fragment to be sequenced. The use of radioactive labeling and the technical complexity of Maxam-Gilbert sequencing method discouraged its extensive use after refinements in the Sanger methods had been made. The chain-termination method developed by Frederick Sanger and coworkers in 1977 soon became the method of choice, owing to its relative ease and reliability (Sanger and Coulson, 1975; Sanger *et al*., 1977). It uses fewer toxic chemicals and lower amounts of radioactivity than the Maxam and Gilbert method. Because of its comparative ease, the Sanger method was soon automated and was the method used in the first generation of DNA sequencers.

Since the early 1990s, DNA sequence production has almost exclusively been carried out with capillary-based, semi-automated implementations of the Sanger biochemistry (Sanger *et al*., 1977 b; Swerdlow *et al*., 1990; Hunkapiller *et al*., 1991). The Sanger sequencing method is a mixed-mode process involving the synthesis of a complementary DNA template using natural 2'-deoxynucleotides (dNTPs) and termination of synthesis using 2',3'-dideoxynucleotides (ddNTPs) by DNA polymerase (Sanger *et al*., 1977). In high-throughput production pipelines, DNA to be sequenced is prepared by one of two approaches. First, for shotgun *de novo* sequencing, genomic DNA is fragmented and cloned into a high-copy-number plasmid, which is then used to transform *Escherichia coli*. Secondly, for targeted

resequencing, PCR amplification is carried out with primers that flank the target. The output of both approaches is an amplified template. This may be many clonal copies of a single plasmid insert present within a spatially isolated bacterial colony that can be picked and DNA isolated for each sequencing reaction, or many PCR amplicons present within a single reaction volume.

The sequencing reaction takes place in cycles of template denaturation, primer annealing and extension with the primer being complementary to known sequence immediately flanking the region of interest. Each round of primer extension is randomly terminated by the incorporation of fluorescently labelled dideoxynucleotides (ddNTPs). These processes result in the generation of a set of nested DNA fragments, which differ in nucleoside monophosphate units. The DNA fragments randomly terminated by fluorescently-labeled dideoxyribonucleotides in the polymerase reaction are then separated by their size using high-resolution gel electrophoresis and analyzed to reveal the DNA sequence.

Since about 2001, automated DNA sequencers have used capillary electrophoresis with laser-induced fluorescence (LIF) detection, the standard detection method in DNA sequencing by capillary electrophoresis (Dolník, 1999). Fluorescence detection, the most widely used detection method for four-colour DNA sequencing (Smith *et al*., 1986; Prober *et al*., 1987), is based on resolution of the emission signal from a dye-labeled nucleotide into colour, with subsequent assignment in the DNA sequence. Laser excitation of fluorescent labels as fragments of discreet lengths exit the capillary, coupled to four-color detection of emission spectra, provides the readout that is represented in a Sanger sequencing 'trace'. Software translates these traces into DNA sequence, while also generating error probabilities for each base-call (Ewing & Green, 1998; Ewing *et al*. 1998). The approach taken for subsequent

analysis, for example genome assembly or variant identification, depends on what is being sequenced and the reason for the sequence. After three decades of gradual improvement, the Sanger biochemistry can be applied to achieve long read-lengths of up to about 1,000 bp, and per-base 'raw' accuracies as high as 99.99% (Shendure & Ji, 2008). However, it is more expensive and impractical for larger sequencing projects compared with next-generation sequencing technologies.

2.82 Next-generation sequencing methods

Next-generation sequencing methods also known as second-generation sequencing methods are high-throughput sequencing technologies that parallelize the sequencing process, producing thousands or millions of sequences at once (Hall, 2007; Church, 2006). These include pyrosequencing or 454 sequencing (Roche Applied Science; Basel), sequencing by synthesis or Solexa technology (used in the Illumina [San Diego] Genome Analyzer), sequencing by ligation or the SOLiD platform (Applied Biosystems; Foster City, USA), the Polonator (Dover/Harvard) and the HeliScope Single Molecule Sequencer technology (Helicos; Cambridge, USA). These are cyclic-array sequencing platforms that sequence a dense array of DNA features by iterative cycles of enzymatic manipulation and imaging-based data collection (Mitra and Church, 1999; Mitra *et al*., 2003). The multiple, fragmented sequence reads produced must be assembled together on the basis of their overlapping areas. These platforms have conceptually similar work flows. The shared attributes include the following:

- random fragmentation of DNA, followed by *in vitro* ligation of common adaptor sequences to accomplish library preparation,

- library amplification on a solid surface, which may be bead or glass,

- direct step-by-step detection of each nucleotide base incorporated during the sequencing reaction,

- massively parallel sequencing and shorter read lengths than capillary sequencers.

All next-generation platforms also offer paired end read capability where sequences can be derived from both ends of the library fragments in separate reactions. However, they are quite diverse in how the array of DNA is generated as well as in sequencing biochemistry. The generation of clonally clustered amplicons to serve as sequencing features can be achieved by several approaches, including *in situ* polonies (Mitra and Church, 1999), emulsion PCR (Dressman *et al*., 2003) for the 454, the Polonator and SOLiD platforms or bridge PCR (Adessi *et al*., 2000; Fedurco *et al*., 2006) for the Solexa technology.

Pyrosequencing is a nonfluorescence technique that measures the release of inorganic pyrophosphate (PPi), which is proportionally converted into visible light by a series of enzymatic reactions (Ronaghi *et al*. 1996, 1998). The 454 sequencing is performed by the pyrosequencing method based on sequencing by synthesis and using the release of pyrophosphate whenever a complementary nucleotide is incorporated in an open 3´ DNA strand during the sequencing reaction. The released PPi is used in a reaction catalysed by sulfurylase and resulting in the release of adenosine triphosphate (ATP). The ATP is used by luciferase to convert luciferin to oxyluciferin resulting in the emission of light. The emitted light is captured by charge-coupling device (CCD) camera and recorded as a series of peaks called a pyrogram, which corresponds to the order of complementary dNTPs incorporated and reveals the underlying DNA sequence. Nucleotides not incorporated in the DNA strand are degraded by apyrase to prevent background noise. Depending on the type of method used, the sequencing biochemistry may be different.

Second-generation or cyclic-array sequencing strategies have several advantages relative to Sanger sequencing. The *in vitro* construction of sequencing library and *in vitro* clonal amplification to generate sequencing features circumvents several bottlenecks like transformation of *E. coli* and colony picking that restrict the parallelism of conventional sequencing. This enables much higher degree of parallelism resulting in the production of thousands or millions of sequences at once (Hall, 2007). Next-generation sequencing technologies produce shorter reads, anywhere from 25–500bp, but many hundreds of thousands or millions of reads in a relatively short time, on the order of a day (Karl *et al*., 2009). This results in high coverage, although the assembly process is much more computationally expensive. They are, therefore, vastly superior to Sanger sequencing due to the high volume of data and the relatively short time it takes to sequence a whole genome. Moreover, as observed previously (Quail *et al*., 2012; Liu *et al*., 2012), the cost per 1 million bases using next-generation sequencing methods is much cheaper compared with Sanger sequencing. The major disadvantage is that the accuracies are usually lower, although this is compensated for by the high coverage (Metzker, 2010).

There are important differences among next-generation sequencing platforms that may result in advantages with respect to specific applications such as the Illumina (Solexa) and the recently developed Ion Torrent Sequencing Technologies.

2.821 Illumina sequencing technology

The Illumina (Solexa) sequencing technology is based on the reversible terminator chemistry (Bentley *et al*., 2008). Single-stranded, adapter-ligated fragments are bound to the surface of the flow cell exposed to reagents for polyermase-based extension. In the generation of amplified sequencing features by bridge PCR (Adessi *et al.*, 2000; Fedurco *et al*., 2006), both

forward and reverse PCR primers are tethered to a solid substrate by a flexible linker, such that all amplicons arising from any single template molecule during the amplification remain immobilized and clustered to a single physical location on an array. Each of the resulting 'clusters' consists of ~1,000 clonal amplicons. On the Illumina platform, the bridge PCR relies on alternating cycles of extension with *Bst* polymerase and denaturation with formamide. Repeated denaturation and extension result in localized amplification of single molecules in millions of unique locations across the flow cell surface. This process occurs in what is referred to as Illumina's "cluster station", an automated flow cell processor.

Following cluster generation, the amplicons are single-stranded (linearization) and a sequencing primer is hybridized to a universal sequence flanking the region of interest. Each cycle of sequence interrogation consists of single-base extension with a modified DNA polymerase and a mixture of four modified nucleotides (Shendure & Ji, 2008). These nucleotides are 'reversible terminators', in that a chemically cleavable moiety at the 3′ hydroxyl position allows only a single-base incorporation to occur in each cycle; and one of four fluorescent labels, also chemically cleavable, corresponds to the identity of each nucleotide (Turcatti *et al*., 2008). After single-base extension and acquisition of images in four channels, chemical cleavage of both groups sets up for the next cycle. Read-lengths up to 36 bp are currently routine; longer reads are possible but may incur a higher error rate. Although read-lengths are limited by multiple factors that cause signal decay and dephasing, such as incomplete cleavage of fluorescent labels or terminating moieties, the large number of reads generated results in high coverage. Additionally, this method only uses DNA polymerase as opposed to multiple, expensive enzymes required by pyrosequencing (Pettersson *et al*., 2008).

2.83 Ion Torrent sequencing technology

Also known as semiconductor sequencing, the Ion Torrent sequencing is a method of DNA sequencing based on the detection of hydrogen ions that are released during the polymerization of DNA. The technology was licensed from DNA Electronics Ltd (Davies, 2011; GW, 2010), developed by Ion Torrent Systems Inc. and was released in February 2010 (Rusk, 2011). The new semiconductor chip technology is capable of directly translating chemical signals into digital information and its first application is sequencing DNA. The first-generation Ion 314™ chip contains 1.4 million wells that sense and record the sequencing reaction. Figure 2.5 is a schematic illustration of a single well of an Ion Torrent sequencing chip showing pH-sensing layer following the incorporation of complementary dNTP.

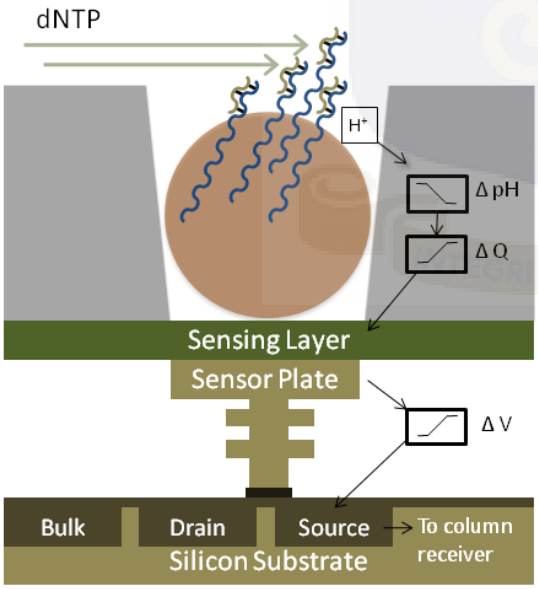


Figure 2.5: Schematic cross-section of a single well of an Ion Torrent sequencing chip. (http://www.iontorrent.com/lib/images/PDFs/amplicon_application_note_040411.pdf)

The Ion Torrent sequencing chemistry is remarkably simple. Naturally, the incorporation of a dNTP into a growing DNA strand involves the formation of a covalent bond and the release of pyrophosphate and a positively charged hydrogen ion (Rusk, 2011; Davies, 2011) resulting in a detectable local change of pH. Each Microwell on a semiconductor chip contains approximately one million copies of one single-stranded template DNA molecule to be sequenced and one DNA polymerase. These microwells are sequentially flooded with unmodified nucleotides one after the other (Rusk, 2011; Pennisi, 2010; Perkel, 2011). The ion sensitive layer, below which is an Ion-sensitive field-effect transistor (ISFET) ion sensor, is beneath the layer of microwells. The hydrogen ion released following the incorporation of complementary dNTP changes the pH of the solution and triggers the ISFET ion sensor to detect the change in pH (Rusk, 2011; Pennisi, 2010) in that microwell, essentially going directly from chemical information to digital information. The series of electrical pulses transmitted from the chip to a computer is translated into a DNA sequence, with no intermediate signal conversion required (Pennisi, 2010; Davies, 2010). The use of labeled nucleotides and optical measurements are avoided because nucleotide incorporation events are measured directly by electronics. The noncomplementary dNTP molecules are washed out before the next cycle when a different dNTP species is introduced. Signal processing and DNA assembly can then be carried out in software.

2.84 Relative importance of Ion Torrent sequencing

The major benefits of ion semiconductor sequencing are rapid sequencing speed and low upfront and operating costs (Perkel, 2011; Davies, 2010) unlike other next-generation sequencing (NGS) methods. This has been enabled by the avoidance of modified nucleotides and optical measurements. Because there is direct detection and, therefore, no scanning, no cameras and no light, each nucleotide incorporation is measured in seconds enabling very

short run times. Moreover, Ion Torrent sequencing technology is uniquely suited for amplicon sequencing and small genomes such as those of bacteria because it is simple, fast, scalable and cost effective. However, other NGS technologies are more cumbersome, too slow or more expensive to sequence smaller genomes or regions of genome. In spite of the advantages of using Ion Torrent sequencing method, it is limited by the difficulty to differentiate signals generated from a high repeat number from repeats of a similar, but different number. Moreover, it has shorter read length compared to other sequencing methods such as Sanger sequencing or pyrosequencing. As observed by Rusk (2011) and Perkel (2011), the read length achieved by Ion Torrent Systems Inc. is 200 base pairs per run.

2.90 Loop-mediated isothermal amplification (LAMP), PCR and DNA amplification

2.91 Characteristics of LAMP technique

Loop-mediated isothermal amplification (LAMP) is a novel nucleic acid amplification method that amplifies DNA with high specificity, efficiency and rapidity under isothermal conditions using a set of four specially designed primers and a DNA polymerase with strand displacement activity (Notomi *et al*., 2000; Nagamine, *et al*., 2002). By designing four different primers to recognize six distinct regions on the target gene, the LAMP method is able to specifically amplify the target gene.

The LAMP method developed by Notomi and colleagues (2000) can amplify a few copies of DNA to 10^9 in less than an hour under isothermal conditions and with greater specificity. Based on its isothermal condition, the LAMP method has no need for time loss for thermal change, resulting in extremely high amplification efficiency. Therefore, the LAMP assay has the advantage in specificity, selectivity and rapidity over other nucleic acid amplification

methods (Mori *et al*., 2001). It is simple and cost-effective, as it does not require special reagents or sophisticated equipments. Moreover, nucleic acid amplification can be done with RNA templates simply by adding reverse transcriptase and following the same procedure as with DNA templates.

2.92 Gene amplification efficiency: LAMP and PCR methods

Polymerase chain reaction (PCR) is a technique in molecular genetics that permits artificial amplification of minute quantities of deoxyribonucleic acid (DNA) to levels that can easily be detected in the laboratory. PCR uses a thermostable DNA polymerase, from *Thermus aquaticus*, which enables the amplification reaction to be performed at higher temperatures and significantly improves the specificity, yield, sensitivity and length of products that can be amplified (Saiki *et al*., 1988). The PCR can synthesize millions of copies of a specific DNA sequence in a brief in *vitro* reaction (Erlich *et al*., 1988) and has been found to amplify single-copy genomic sequences by a factor of more than 10 million with very high specificity. PCR-based detection is widely used (Saiki *et al*., 1985; Saiki *et al*., 1988). It is, therefore, at the forefront of molecular diagnostic technology today and represents the rapid, sensitive and specific method, which provides the possibility to detect few copies of the genetic material of interest.

Despite the simplicity and the obtainable magnitude of amplification, PCR detection is costly, requires skilled laboratory workers and is, therefore, rarely available as a routine diagnostic tool especially in resource-poor countries where there are inadequate laboratory facilities and trained personnel and financial constraints. Unlike PCR, which requires high precision and expensive thermal cycler, the LAMP method does not require an expensive thermocycler because all reactions can be performed at a constant temperature ranging from 60°C to 65°C.

The amplification specificity for LAMP is extremely high and is known to be better than that of PCR (Mori *et al*., 2001) because the LAMP reaction requires a set of four oligonucleotide primers that recognize six distinct regions on the target DNA. Although the use of multiple primers, such as in nested PCR has improved amplification specificity for the target sequence, residual co-amplification of irrelevant sequences still causes a general setback in nucleic acid amplification, particularly for diagnostic use.

Polymerase chain reaction (PCR) requires more elaborate method for detection of amplified product. However, visualization of DNA products on gel electrophoresis is not required for assessing successful DNA amplification because a positive LAMP reaction causes the solution to become cloudy due to the formation of magnesium pyrophosphate byproduct (Mori *et al*., 2001). Although the detection limit of LAMP is expected to be equal to or higher than that of PCR, the detection time is shorter (Dukes *et al*., 2006; Hara-Kudo *et al*., 2007; Goto *et al*., 2007) and the LAMP reaction can be accelerated by using two specially designed loop primers (Nagamine *et al*., 2002). For these reasons, gene amplification by the LAMP method is superior to that by PCR (Goto *et al*. 2009). LAMP method is simpler and more cost-effective, as it does not require special reagents or sophisticated equipments. The development of a diagnostic technique based on LAMP for Buruli ulcer particularly in endemic and resource-poor countries cannot be overemphasized.

2.93 Application of LAMP and PCR techniques

The loop-mediated isothermal amplification (LAMP) method has wide application in various fields including clinical diagnosis and the detection of food-borne pathogens in the food industry. Using LAMP and PCR methods, Denguo and colleagues (2008) carried out

comparative analysis of *Salmonella* DNA detection sensitivity in polluted raw milk and observed differences in detection limit or sensitivity based on the purity of DNA template used. They observed plentiful DNA amplification products when pure DNA template was used. However, contrary to previous suggestion that DNA extraction step can be omitted in LAMP assay on account of superior toleration of LAMP for biological substances to PCR (Kaneko *et al*., 2007), there was no amplified product when DNA extraction step was omitted. In other words, the detection limit or sensitivity of LAMP was lower than that of PCR when impure DNA template was used. This may be explained by the presence of a wide range of inhibitors including organic and inorganic substances such as enzymes, polysaccharides, fats, proteins and salts (Kaneko *et al*., 2007) or indirect influence of the primers on the LAMP reaction, as DNA purity did not affect the detection of other foodborne pathogens in raw milk with LAMP. This study suggests further research on the stability and sensitivity of LAMP application in clinical diagnosis and food detection fields.

Based on the *invA* gene present in all *Salmonella* isolates, Wang *et al*. (2008) developed a LAMP method for rapid detection of the food-borne pathogen *Salmonella* and evaluated its specificity and sensitivity. The LAMP method was found to be very specific and the presence of 100 ng of non-*Salmonella* genomic DNA in the reaction with 1 pg *Salmonella*, the target DNA, neither adversely affected the amplification efficiency nor generated significant background when separated by agarose gel electrophoresis. This result is consistent with that of Notomi and colleagues (2000). Using serial 10-fold dilution of the extracted DNA and comparing with that of conventional PCR, the detection limits were found to be 100 fg DNA/tube and 1 pg DNA/tube for the LAMP assay and PCR respectively. Further comparison of the sensitivity of LAMP and PCR methods revealed a 10-fold sensitivity for the LAMP method higher than that of PCR.

The LAMP technology has been useful for the detection of parasites in the *Trypanosoma brucei* group. In evaluating LAMP-based method for specificity and sensitivity and comparing with PCR-based methods, Kuboki and colleagues (2003) observed up to 100 times higher sensitivity of LAMP-based method for *in vitro* detection of trypanosomes than that of PCR-based methods. Moreover, *in vivo* studies in mice infected with human-infective *T. brucei gambiense* detected trypanosome DNA by LAMP reaction at all days postinfection, clearly indicating the extremely high sensitivity of the LAMP reaction. This study highlights the potential clinical importance of LAMP as a diagnostic tool for identification of African trypanosomiasis.

Moreover, LAMP has demonstrated the potential to be useful in field conditions for diagnosis of trypanosome infections without being affected by ambient temperatures of tropical and sub-tropical countries. In evaluating the stability of LAMP reagents stored at 25^{o}C, 37^{o}C and the recommended storage temperature of –20^{o}C, Thekisoe and colleagues (2009) observed no significant differences in detection sensitivity of LAMP among the reagents stored at these temperatures. There was amplification of serially diluted genomic DNAs (extracted by different methods) of *T. b. gambiense* with high sensitivity. However, LAMP detection sensitivity was poor when fresh blood, as DNA template, was added directly into reactive solution, in agreement with previous findings on the use of impure DNA template (Denguo *et al*., 2008).

CHAPTER THREE

3.0 Methodology

3.10 Description of study area

The study was carried out in eight Buruli ulcer (BU) endemic communities in the Asante Akim North District (AAND) in the Ashanti Region of Ghana. The district is located in the eastern part of the Ashanti Region and shares boundaries with Sekyere East on the north, Kwahu South on the east, Asante Akim South on the south and Ejisu-Juaben Municipal on the west. It occupies a total area of 1,462 km^2 found at 6°37′5″N 1°12′36″W, and has a population of 169,976 as at the 2010 national census (MCE, 2012). Figure 3.1 is a map showing BU endemic communities in AAND, distances between the various communities and the network of water bodies. Farming is the predominant occupation among residents of the district.

Buruli ulcer disease (BUD) has been prevalent in the district since the early 1970s, starting from villages in the Afram Plains sector of the district and gradually spreading to other towns in the district. The Agogo Presbyterian Hospital (APH), the district hospital, has been in the forefront of BU management since the early 1970s and is one of the first hospitals to use the services of trained community volunteers to identify cases of the disease in the district. Currently, there are about 50 trained volunteers in the district involved in active case search in their respective communities and are making immense contribution to the management of the disease. The hospital is a designated Training Centre for BU management by the World Health Organisation and the Ministry of Health, Ghana.

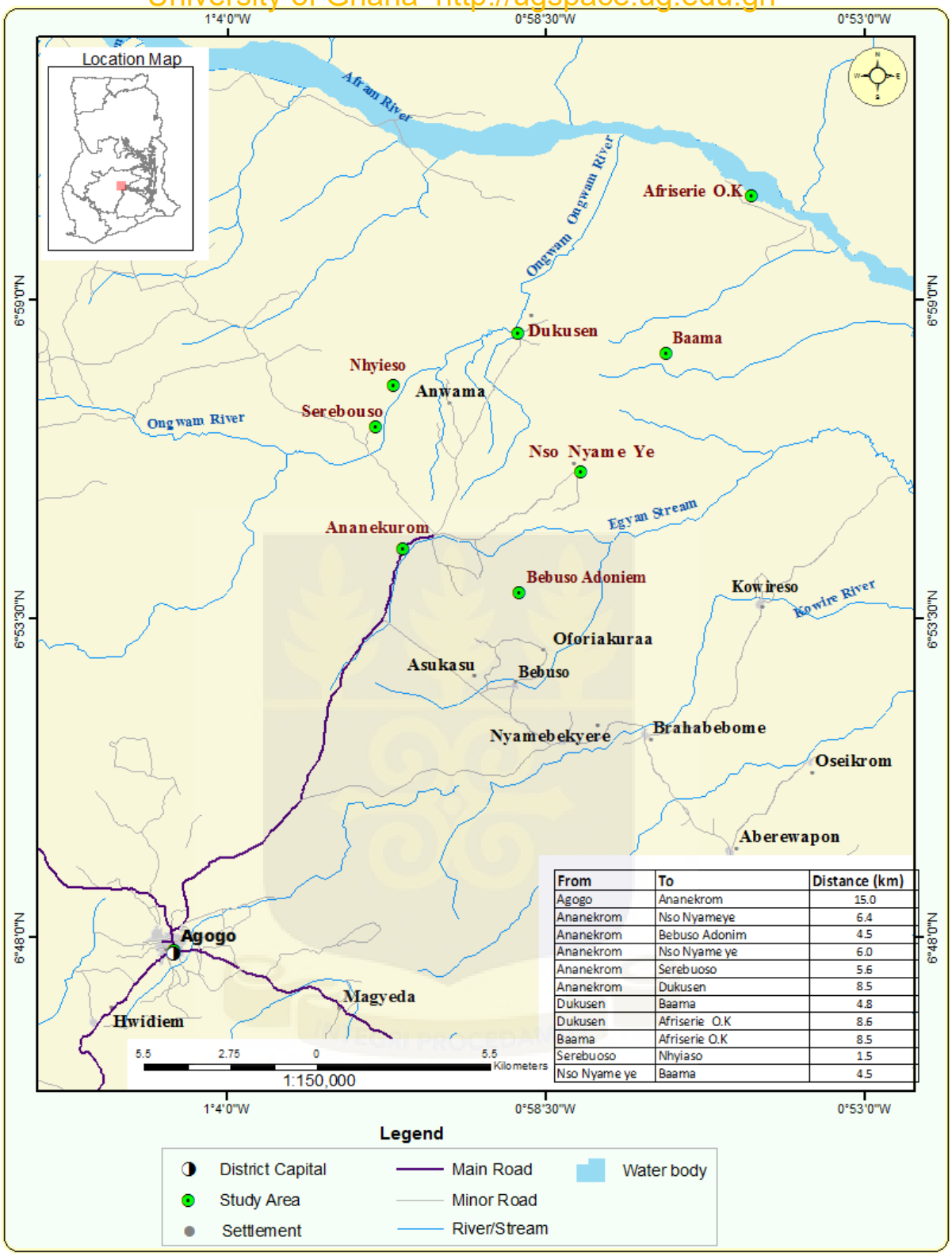


Fig. 3.1: Buruli ulcer endemic communities in Asante Akim North District.

3.11 Statement on ethical approval

Ethical approval for the use of clinical samples was obtained from the Ethical Review Board of the Noguchi Memorial Institute for Medical Research (NMIMR), University of Ghana, Legon. Specimens used were anonymously taken from a collection of patients' specimens processed for diagnosis of BU from Agogo Presbyterian Hospital in Ghana.

3.12 Collection and processing of samples

3.121 Collection of clinical samples and confirmation of Buruli ulcer diagnosis

Community-based volunteers referred most of the suspected cases to the APH. The patients are initially diagnosed based on their clinical presentation, which include the presence of nodule (s) and painless or minimally painful ulcer with undermined edge and a whitish-yellow necrotic base. In addition, residence or previous residence in an endemic area, persons under 15 years of age with about 85% of lesions on the limbs are important diagnostic clues. Swabs and fine-needle aspiration (FNA) sampling techniques were used to collect samples from ulcerative and non-ulcerative lesions respectively for laboratory confirmation of infection, as recommended by the World Health Organisation (WHO, 2010). Samples obtained by swabs were taken from the undermined edges of ulcers where the bacteria are usually concentrated. Both swab and FNA samples were kept in transport medium (liquid Middlebrook 7H9 broth supplemented with polymyxin B, amphotericin B, nalidixic acid, trimethoprim and azlocillin, PANTA). All samples were labelled and held at 4 ºC in an insulated container with ice packs and transported to the Noguchi Memorial Institute for Medical Research (NMIMR) for laboratory analysis.

The swab and FNA samples were decontaminated using the NaOH solution, malachite green, cycloheximide and oxalic acid method. This method has been found to produce less

contamination (Portaels *et al*., 1988) and to yield higher viable counts of mycobacteria (Livanainen, 1995) compared with decontamination with NaOH followed by oxalic acid. Cycloheximide and malachite green are antifungal agents and are, therefore, useful in the isolation of *Mycobacterium* species. A volume of 500 µl of each sample in phosphate buffered saline (PBS) was pipetted and dispensed into eppendorf tube. The samples were span at 14000 rpm for 5 minutes to get sediments and the supernatant carefully removed with pipette. To each of the sediments, 500 µl of NaOH, 100 µl of cycloheximide and 500 µl of 0.2% malachite green were added and stirred by drawing and dispensing the mixture with pipette. The tubes and their contents were left for 30 minutes at room temperature, centrifuged for 5 minutes at 14000 rpm and 700 µl supernatant carefully removed. To each of the mixture, 1000 µl oxalic acid was added to neutralize it and left for 20 minutes followed by centrifugation at 14000 rpm for 5 minutes. The supernatant for each sample was carefully removed and four drops of PBS added to sediment and mixed for inoculation onto culture medium.

Decontaminated suspensions were inoculated onto Lowenstein-Jensen medium and incubated at 31ºC for the cultivation of *M. ulcerans*. DNA was extracted from portions of the suspensions, as described in Section 3.123, and used as template in IS*2404* PCR for the identification of *M. ulcerans*. In addition to *M. ulcerans* isolates from the Asante Akim North District (AAND), eight (8) isolates from different endemic areas were included in the study, the details of which are described in the following section.

3.122 *M. ulcerans* isolates used for the genotyping (Whole genome sequencing)

A total of 19 *M. ulcerans* isolates were investigated. Eleven (11) were recovered from patients from the AAND who reported at the Agogo Presbyterian Hospital. Also included in the study were isolates from two other endemic areas, Amansie West and Ga Districts of Ghana, described previously (Roltgen *et al*., 2010). Two of the isolates, Mu_06-3845 and Mu_M3 were cultivated from aquatic insects, while the remaining ones were isolated from humans. Portaels and colleagues (2008) isolated *M. ulcerans* Mu_06-3845 from a Hemiptera (Water Strider, *Gerris* sp.) collected from a swamp in a BU endemic region in Benin and it represents the first fully characterized culture of the agent of BU from an environmental source. The Mu_M3 is a recent isolate also from a Hemipteran aquatic insect in the family Belostomatidae (giant water bugs). Table 3.1 gives information on *Mycobacterium ulcerans* isolates used in the sequencing study.

Table 3.1: Information on *Mycobacterium ulcerans* isolates used in sequencing study

Identifier of Isolate	Place of origin	Year of isolation	Source	Reference
Mu_NM33.04	Amansie West District, Ghana	2004	Human	Roltgen *et al*., 2010
Mu_NM43.02	Ga District, Ghana	2002	Human	Roltgen *et al*., 2010
Mu_NM14.01	Ga District, Ghana	2001	Human	Roltgen *et al*., 2010
Mu_NM54.02	Ga District, Ghana	2002	Human	Roltgen *et al*., 2010
Mu_NM49.02	Ga District, Ghana	2002	Human	Roltgen *et al*., 2010
Mu_06-3846	Lalo, Couffo, Benin	1997	Human	Doig *et al*., 2012
Mu_07-1082	Adjohoun, Ouémé, Benin	2003	Human	Doig *et al*., 2012
Mu_06-3845	Houedja, Ouinhi, Zou, Benin	2000	Aquatic insect	Portaels *et al*., 2008
Mu_F75	Asante Akim North District, Ghana	2010	Human	This study
Mu_F65	Asante Akim North District, Ghana	2011	Human	This study
Mu_F70	Asante Akim North District, Ghana	2010	Human	This study
Mu_S77	Asante Akim North District, Ghana	2010	Human	This study
Mu_M3	Asante Akim North District, Ghana	2010	Aquatic insect	This study
Mu_F74	Asante Akim North District, Ghana	2011	Human	This study
Mu_F85	Asante Akim North District, Ghana	2010	Human	This study
Mu_F13	Asante Akim North District, Ghana	2011	Human	This study
Mu_S38	Asante Akim North District, Ghana	2010	Human	This study
Mu_S43	Asante Akim North District, Ghana	2010	Human	This study
Mu_F64	Asante Akim North District, Ghana	2010	Human	This study

3.123 DNA extraction from clinical samples

The extraction of DNA from clinical samples was carried out as per the modified Boom protocol described previously (Durnez *et al*., 2009). From each specimen, 250 µl was added to 250 µl of 2 × lysis buffer (1.6 M GuHCl, 60 mM Tris pH 7.4, 1% Triton X-100, 60 mM EDTA, Tween-20 10%), 50 µl Proteinase K (20 mg/ml) and 500 µl glass beads in a labeled 1.5 ml microcentrifuge tube. For quality control, bacterial suspension in sterile UV-treated milli-Q water with a concentration of about 250 AFB/extract and sterile UV-treated milli-Q water served as positive and negative extraction controls respectively. The specimens were vortexed and incubated overnight at 60 °C in a shaking water bath at a speed of 200 rpm. To capture DNA, 40 µl diatomaceous earth stock solution (10 g diatomaceous earth, Sigma Aldrich ChemiGmbH, in 50 ml sterile milli-Q water containing 500 µl of 37% (wt/vol) HCl) was added to each specimen. The specimens were vortexed and incubated horizontally at 37 °C in a shaker incubator at 200 rpm for 60 min.

The specimens were centrifuged at 14000 g for 10 s and the supernatants were carefully removed using pipette. The resulting pellets were washed twice with 70% ethanol. In each case, 900 µl 70% ethanol (2-8 °C) was added to each pellet, vortexed for 5 sec and centrifuged for 1 min. at 14000 g. The supernatants were removed followed by 900 µl acetone washing of the pellets, vortexing, centrifugation and removal of the supernatants. The washed pellets were dried at 50 °C for 20 min in an activated thermoblock placed in the laminar flow. The dried pellets were resuspended in 100 µl TE to release DNA into solution and incubated at 58 °C for 20 min with regular vortexing to have a good distribution of the diatomaceous earth in the TE. The specimens were centrifuged at 14000 g for 10 s and the resulting supernatants containing the DNA were kept at −20 °C until further analysis.

3.20 *Mycobacterium ulcerans* DNA preparation and genome sequencing

3.21 DNA library preparation

3.211 Preparation of adapter-compatible DNA

Eleven isolates used in the study were sequenced using the Ion Torrent platform. Enzymatic fragmentation of *M. ulcerans* genomic DNA (gDNA) into blunt-ended fragments was carried out using Ion Shear™ Plus Reagents Kit in the Ion Xpress™ Plus Fragment Library Kit as described by the manufacturers (Ion Torrent, Life Technologies). Based on a 200 bp sequencing chemistry, library size of 200-base-read and, therefore, a median fragment size of 200–300 bp were used. Ten (10) µl at 100 ng/µl in nuclease-free water was prepared for 1 µg input. The Ion Shear™ Plus 10X Reaction Buffer and the Ion Shear™ Plus Enzyme Mix II were each vortexed for 5 seconds, pulse-spun to bring the contents to the bottom of the tubes, and placed on ice. Ten (10) µl of gDNA (100 ng/µl) was dispensed into a 1.5 ml Eppendorf LoBind^{R} tube, followed by 5 µl Ion Shear™ Plus 10X Reaction Buffer and 25 µl nuclease-free water, mixed vigorously by vortexing for 5 seconds and pulse-spun to bring the contents to the bottom of the tube. Ten (10) µl of the Ion Shear™ Plus Enzyme Mix II was added to the sample and mixed with the DNA and buffer to form a total reaction volume of 50 µl. Mixing of the reaction was rapidly done by pipetting up and down 8–10 times, incubated in a heat block at 37°C for 15 minutes and 5 µl of Ion Shear™ Stop Buffer added immediately after incubation. The mixture was mixed thoroughly by vortexing for at least 5 seconds and stored on ice.

The fragmented DNA was purified by adding 99 µl of Agencourt^{R} AMPure^{R} XP Reagent (1.8X sample volume) to the sheared DNA sample, pipetted up and down 5 times to thoroughly mix the bead suspension with the DNA. The mixture was pulse-spun and incubated at room temperature for 5 minutes. The reaction tube was pulse-spun and placed in

a magnetic rack until the solution cleared of brown tint when viewed at an angle. The supernatant was carefully removed and discarded without disturbing the bead pellet. Freshly prepared 70% ethanol (500 µl) was added and incubated for 30 seconds, turning the tube around twice in the magnet to move the beads around. The supernatant was removed and discarded without disturbing the pellet after the solution cleared. Ethanol washing was repeated and the tube pulse-spun and placed back in the magnetic rack. Any remaining supernatant was carefully removed with a 20 µl pipettor without disturbing the pellet. The beads were air-dried at room temperature for < 5 minutes. The tube was removed from the magnetic rack, and 25 µl of Low TE added directly to the pellet to disperse the beads. The suspension was pipetted up and down for 5 times and the sample vortexed for 10 seconds to mix thoroughly, pulse-spun and placed in the magnetic rack for at least 1 minute until the solution cleared. The supernatant containing the eluted DNA was transferred to a new 0.2 mL PCR tube without disturbing the pellet.

3.212 Adapter ligation, purification of ligated DNA and size selection

The reagents for non-barcoded libraries (~25 µl DNA, 10 µl 10X Ligase Buffer, 10 µl Adapters, 2 µl dNTP Mix, 41 µl Nuclease-free water, 4 µl DNA Ligase and 8 µl Nick Repair Polymerase) were combined in a 0.2 mL PCR tube and mixed well by pipetting up and down. The tube was then placed in a thermal cycler and run for 15 min at 25°C, 5 min at 72°C and held at 4°C before transfer to a 1.5 mL Eppendorf LoBind^{R} tube for the next cleanup step. A volume of 140 µl (1.4X sample volume) of Agencourt^{R} AMPure^{R} XP Reagent was added to the ligated DNA sample and purified as described above. However, the supernatant containing the eluted DNA was transferred to a new 1.5 mL Eppendorf LoBind^{R} tube without disturbing the pellet.

The library was size-selected with the E-Gel® SizeSelect™ Agarose Gel. Prior to loading, 20 µl of Low TE was added to the purified ligated DNA to bring the total volume to 40 µl and the 50 bp DNA Ladder was diluted in Low TE buffer to 25 ng/µl (1:40 dilution). Twenty (20) µl of ligated DNA was added to the loading well (top row). Ten (10) µl of diluted DNA ladder was added into the middle well, lane M, and no more than 250 ng (10 µl of 1:40 dilution) of the DNA Ladder was loaded for the 200-base-read libraries. Added to each of the empty wells in the top row and large wells in the bottom row (collection wells) was 25 µl nuclease-free water and 10 µl to the center well (lane M) of the bottom row. The gel was ran for 12–14 minutes and monitored till the desired fragment size range was in the collection well before stopping the run. Using a pipette, the sample was carefully collected from the collection wells.

3.213 DNA Library amplification and purification

The volume of the unamplified library (~60 µl) was adjusted to 50 µl, volume of amplification reaction. This volume (50 µl) was combined with 200 µl of Platinum^{R} PCR SuperMix High Fidelity and 10 µl Library Amplification Primer Mix to form a total reaction volume of 260 µl. The 260 µl reaction mix was split into two 0.2 ml PCR tubes, each containing about 130 µl. The tubes were placed into a thermal cycler with the number of cycles minimized by setting at five to avoid over-amplification and introduction of PCR-induced errors. The thermal profile for PCR amplification of *M. ulcerans* genomic DNA included an initial denaturation step of 95°C for 15 sec following 5 min of holding, annealing at 58°C for 15 sec, elongation at 70°C for 1 min and a holding stage at 4°C. The previously split PCRs were combined in a new 1.5-ml Eppendorf LoBind^{R} tube. A volume of 390 µl (1.5X sample volume) of Agencourt^{R} AMPure^{R} XP Reagent was added to the sample and purified as described earlier. However, 20 µl of Low TE was directly added to the pellet to

disperse the beads and the supernatant containing the eluted DNA transferred to a new 1.5 ml Eppendorf LoBind[R] tube without disturbing the pellet. An aliquot of the library (1 µl, 1:10) was analyzed on a Bioanalyzer[R] instrument with an Agilent[R] High Sensitivity DNA Kit to assess the size distribution and to determine the library dilution required for template preparation.

3.22 DNA template preparation for sequencing reaction

Following dilution of the stock library to the optimal input concentration (~26 pM) to achieve 10-30% template positive Ion PGM™ Template OT2 200 Ion Sphere™ Particles (ISPs), 25 µl of the diluted library was used in the amplification reaction. The preparation of reagents for the amplification solution (Appendix I) was based on the manufacturer's information provided in the Ion PGM™ Template OT2 200 Kit User Guide. Nuclease-free water (25 µl), 500 µl of the reagent mix, 300 µl of PCR reagent B, enzyme mix (50 µl), 25 µl of diluted library solution (to target 10–30% of positive unenriched Ion PGM™ Template OT2 200 Ion Sphere™ Particles) were dispensed into a 1.5-ml Eppendorf LoBind® Tube at 15°C to 30°C in that order and pipetted thoroughly to mix. The solution was vortexed at maximum speed for 5 seconds and centrifuged for 2 seconds.

The Ion PGM™ Template OT2 200 Ion Sphere™ Particles (ISPs), in solution, was vortexed at maximum speed to resuspend the particles, centrifuged for 2 seconds and pipetted up and down to mix. The ISPs (100 µl) was added to the amplification solution (900 µl) and the complete amplification solution (1000 µl) vortexed at maximum speed for 5 seconds and proceeded immediately to the next stage of the amplification solution preparation and installation. Following 2 seconds centrifugation, the entire 1000 µl amplification solution was pipetted and dispensed through the sample port on the Ion PGM™ OneTouch Plus Reaction

Filter Assembly. Ion OneTouch™ Reaction Oil (1500 µl) was added through the sample port; 1000 µl was first added followed by 500 µl. The filled Ion PGM™ OneTouch Plus Reaction Filter Assembly was installed on the Ion OneTouch™ 2 Instrument so that it was firmly seated and the instrument ran for 16 hours. The instrument centrifuged the samples for 10 minutes after the run. This was followed immediately by recovery of the ISPs.

The recovery tubes were carefully removed from the instrument and a pipette used to remove all but 50 µl of the recovery solution from each of the recovery tubes without disturbing the pellet of the ISPs. The ISPs were resuspended in the remaining recovery solution using a new pipette tip for both tubes and pipetted up and down until each pellet dispersed in the solution. The suspensions from the recovery tubes were transferred into well 1 of the 8-well strip from the Ion OneTouch™ ES Supplies Kit for quality assessment of the unenriched ISPs using QubitR 2.0 Fluorometer.

To enrich the ISPs, fresh Melt-Off Solution was prepared by adding 40 µl 1 M NaOH to 280 µl Tween^{R} Solution to form a total volume of 320 µl. This was followed by washing and resuspension of Dynabeads® MyOne™ Streptavidin C1Beads (Appendix II) and filling of the 8-well strip, which already contains 100 µl template-positive ISP sample in Well 1. Dispensed into Well 2 was 130 µl of DynabeadsR MyOne™ Streptavidin C1 Beads resuspended in MyOne™ Beads Wash Solution with Wells 3, 4 and 5 having 300 µl each of Ion OneTouch™ Wash Solution. Wells 6 and 8 remained empty, but 300 µl of freshly-prepared Melt-Off Solution was dispensed into Well 7. An opened 0.2-mL PCR tube with neutralization solution and the 8-well strip were loaded on the Ion OneTouch™ ES and the contents of Well 2 pipetted up and down to resuspend the beads before performing the run for ~35 minutes. Immediately after the run, the PCR tube containing the enriched ISPs was

securely closed, removed from the instrument and the contents mixed by gently inverting it 5 times, ensuring that the sample was in ~230 µl of Melt-Off Solution, Ion OneTouch™ Wash Solution and Neutralization Solution.

3.23 *M. ulcerans* genome sequencing

Genome sequencing was carried out on Ion Torrent Personal Genome Machine (PGM) using 316D chips (one chip per genome) and 200 bp sequencing chemistry. The sequencing primer was thawed on ice prior to preparing the template-positive ISPs for sequencing. The entire volume of the enriched ISPs was transferred to a new 0.2-mL PCR tube followed by vortexing of the Control Ion Spheres™ (CISs) and centrifuging the CISs for 2 seconds before taking aliquots. Five (5) µl of the CISs and 100 µl of Annealing Buffer were added to the PCR tube containing the enriched particles, mixed by pipetting up and down and centrifuged for 2 minutes at 15,500 × g. The supernatant was carefully removed with a pipette tip, leaving ~20 µl in the bottom of the tube. Annealing Buffer (150 µl) was added to the tube, mixed as previously described and the ISPs collected by centrifuging the tube for 2 minutes at 15,500 × g. The supernatant was carefully removed, leaving slightly less than 15 µl in the bottom of the tube by visually comparing the volume to a separate tube containing 15 µl of liquid. Sequencing Primer (12 µl) was added to the ISP sample and the mixture pipetted thoroughly up and down to disrupt the pellet completely before placing the tube in the thermal cycler, already programmed for 95°C for 2 minutes and then 37°C for 2 minutes, for running. After cycling, the reaction remained in the cycler at room temperature while chip check was in progress.

After the chip check (Appendix III) to ensure proper functioning, 3 µl of sequencing polymerase was added to the ISPs removed from the thermal cycler, mixed and incubated at

room temperature for 5 minutes. Thirty (30) µl of 50% annealing buffer/ 50% water mixture was added to the sample, bringing the total to 60 µl and the ISPs sonicated for 10 seconds using a standard 40 KHz sonication bath and loaded on the chip, which was loaded on the PGM™ Sequencer to begin sequencing.

3.24 Sequencing data analysis

Quality trimming, adapter trimming and length trimming were carried out for the raw sequence reads prior to mapping to the reference sequence. Quality trimming was based on Phred score >10. Ion Torrent adaptor sequences were removed and a minimum read length deemed to have passed the length trimming was 24bp. All reads including those from eight isolates, previously sequenced by Doig and colleagues (2012) using an Illumina GAIIx DNA sequencer, were mapped to *M. ulcerans* Agy99 genome used as a reference (*M. ulcerans* Agy99, CP000325). Agy99 was isolated in 1999 from an ulcerative lesion on the right elbow of a female patient from the Ga district of Ghana (Stinear *et al*., 2007).

Read mapping and SNP detection were done using an in-house Python utility called Nesoni, which uses a Short Read Mapping Package, SHRiMP2 (David *et al*., 2011) for read mapping. Nesoni identified SNPs and indels up to ~10 bp and predicted the consequences of SNPs and indels on protein-coding sequences. A minimum coverage of 10 reads was required to call a SNP. SNPs were identified using Nesoni v0.35 (http://www.bioinformatics.net.au), which used the sequence reads for each genome aligned with the Agy99 reference to construct a tally of putative differences at each nucleotide position, (indels excluded). This tally was then employed in a Bayesian model to decide whether a base (or deletion) could be called for the position, and if so, whether it differed from the reference sequence.

Phylogenetic analysis was performed using a distance method, based on pairwise nucleotide sequence alignments among all strains for the set of variable nucleotide positions. Phylogeny was inferred by neighbor-joining, using uncorrected *p* distances with bootstrapping as implemented in SplitsTree4 v4.11.3.

3.30 Materials and methods for VNTR study

3.31 Isolation of *Mycobacterium ulcerans,* decontamination and DNA extraction

In all, seventeen (17) *M. ulcerans* clinical isolates obtained from the Asante Akim North, Suhum-Kraboa-Coaltar and Dormaa Districts of Ghana were used in this study. The isolates were recovered from swabs and fine needle aspirates of clinically-diagnosed patients of BU. Table 3.2 gives information on *M. ulcerans* isolates used in the variable number of tandem repeat (VNTR) study. Decontamination of portions of clinical samples and cultivation of MU as well as DNA extraction of remaining portions of samples were carried out as described earlier.

Table 3.2: Information on *M. ulcerans* human isolates used in the VNTR study

Identifier of Isolate	Place of origin	Year of isolation	Reference
Mu_S8	Suhum-Kraboa-Coaltar District, Ghana	2012	This study
Mu_S21	Dormaa District, Ghana	2013	This study
Mu_S23	Asante Akim North District, Ghana	2009	This study
Mu_S25	Asante Akim North District, Ghana	2009	This study
Mu_F32	Asante Akim North District, Ghana	2010	This study
Mu_F37	Asante Akim North District, Ghana	2010	This study
Mu_F41	Asante Akim North District, Ghana	2010	This study
Mu_F45	Asante Akim North District, Ghana	2010	This study
Mu_F54	Asante Akim North District, Ghana	2010	This study
Mu_F61	Asante Akim North District, Ghana	2010	This study
Mu_F63	Asante Akim North District, Ghana	2010	This study
Mu_F65	Asante Akim North District, Ghana	2011	This study
Mu_F67	Asante Akim North District, Ghana	2010	This study
Mu_F77	Asante Akim North District, Ghana	2010	This study
Mu_F74	Asante Akim North District, Ghana	2011	This study
Mu_F85	Asante Akim North District, Ghana	2010	This study
Mu_S43	Asante Akim North District, Ghana	2010	This study

3.32 Location of tandem repeats in *M. ulcerans* and VNTR PCR

The assembled *M. ulcerans* genome sequence available at the Sanger Centre website (http.a://www.sanger c.uk/Projects/M_ulcerans/) was screened for the presence of tandem repeat (TR) loci by using Tandem Repeat Finder program. The study focused on TRs of the minisatellite category. This was based on the ease with which allelic differences can be resolved by agarose gel electrophoresis and by the relatively large range of variability of sequences of this type in *M. tuberculosis* (Mazars *et al*., 2001; Supply *et al*., 2000). Tandem repeats of the minisatellite category with repeat unit ranging from 9 bp to 80 bp (VNTR) were screened against a set of criteria. Out of the TR loci with period size ranges between 9 and 30 bp identified (Table 3.3), 25 loci with more than 95% nucleotide identity between individual repeat units and with two or more copies of repeats were selected for study. These criteria were based on previous observation that the presence of at least two identical or nearly identical repeats is necessary and sufficient to generate TR variability in the case of *M. tuberculosis* minisatellites (Supply *et al*., 2000). Table 3.3 shows detailed characteristics of the VNTR loci.

The primers for PCR amplification were designed based on sequences flanking each of the VNTR loci. The primers were designed using Primer 3 software. Table 3.4 gives sequences for primers of the twenty five newly-described VNTR loci used in this study. The Variable number of tandem repeat (VNTR) PCR was carried out using a Hotstar Taq DNA polymerase kit (QIAGEN). Sample DNA (3 µl) was added to 27 µl of a PCR mix containing 15.5 µl of water, 3 µl of 10× PCR buffer (containing 1.5 mM $MgCl_2$ at the final concentration), 0.4 µM of each primer, 0.2 mM of each deoxynucleotide triphosphate (Roche), 1× Q-solution, and 0.1 µl of HotstarTaq DNA polymerase (0.5 U).

The PCRs were run on a PTC 100 thermocycler (MJ Research, Waltham, Mass.) at 95°C for 15 min, followed by 40 cycles of 94°C for 30 s, 59°C for 1 min, and 72°C for 1 min 30 s and a final extension at 72°C for 10 min. A total of 3 µl of the PCR products was electrophoretically separated using a 2% small-fragment agarose gel (Eurogentec, Seraing, Belgium) in 0.5× TAE (20 mM Trisacetate, 0.5 mM EDTA at the final concentration) buffer at 100 V for 30 minutes, and the gel was then stained with ethidium bromide. The sizes of the amplicons were estimated by comparison with a 100-bp stepladder (Promega, Leiden, The Netherlands). Electrophoretic conditions were varied to 2% agarose gel in 0.5× TAE at 60 V for 3 hours to resolve better DNA bands of samples that appeared to differ in band sizes.

Table 3.3: Characteristics of VNTR loci in *Mycobacterium ulcerans*

Locus	Indices	Period size (bp)	Copy number	Percent matches	Entropy (0-2)
3	568650--568720	24	3	91	1.72
5	632030--632063	16	2.1	100	1.86
7	684443--684467	12	2.1	100	1.81
8	697790--697845	28	2	96	1.82
9	711175--711212	19	2	100	1.82
14	802540--802580	21	2	90	1.93
15	830607--830641	15	2.3	95	1.82
16	843055--843089	16	2.2	100	1.82
17	910531--910565	12	2.9	91	1.89
18	932775--932831	30	1.9	92	1.87
19	933011--933078	30	2.3	92	1.8
20	967661--967691	16	1.9	100	1.84
21	976248--976278	15	2.1	93	1.89
22	994984--995013	13	2.3	100	1.93
23	1025256--1025284	9	3.2	100	1.79
24	1093192--1093217	13	2	100	1.95
25	1168267--1168304	19	2	100	1.95
26	1292420--1292453	17	2	100	1.9
28	1452103--1452128	13	2	100	1.74
29	1461867--1461907	20	2	100	1.98
30	1496558--1496582	12	2.1	100	1.84
31	1502010--1502039	15	2	93	1.81
32	1689909--1689945	15	2.1	100	1.86
33	1601411—1601449	19	2.1	100	1.94
34	1524901--1524931	15	2.5	100	1.88

Table 3.4: Primers for VNTR loci in *Mycobacterium ulcerans* used in the study

Locus	Primer sequence (5´ to 3´)	Locus	Primer sequence (5´ to 3´)
3	F3 GTTGAAGTTCCAGCCTCG R3 TCGGTGCCACAGATACTC	22	F22 TTCCATCATTGTCAACTGTG R22 GGCAACGAGAGTCTTCTTA
5	F5 GACGAAGTTCAGGCAGAC R5 TCACCGATACCGACAATCA	23	F23 CTGGTGATGGTGGTGTTG R23 CCGTCATAGCCGAACAAG
7	F7 ATAGGCGGTTGCGAATTG R7 CTGCTGCTGACCTACCAA	24	F24 ACCGACAGTAGAGCGAAT R24 GGACAAGCGATCCAGTTG
8	F8 GATGAAGTGCCTGGTCTG R8 CGTCAACTACCGATAGAACT	25	F25 CTCCTGGTCCAGATTGATAC R25 CACTTGCATGGTGACAGT
9	F9 CGTTCAGTTCGGTTCCAA R9 GGCTTCATCCAGGTATCG	26	F26 GCTCGGAACTTCTGGATG R26 CTGCTGTCTCCTGAACCT
14	F14 CATCATCGCCGTCATCTG R14 AACGACCGCAACTACCTA	28	F28 TGAAGATCGCAATCCAGAA R28 GGTAGACATCGTCGTTGAT
15	F15 GCACAGCCTCTACAGAAT R15 GCCGATACCAACCTGATT	29	F29 AATTGTGGATTGAGGAGGTC R29 CGCTTCCGATGATGTGTT
16	F16 CGACAAGACGATTCAGTTC R16 TCGGATAGTGGTTGTAGGT	30	F30 GACGGTTGAACACCTCAA R30 TCGAGACGGTACAGATCG
17	F17 CGAGCAGATCCGAATAGC R17 TTTGTCCCACCCTTACCA	31	F31 GCTACCGATACCGATGTTG R31 AGTTGACCAGTGGCTCAT
18	F18 GATGTTGTTATTGCCGTAGT R18 CCTCAAGTTGGTTCAGTCA	32	F32 CCGCAGTAAGGTCGTAAG R32 CGAACTTCCACCACTTCA
19	F19 TGTTGTTATTGCCGTAGTTG R19 CCTCAAGTTGGTTCAGTCA	33	F33 TGATGGCGTTCAAGAAGG R33 GTGCTCTAACAGGACCAA
20	F20 GTGAATCAGTACCAGTATCTGT R20 GTATGCGATGGACGAGTC	34	F34 AGGTTGACGGTCGGTTAC R34 GGTTGATCCAATACGCCTAC
21	F21 GGATGCCAGTCAGTGATG R21 TGCGGAACAGTCAATACG		

3.40 Materials and Methods for LAMP study

3.41 Clinical specimens and DNA extraction

Thirty clinical specimens consisting of 20 swabs and 10 fine needle aspirates taken respectively from ulcers and pre-ulcerative lesions of suspected Buruli ulcer (BU) patients were used in this study. The fine needle aspirate specimens were kept in 1 ml phosphate buffered saline (PBS), but swabs were stored dry in sterile tubes. Each swab was transferred into a tube containing 2 ml milli-Q purified water (Millipore Corporation, Billerica, MA) and gently vortexed for 5 sec and then removed. The extraction of DNA from clinical specimens was carried out as per the modified Boom protocol (Durnez *et al*., 2009) described earlier.

To investigate the performance of the LAMP assay on crude DNA preparations, two types of DNA extracts for each clinical specimen were obtained. One crude extract consisted of 250 µl suspensions of the specimen boiled for 10 min followed by centrifugation at 14,000 rpm for 5 min (boiled extract). The other crude extract used was a 250 µl suspension of the unboiled specimen.

3.42 Loop-mediated Isothermal Amplification (LAMP) assay

The LAMP assay was performed using a set of 6 primers comprising 2 outer primers (Buruli-F3: CGAGAACAGCCTGCACTG, and Buruli- B3: CGGTTGGCGGTCAAAGC). Two inner primers (Buruli-FIP:GTGCGCCGTGTCCGGTATGGATACGCGATGTCACCTTC and Buruli- BIP: AGGTCCTAGCAACGCTACGCAAATCCGGCAGGCTTCGG), 2 loop primers Buruli-LF: GCCTTTGACGGTCTTCGTC, and Buruli- LB: (CACCGCGATCAATCTGCAC). The primers were designed using Primer Explorer (version 4; EikenChemical, Tokyo, Japan; http://primerexplorer.jp/elamp4.0.0/index.html).

Pocket warmer LAMP (pwLAMP) was performed using a loopamp DNA amplification kit (Eiken Chemical) described previously (Kaneko *et al*., 2007). Each 25 µl reaction mixture contained 1.6 µM each of FIP and BIP, 0.2 µM each of F3 and B3, 0.8 µM each of LF and LB, 2× reaction mixture (12.5 µl), 1 µl of *Bst* DNA polymerase, 1 µl of fluorescence detection reagent (Eiken Chemical), 3.5 µl distilled water and 1 µl sample. Reaction tubes were incubated at 60ºC for 60 min in the heat block (GeneAmp 9700, Applied Biosystems, Foster City, CA) while with the pwLAMP, the tubes were sandwiched in a twofold pocket warmer (Hokaron Haru-type, Lotte Health Products, Tokyo, Japan) surrounded by a paper towel and put in a Styrofoam box for 120 minutes (60 minutes reaction incubation). The reaction was terminated at 85ºC for 5 minutes and the results were read by eye in ambient light and also using UV illumination.

3.43 Detection limit of LAMP

Ten *M. ulcerans* strains grown on LJ slants were harvested and DNA was extracted as previously described (Durnez *et al*., 2009). Serial dilutions of purified *M. ulcerans* DNA containing 300,000, 30,000, 300, 30 and 3 copies of IS*2404* element per 5 µl were prepared. The number of copies of the insertion sequence element was determined based on the genome size of 5,806 kb and the presence of an average number of 207 copies of IS*2404*. This was used to determine the detection limit of the LAMP assays.

Standard curve of *M. ulcerans* genomic DNA (gDNA) specific for the IS*2404* target was added to the assays. This was achieved by preparing a standard curve in which the target of interest was present at 30,000 copies, 3,000 copies, 300 copies, 30 copies, and 3 copies. This was accomplished by calculating the mass of the genome and dividing this mass by the copy number of the target of interest per genome. In doing so the mass of *M. ulcerans* gDNA that

contained one copy of the target of interest was determined. The mass of *M. ulcerans* gDNA containing a copy number of interest, that is 300,000 to 30 copies, can be calculated subsequently by multiplying the copy number with the mass of *M. ulcerans* gDNA of one copy of the target of interest.

$$m_{x\_copies\_target} = \frac{m_{MUgenome}}{\#copies_{\_target}} \times x$$

The concentrations of *M. ulcerans* gDNA needed to achieve the copy number of interest is determined by dividing the last mentioned mass by the volume to be pipetted into each reaction. With this concentration a serial dilution of the *M. ulcerans* gDNA can be prepared from a stock concentration of *M. ulcerans* gDNA of known concentration.

To evaluate the specificity of the LAMP method for *M. ulcerans*, DNA extracts of eight *Mycobacterium* sp. (*Mycobacterium marinum, Mycobacterium tuberculosis, Mycobacterium avium, Mycobacterium intracellulare, Mycobacterium kansasii, Mycobacterium abscessus, Mycobacterium chelonae*), two *Mycobacterium ulcerans* strains (*Mycobacterium shinsuence*, Japanese strain and one African strain) and Jurkat (human T cell line) were examined.

3.44 Polymerase chain reaction for IS*2404*

Polymerase chain reaction (PCR) targeting IS*2404* was performed as described previously (Stinear *et al*., 1999). The first and second round PCRs used primers pGp1: 5´-AGGGCAGCGCGGTGATACGG-3´ and pGp2: 5´- CAGTGGATTGGTGCCGATCGAG-3´

and pGp3: 5´- GGCGCAGATCAACTTCGCGGT-3´ and pGp4: 5´ CTGCGTGGTGCTTTACGCGC-3´, respectively.

For the First round, the 30 µl reaction volume contained 3 µl DNA, 25 pmol/µl of each primer (pGp1 and pGp2), 3 µl of 10× PCR buffer (containing 1.5 mM magnesium chloride), 6.0 µl Q-solution, 0.2 mM deoxynucleotide triphosphates (dNTPs) and 1.0 U HotStar *Taq* polymerase (QIAGEN). For the second run, 1 µl of the first run product was added to 24 µl reaction volume containing, 25 pmol/µl of each primer (pGp3 and pGp4), 2.5 µl of 10× PCR buffer, 5.0 µl Q-solution, 0.2 mM dNTPs and 1.0 U HotStar *Taq* polymerase. Amplification for both rounds for 40 and 35 cycles respectively was carried out in an Eppendorf mastercycler thermal cycler as follows: denaturation at 95^{o}C for 15 min, 94^{o}C for 30 sec, 64^{o}C for 1 min, 72^{o}C for 1 min, 30 sec and a final extension at 72^{o}C for 10 min.

The second round PCR products were electrophoresed in a 2% TAE (0.04 M Tris-acetate and 0.001 M EDTA pH 8.0) agarose gel with ethidium bromide. The size of amplicons was estimated by comparison with 100 bp plus DNA ladder (Fermentas Life Sciences, EU) and visualized using Kodak Gel logic 100 Molecular Imaging System.

3.45 Statistical analysis for the LAMP study

A chi-squared test was performed to assess the statistical significance between the pw-LAMP method and IS*2404* PCR for *M. ulcerans* DNA amplification using Statistical Package for Social Sciences (SPSS, version 16.0; SPSS Inc., Chicago, IL) software.

CHAPTER FOUR

4.0 Description and Presentation of Results

4.10 Results for the whole genome sequencing study

4.11 Summary mapping report for *Mycobacterium ulcerans* isolates

Whole genome sequencing with single-end reads generated by Ion Torrent Sequencer was carried out for eleven of the nineteen (19) *M. ulcerans* isolates. For Mu_F64, one of the eleven isolates, over three million (3,304,625) reads that passed the trim were obtained with an average length of 191.70 bases (Table 4.1). These reads, constituting 95.73% of the total, were mapped to the reference isolate (Mu_Agy99). The remaining eight isolates, previously sequenced by Doig and colleagues (2012) using an Illumina GAIIx DNA sequencer, produced paired reads of varying sizes when the data were processed. With the exception of the insect isolate (Mu_M3) that had 59.79% mapped reads, all other isolates had higher percentage of mapped reads ranging from 94.48 to 97.35% (Table 4.1). Details of reads statistics for all the *M. ulcerans* isolates are shown in the table at Appendix IV.

Table 4.1: Mapped reads statistics for *Mycobacterium ulcerans* sequenced isolates

Isolate	Summary statistics					
	Sequence parameter	**Count**	**Percentage of reads (%)**	**Average length**	**Number of bases**	**Percentage of bases (%)**
Mu_Agy99	References	1	-	5,631,604.00	5,631,604	-
Mu_06-3845	Mapped reads	13,018,332	95.67	35.98	468,439,948	95.67
	Total reads	13,607,626	100.00	35.98	489,629,273	100.00
Mu_NM33.04	Mapped reads	24,837,745	97.25	31.70	787,305,172	96.96
	Total reads	25,541,270	100.00	31.79	812,003,478	100.00
Mu_NM43.02	Mapped reads	30,098,279	97.34	32.44	976,525,047	97.12
	Total reads	30,920,162	100.00	32.52	1,005,532,747	100.00
Mu_NM14.01	Mapped reads	30,162,918	97.19	32.22	971,893,819	96.93
	Total reads	31,033,720	100.00	32.31	1,002,641,795	100.00
Mu_NM54.02	Mapped reads	28,554,059	97.35	29.87	852,846,250	96.92
	Total reads	29,331,176	100.00	30.00	879,987,169	100.00
Mu_06-3846	Mapped reads	8,820,473	96.38	21.56	189,862,848	94.34
	Total reads	9,151,652	100.00	21.99	201,252,432	100.00
Mu_07-1082	Mapped reads	11,222,261	94.48	35.99	403,900,342	94.57
	Total reads	11,878,192	100.00	35.95	427,072,561	100.00
Mu_NM49.02	Mapped reads	28,606,620	96.98	32.96	942,863,737	97.03
	Total reads	29,497,100	100.00	32.94	971,675,695	100.00
Mu_F64	Mapped reads	3,304,625	95.73	191.70	633,483,852	95.72
	Total reads	3,452,022	100.00	191.72	661,829,520	100.00
Mu_F65	Mapped reads	3,281,977	95.98	191.13	627,278,983	95.70
	Total reads	3,419,531	100.00	191.69	655,484,005	100.00
Mu_F70	Mapped reads	3,165,967	96.80	192.13	608,291,220	96.75
	Total reads	3,270,603	100.00	192.24	628,749,072	100.00
Mu_S77	Mapped reads	2,714,298	95.58	194.43	527,727,799	95.61
	Total reads	2,839,711	100.00	194.37	551,967,423	100.00

Mu_Agy99	References	1	-	5,631,604.00	5,631,604	-
Mu_F74	Mapped reads	2,010,482	95.25	200.00	402,104,611	95.30
	Total reads	2,110,744	100.00	199.91	421,955,356	100.00
Mu_M3	Mapped reads	3,995,559	59.79	5,631,604.00	5,631,604	-
	Total reads	6,683,202	100.00	151.00	1,009,163,502	100.00
Mu_F85	Mapped reads	3,354,736	95.56	209.19	701,773,226	95.59
	Total reads	3,510,580	100.00	209.13	734,184,870	100.00
Mu_F13	Mapped reads	3,000,555	95.74	196.47	589,507,553	95.74
	Total reads	3,134,134	100.00	196.47	615,760,464	100.00
Mu_S38	Mapped reads	3,162,095	94.52	212.75	672,736,392	94.50
	Total reads	3,345,466	100.00	212.80	711,912,245	100.00
Mu_S43	Mapped reads	3,891,101	95.96	194.75	757,796,440	95.94
	Total reads	4,054,951	100.00	194.79	789,852,467	100.00
Mu_F75	Mapped reads	2,696,054	96.10	162.91	439,217,416	95.95
	Total reads	2,805,419	100.00	163.16	457,742,097	100.00

4.12 Mapping of reads and SNP detection

All the valid reads were mapped to the *M. ulcerans* (MU) reference Strain, Agy99, reported to have been isolated in 1999 from a female patient from the Ga district of Ghana (Stinear *et al*., 2007). Mapping and SNP detection were done using the Nesoni software, which uses a Short Read Mapping Package, SHRiMP2 (David *et al*., 2011). Nesoni identified SNPs and indels up to ~10 bp.

All the reads used for SNP detection satisfied the minimum requirement of 10 reads to call a SNP. Coverage was very high for many of the SNPs called. For example *M. ulcerans* isolate Mu_F64 had 2896 as the highest coverage for SNP detection and a minimum coverage of

eleven (11), Table 4.2. Figures 4.1 and 4.2 show reads coverage for all the *M. ulcerans* isolate used in the study. High frequencies and probabilities (> 90.15%) were recorded for all the single nucleotide polymorphisms (SNPs), implying identification of SNPs with high degrees of confidence. Individual MU isolates from the Asante Akim North District (AAND) that form the tight cluster (Mu_S43 to Mu_S77, Fig. 4.6) differ by only 2 SNPs, but have about two times SNP numbers recorded for other isolates within the same district. For example Mu_F65 had 451 SNPs compared to 227 SNPs for Mu_F70. Again, SNP numbers for the tight cluster isolates were about four times that of individual isolates from the Ga District. MU isolates from Benin had SNP numbers (254 &258) about two times SNP numbers for Ga District isolates except Mu_06_3846 that had unusually small SNP number (24). Details of SNP numbers and their respective ranges of coverage, frequencies and probabilities are shown in Table 4.2.

Table 4.2: Type of variants observed in *Mycobacterium ulcerans* isolates in comparison with reference isolate (Mu_Agy99)

M. ulcerans Isolate	Type and number of variant compared to Agy99			Coverage for SNPs	Frequencies (%)	Probabilities (%)
	SNP	MNV	InDel			
Mu_NM33.04	25	0	0	53 – 550	51.56 – 100.00	100
Mu_NM43.02	123	2	1	35 – 634	52.45 – 100.00	100
Mu_NM14.01	119	2	1	39 – 635	56.00 – 100.00	100
Mu_NM54.02	129	3	1	16 – 498	51.40 – 100.00	100
Mu_NM49.02	129	2	1	23 – 549	55.60 – 100.00	100
Mu_06-3846	24	1	7	15 – 34	53.30 – 100.00	95.15 – 100.00
Mu_07-1082	254	0	4	17 – 286	50.50 – 100.00	99.74 – 100.00
Mu_06-3845	258	2	3	17 – 151	50.00 – 100.00	99.70 – 100.00
Mu_F75	245	40	8	15 – 1637	50.00 – 100.00	95.69 – 100.00
Mu_F65	451	42	164	10 – 1259	35.10 – 100.00	90.41 – 100.00
Mu_F70	227	36	6	16 – 1071	46.90 – 100.00	97.85 – 100.00
Mu_S77	411	42	31	15 – 2026	46.70 – 100.00	90.89 – 100.00
Mu_M3	2,475	257	53	15 – 870	20.90 – 100.00	90.15 – 100.00
Mu_F74	414	41	31	15 – 1718	38.10 – 100.00	90.75 – 100.00
Mu_F85	411	40	29	15 – 2726	45.10 – 100.00	97.58 – 100.00
Mu_F13	249	40	11	16 – 2052	47.90 – 100.00	93.03 – 100.00
Mu_S38	243	38	6	15 – 2401	42.30 – 100.00	96.60 – 100.00
Mu_S43	425	40	21	15 – 3241	40.50 – 100.00	99.86 – 100.00
Mu_F64	254	37	149	11 – 2896	42.90 – 100.00	94.34 – 100.00

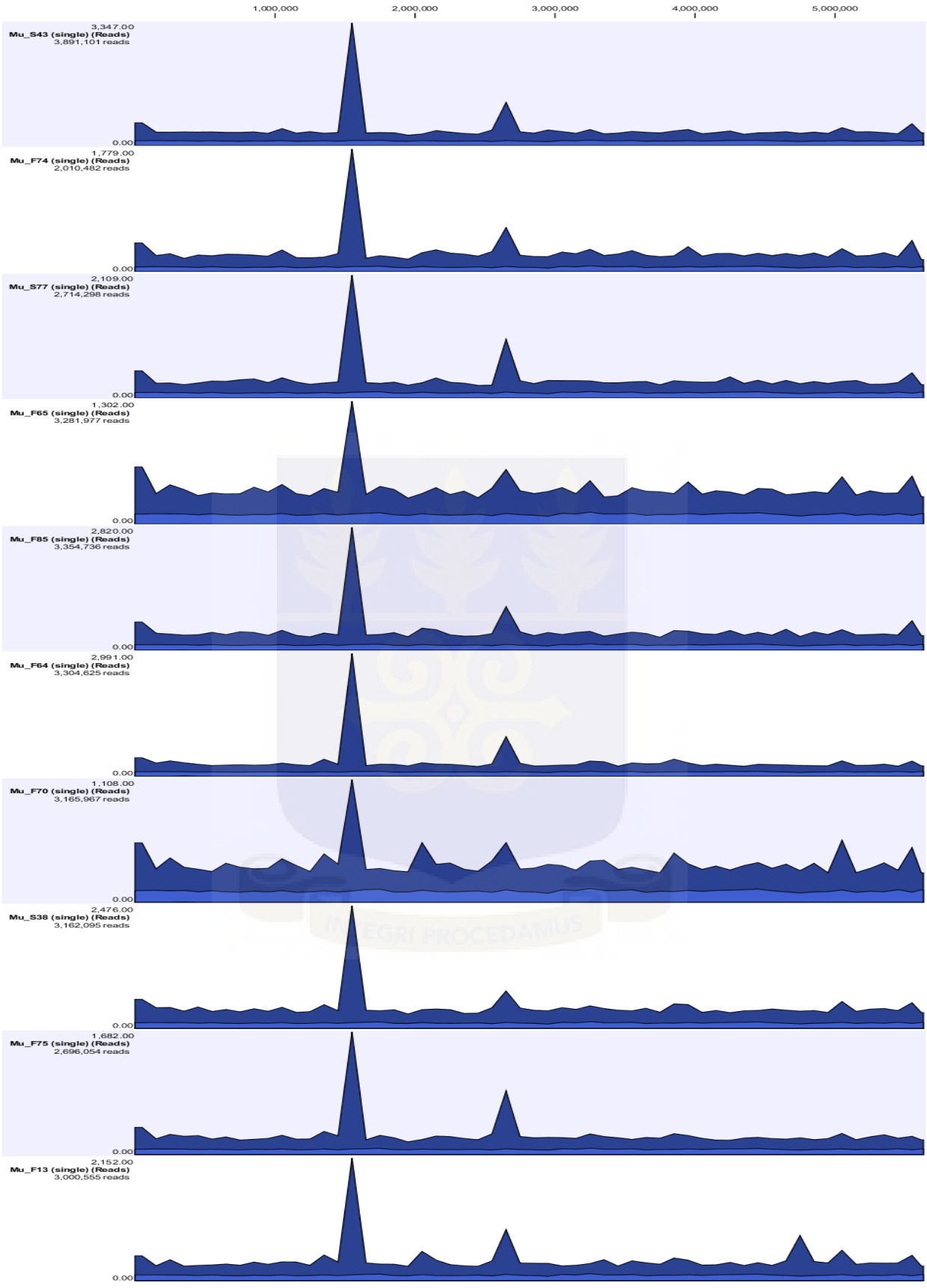


Fig. 4.1: Single read coverage for *M. ulcerans* isolates from Asante Akim North District

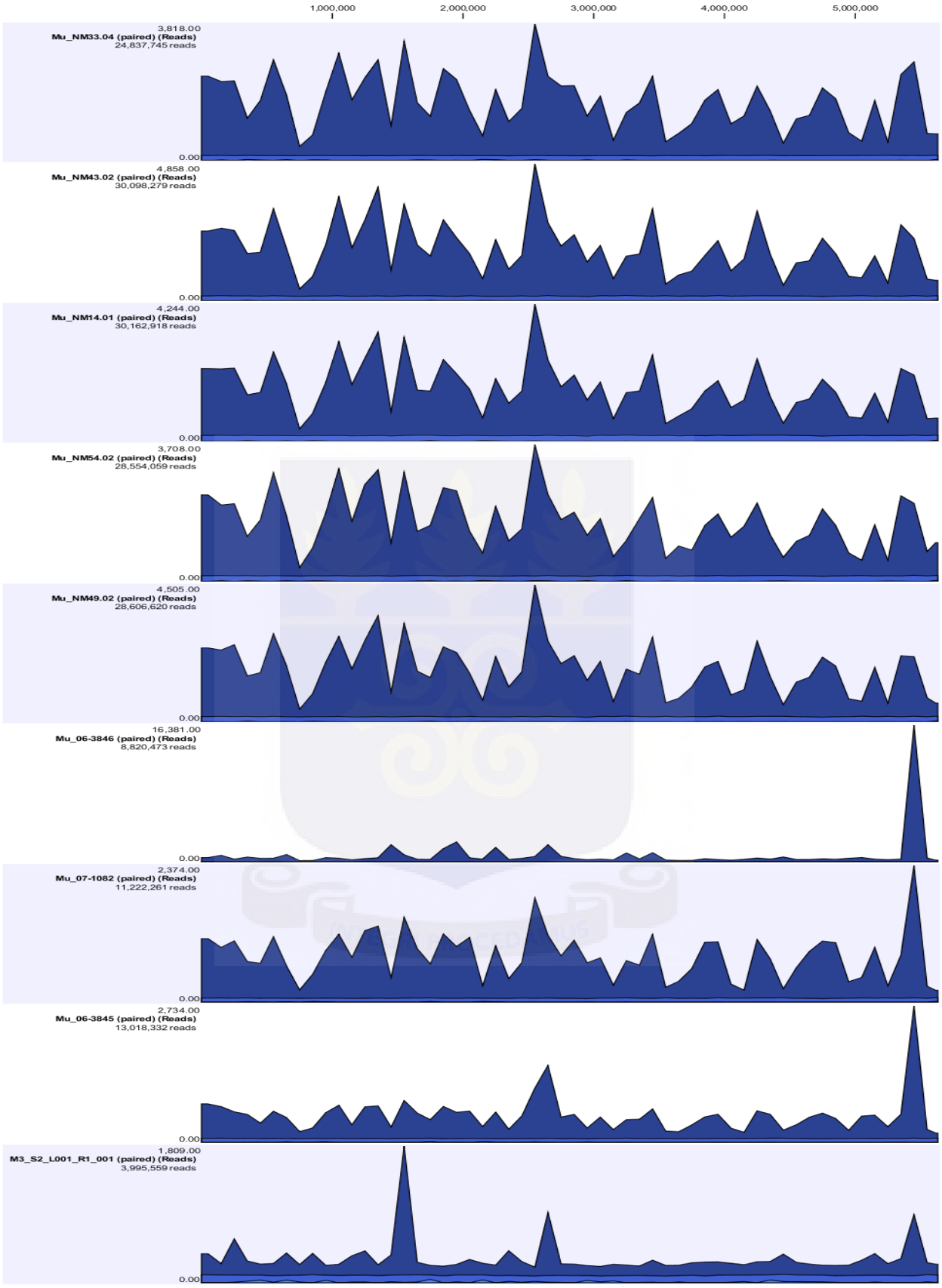


Fig. 4.2: Paired read coverage for *M. ulcerans* isolates from Benin, Amansie West and Ga Districts

4.13 Distribution of variants in *M. ulcerans* genome

The extent of distribution of variants across the entire genomes of the various *M. ulcerans* isolates compared to the reference isolate (Agy99) are shown in Figures 4.3 and 4.4. The first five isolates (Mu_S43 to Mu_F85) in Figure 4.3 have very similar pattern of variant distribution in their genomes, reflecting their close genetic relatedness. Single nucleotide polymorphisms (SNPs) identified in all MU isolates are shown in Table 4.3. Figure 4.5 shows SNPs obtained for the various isolates from concatenated sequences of the 299 SNP positions in the reference genome.

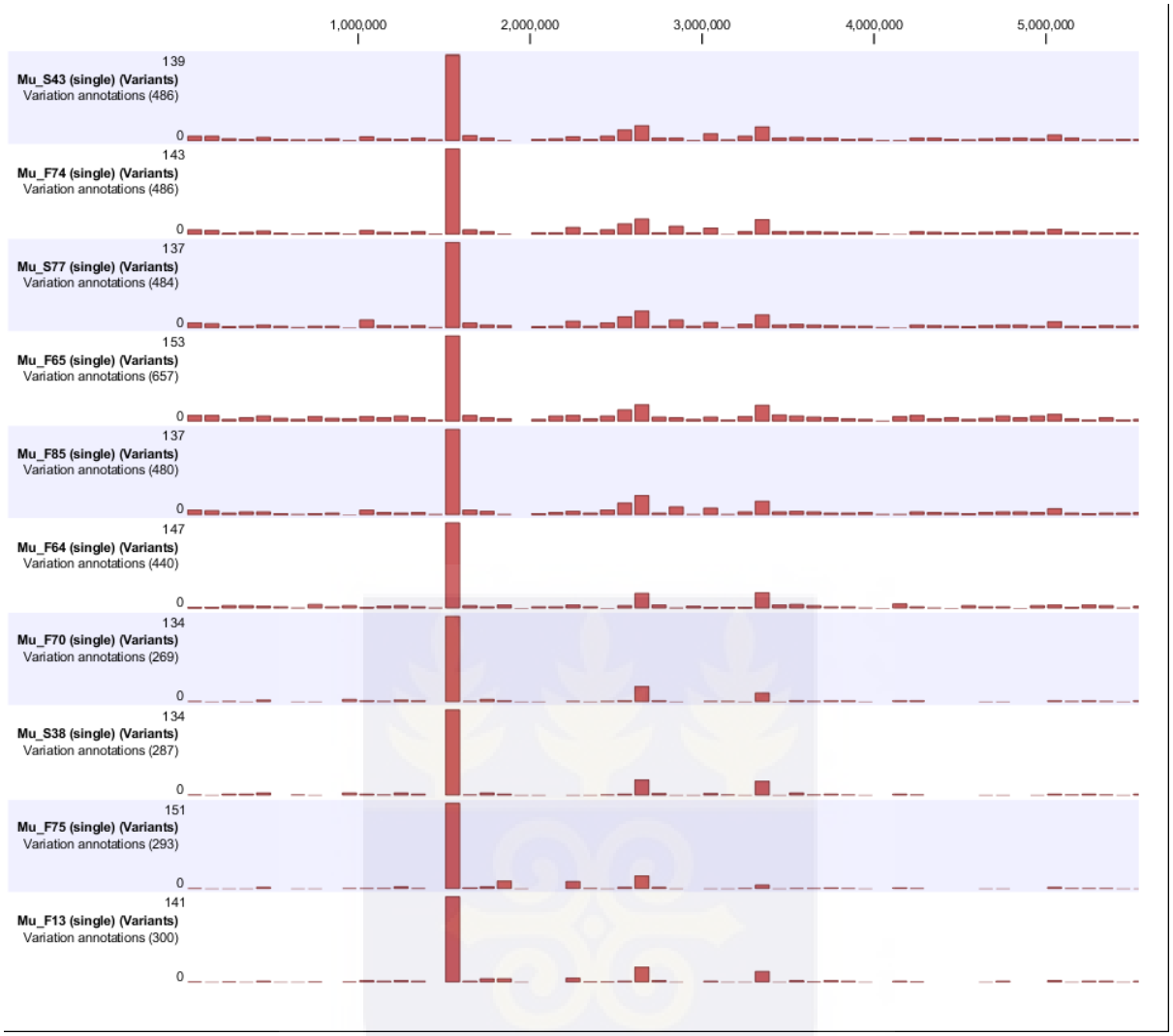


Fig. 4.3: Regions in *M. ulcerans* genome showing the extent of variations for single reads compared to the reference isolate

Fig. 4.4: Regions in *M. ulcerans* genome showing the extent of variations for paired reads compared to the reference isolate

Table 4.3 Single nucleotide polymorphisms differentiating two major *M. ulcerans* clusters in the AAND of Ghana

Position	**Agy99**	B	F64	F75	F13	S38	F70	D	M3	E	F	H	3846	F65	F74	F85	S77	S43	1082	3845
5217	C	C	T	T	T	T	T	C	C	C	C	C	C	C	C	C	C	C	C	C
34246	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	C	C
34723	T	T	T	T	T	T	T	T	T	C	C	C	T	T	T	T	T	T	T	T
43972	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C
56032	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T
77657	C	C	C	C	C	C	C	C	C	T	T	T	C	C	C	C	C	C	C	C
111940	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A
133648	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G
153193	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G
163740	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C
165146	G	G	G	G	G	G	G	G	G	G	G	G	G	C	C	C	C	C	G	G
179989	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	C	C
187847	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C	C	C	C	T	T
275227	A	A	A	A	A	A	A	G	G	G	G	G	A	A	A	A	A	A	A	A
286692	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C
343811	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C
366009	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C
376413	G	G	G	G	G	G	G	G	G	A	A	A	G	G	G	G	G	G	G	G
446532	T	T	C	C	C	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T
468708	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A
479253	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
482675	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C
487781	A	A	A	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A	A	A
495906	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C
508649	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	C	C
604140	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T
642364	T	T	T	T	T	T	T	C	C	C	C	C	C	C	C	C	C	C	C	C
686169	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A
806478	C	C	C	C	C	C	C	C	C	C	C	C	C	G	G	G	G	G	C	C

837599	C	C	C	C	C	C	C	T	T	T	T	T	T	C	C	C	C	C	C	C
843298	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	T
844127	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	C	C	C
886275	C	C	C	C	C	C	C	T	T	T	T	T	C	C	C	C	C	C	C	C
971225	C	C	T	T	T	T	T	C	C	C	C	C	C	C	C	C	C	C	C	C
982701	A	A	A	A	A	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A
995111	C	C	C	C	C	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C
1048338	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C	C	C	C	C	C
1061028	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T
1068962	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C	C	C	C	T	T
1090880	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	C	C
1091862	A	A	A	A	A	A	A	C	C	C	C	C	A	A	A	A	A	A	A	A
1096605	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C
1140497	G	G	G	G	G	G	G	G	G	G	T	G	G	G	G	G	G	G	G	G
1177120	A	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
1238622	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1295651	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C
1321620	G	G	G	G	G	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C
1324794	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	G
1331384	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C	C	C	C	T	T
1363462	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	A
1401002	A	A	A	A	A	A	A	G	G	G	G	G	A	A	A	A	A	A	A	A
1403655	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C
1412603	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G
1425859	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T
1483023	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C
1529538	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	T	T
1559039	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	G	G
1596932	A	A	G	G	G	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A
1608301	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C
1620521	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G	G	G	G	A	A
1623795	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A

1623967	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G
1631658	C	C	C	C	C	C	C	T	T	T	T	T	C	C	C	C	C	C	C	C
1634826	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
1644563	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
1660610	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
1667158	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G	G	G	G	A	A
1674860	T	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
1681791	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C
1685280	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	A	A
1706282	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T
1711432	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C
1718395	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C
1788572	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T
1789257	C	C	C	C	C	C	C	C	C	C	C	C	C	G	G	G	G	G	G	G
1877820	C	C	C	C	C	T	T	C	C	C	C	C	C	C	C	C	C	C	C	C
1882597	C	C	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C
1899339	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C
1906107	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C
1909275	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A
1921818	A	A	C	C	C	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A
1953705	C	C	C	C	C	C	C	C	C	T	T	T	C	C	C	C	C	C	C	C
1958865	T	T	T	T	T	T	T	C	C	C	C	C	T	T	T	T	T	T	T	T
1960084	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
1982139	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T
2013818	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	T
2056416	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G
2073675	A	A	A	A	A	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A
2088885	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T
2122245	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	T	T
2153815	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G
2157740	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	A	
2168348	C	C	C	C	C	C	C	C	C	C	C	C	C	A	A	A	A	A	A	A

2211536	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T
2250462	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
2262196	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	T	T
2283073	G	G	G	G	G	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A
2304270	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G
2348892	C	C	C	C	C	C	C	C	C	A	A	A	C	C	C	C	C	C	C	C
2368411	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T
2433592	A	A	A	A	A	A	A	G	G	G	G	G	A	A	A	A	A	A	A	A
2443453	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T
2455123	T	T	T	T	T	T	T	T	T	C	C	T	T	T	T	T	T	T	T	T
2460832	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C
2485389	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G
2492682	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C
2496544	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	T	T
2497586	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	G	G
2500817	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
2528399	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A
2536596	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	C	C
2564007	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G	G	G	G	A	A
2566767	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C
2580522	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	T	T
2580527	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T	T	T	T	T	T
2586113	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	C
2594215	C	C	C	C	C	C	C	G	G	G	G	G	G	G	G	G	G	G	G	G
2646800	A	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A
2691935	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	C
2707658	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G
2737446	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
2750738	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
2885754	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A
2926056	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G
3040231	T	T	T	T	T	T	T	T	T	T	T	T	C	T	T	T	T	T	T	T

3042790	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C
3076035	G	G	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G
3095093	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	A
3216548	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	G	G	
3224762	G	G	G	G	G	G	G	A	A	A	A	A	G	G	G	G	G	G	G	G	
3228996	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	A	A	
3248025	G	G	G	G	G	C	C	G	G	G	G	G	G	G	G	G	G	G	G	G	
3250091	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G	G	G	G	A	A	
3298013	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	T	
3301651	C	C	C	C	C	C	C	C	C	C	G	C	C	C	C	C	C	C	C	C	
3317003	C	C	C	C	C	C	C	T	T	T	T	T	C	C	C	C	C	C	C	C	
3322137	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	A	
3402081	T	T	T	T	T	T	T	G	G	G	G	G	G	G	G	G	G	G	G	G	
3443738	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	
3457302	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	G	G	
3493181	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	G	G	
3589132	A	A	A	A	A	A	A	G	G	G	G	G	A	A	A	A	A	A	A	A	
3590442	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	
3594809	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	
3639112	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	G	
3690405	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
3700529	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	
3715145	T	T	T	T	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	
3786844	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	G	G	
3813515	G	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
3827482	C	C	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	
3847350	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G	
3857417	A	A	A	A	A	A	A	A	A	A	A	A	C	A	A	A	A	A	A	A	
3862080	G	G	G	A	A	A	A	G	G	G	G	G	G	G	G	G	G	G	G	G	
3909489	C	C	C	C	C	C	C	T	T	T	T	T	C	C	C	C	C	C	C	C	
3910253	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	A	A	
4027569	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	A	

4048169	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C
4099018	A	A	A	A	A	A	A	A	A	A	A	A	A	C	C	C	C	C	A	A
4102801	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G	G	G	G	G	G
4120933	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C
4169725	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G
4205303	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G
4206998	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C
4214955	C	C	C	C	C	C	C	C	C	C	C	C	C	A	A	A	A	A	A	A
4234218	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C	C	C	C	T	T
4239479	G	G	G	G	G	G	G	T	T	T	T	T	G	G	G	G	G	G	G	G
4239480	A	A	A	A	A	A	A	T	T	T	T	T	A	A	A	A	A	A	A	A
4240215	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	C
4248562	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G
4278153	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G
4280148	A	A	A	A	A	A	A	G	A	G	G	G	A	A	A	A	A	A	A	A
4326987	A	A	A	A	A	A	A	T	T	T	T	T	T	A	A	A	A	A	A	A
4339286	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A
4368716	G	G	G	G	G	G	G	C	G	G	G	G	G	G	G	G	G	G	G	G
4399434	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T
4423911	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G	G	G	G	A	A
4498200	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G
4498560	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T	T
4514887	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G
4597138	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A
4597315	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	A	A
4606732	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G	G	G
4639498	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A
4643971	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G	G
4656740	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	G	G
4666016	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G
4683466	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C
4721897	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G

4770743	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
4788986	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C
4842750	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	T	T
4870070	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A	A
4895058	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C
4921425	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C
5004495	T	T	T	T	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T
5029290	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
5038962	A	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
5040698	T	T	T	T	G	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
5044283	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	T
5094686	C	C	C	C	C	C	C	C	C	C	C	C	T	C	C	C	C	C	C	C
5119098	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	G	G
5149246	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G
5155941	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A
5245228	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C
5286908	G	G	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
5308823	C	C	C	C	T	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C
5315773	A	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A
5321488	T	T	T	T	T	T	T	C	C	C	C	C	C	T	T	T	T	T	T	T
5321517	T	T	T	T	T	T	T	C	C	C	C	C	C	T	T	T	T	T	T	T
5324878	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G
5346220	G	G	T	T	T	T	T	G	G	G	G	G	G	G	G	G	G	G	G	G
5346604	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	C	C
5382178	C	C	C	C	C	C	C	C	C	T	T	T	C	C	C	C	C	C	C	C
5388859	A	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G
5395961	C	C	C	C	C	C	C	C	C	C	C	C	A	C	C	C	C	C	C	C
5420373	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A	A	A	A	G	G
5460514	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G
5479035	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	T
5499366	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A
5516965	A	A	A	A	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A

5537763	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	A
5562616	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	C
5590526	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G
5593884	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	G	G
5602537	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	G	A	G
5602631	G	G	G	G	G	G	G	G	G	G	G	G	G	C	C	C	C	C	G	G
5603342	A	A	A	A	A	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A
5606137	C	C	C	C	C	C	C	G	C	G	G	G	C	C	C	C	C	C	C	C
5607063	A	A	A	A	A	A	A	A	G	A	A	A	A	A	A	A	A	A	A	A
5610402	G	G	G	G	G	G	G	G	G	G	G	G	A	G	G	G	G	G	G	G
5611273	T	T	T	T	T	T	T	T	T	T	T	T	T	C	C	C	C	C	C	C
5611985	C	C	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T
5618159	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	C

Single Nucleotide polymorphisms are indicated in green highlights. **B**, Mu_NM33.04; **D**, Mu_NM43.02; **E**, Mu_NM49.02; **F**, Mu_NM14.01; **H**, Mu_NM54.02

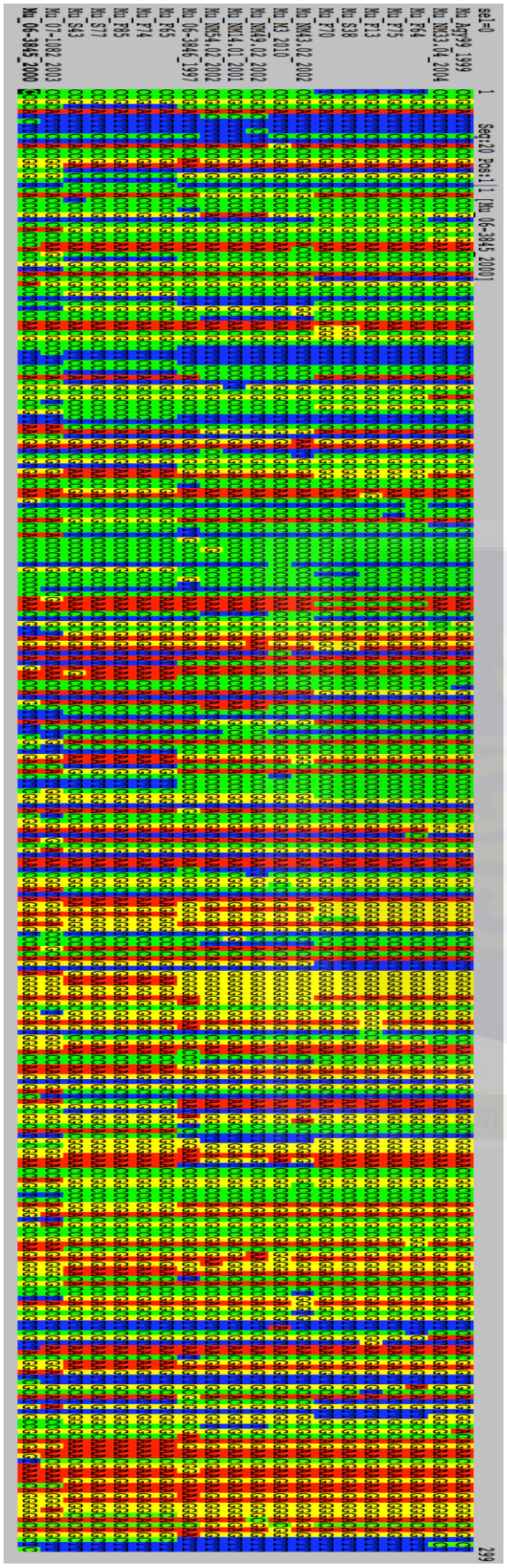


Fig. 4.5: Concatenated sequences of *M. ulcerans* isolates with SNPs compared to the reference isolate

4.14 Phylogenetic relationship among *M. ulcerans* isolates

Phylogenetic relationship among the *M. ulcerans* isolates was inferred using unrooted neighbour joining tree (Fig. 4.6) produced by SplitsTree4 (Huson, 1998) using uncorrected 'P' distances and based on the 299 SNP positions observed when compared with the reference genome. Five different *M. ulcerans* clusters shown in Figure 4.6 were produced; two for the isolates from the Asante Akim North District (AAND) and one each for the isolates from other endemic areas studied. Isolates from AAND formed two distinct clusters. One cluster wherein referred to as the tight cluster consists of four isolates with identical SNP profiles across their entire genome and another isolate (Mu_S43) that differs from these members by only 2 SNPs as shown in Figure 4.6. To evaluate the reliability of the inferred trees, bootstrap analysis was carried out based on 1000 bootstrap replicates with clustal bootstrap values ranging from 98.6 to 100 (Fig. 4.6).

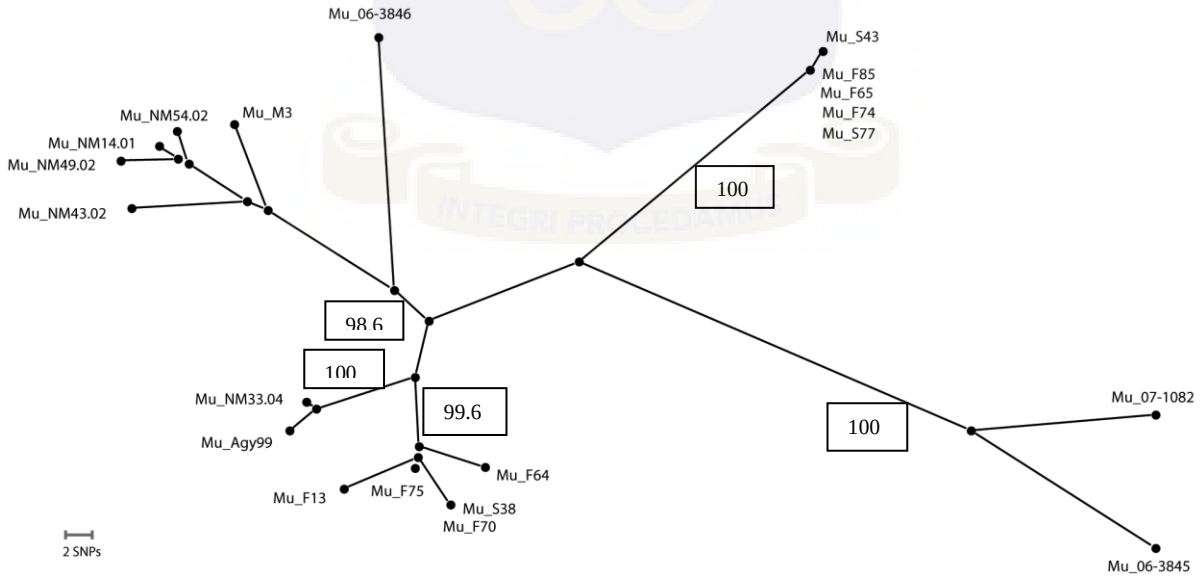


Fig. 4.6: Unrooted neighbour joining tree showing phylogenetic relationship among *M. ulcerans* isolates

4.15 SNPs detection in *M. ulcerans* coding sequences

Table 4.3 describes single nucleotide polymorphisms (SNPs) identified in coding sequences in *M. ulcerans*. The coloured information in the table is unique to the corresponding isolates. Generally, the isolates in the same cluster have more common SNPs in their coding sequences. The insect isolate from the Asante Akim North District, Mu_M3, that clustered with the Ga District isolates shares a lot of SNPs with them (Table 4.4). Similar observation was made in the insect isolate from Benin (Mu_06-3845) identified by Portaels and colleagues (2008) and the human isolate, Mu_07-1082, it clusters with. Out of the number of SNPs in coding sequences identified, about one-third (58) was synonymous SNPs.

Table 4.4: Single nucleotide polymorphisms identified in coding sequences in *M. ulcerans*

Isolate	Genes with SNPs			
	Exact position in genome (bp)	**Nature of sequence**	**Base (codon)**	**Coding enzyme (nucleotide change in gene)**
Mu_NM33.04	1959260..1960300	CDS syn.	825 (275)	Conserved lipoprotein LprF (G → A)
	2245103..2250523	CDS	62 (21)	Phenolpthiocerol synthesis type-I polyketide synthase PpsD (T → C)
Mu_NM43.02 only	487315..488535	CDS	755 (252)	Cytochrome P450 123A3Cyp123A3 (A → C)
	1482868..1483359	CDS syn.	156 (52)	Tuberculin related peptide (C→ T)
	2521286..2533861	CDS syn.	5463 (1821)	Polyketide synthase; Pks12 (A → G)
	4204700..4205953	CDS syn.	651 (217)	Carboxylesterase LipQ (G → A)
	4367925..4370252	CDS	1537 (513)	Pyruvate dehydrogenase E1 component AceE_1 (G → C)
	4788762..4789133	CDS syn.	225 (75)	Conserved hypothetical protein (C → T)
	4868465..4870312	CDS syn.	243 (81)	Hydrolase (A→ G)
	4920725..4921600	CDS	701 (234)	Ku domain DNA repair protein (C → T)
Mu_NM43.02	77264..77878	CDS syn.	222 (74)	Transcriptional regulator (C → T)
Mu_NM14.01	273672..275357	CDS	131 (44)	NADPH: adrenodoxin oxidoreductase FprB (A → G)
Mu_NM54.02	375984..376421	CDS syn.	9 (3)	Methylmalonyl-CoA mutase alpha subunit; McmA2b (G → A)
Mu_NM49.02				
Mu_M3	642133..642480	CDS	232 (78)	*Conserved hypothetical protein (T → C)
	837503..838177	CDS syn.	97 (33)	*50S ribosomal protein L4; RplD (C → T)
	885499..886281	CDS syn.	777 (259)	Pseudouridine synthase a; TruA (C → T)
Mu_NM14.01 only	1140176..1142251	CDS	322 (108)	Beta-glucosidase BglS (G → T)
	3300508..3301824	CDS	1144 (382)	Cytochrome P450 140A5 Cyp140A5

				(C → G)
Mu_NM14.01, Mu_M3, Mu_06-3846	3587910..3589535	CDS	404 (135)	P-loop ATPase (A → G)
	1238239..1239669	CDS	1048 (350)	*Conserved hypothetical membrane protein (C → T)
	1400702..1401955	CDS syn.	954 (318)	Beta-ketoacyl-acyl-carrier-protein synthase I (A → G)
	1629912..1632173	CDS	1747 (583)	Methylmalonyl-CoA mutase large subunit; MutB (C → T)
	1673353..1675179	CDS	1508 (503)	*Fatty-acid-CoA ligase; FadD11 (T → C)
	1957808..1959130	CDS	266 (89)	Beta-lactamase (T → C)
Mu_NM49.02 only	2056197..2057207	CDS syn.	792 (264)	Quinone reductase Qor (G → A)
	3042670..3043665	CDS	121 (41)	Hydrolase (C → T)
	4606646..4607257	CDS syn.	87 (29)	Conserved hypothetical protein (G → A)
	2245103..2250523	CDS	62 (21)	*Phenolpthiocerol synthesis type-I polyketide synthase PpsD (T → C)
	2592553..2595147	CDS	933 (311)	Conserved alanine and proline rich protein (C → T)
	3224434..3224817	CDS	56 (19)	Hypothetical protein (G → A)
	3401365..3402327	CDS	247 (83)	*Membrane-bound C-5 sterol desaturase Erg3 (T → G)
	3826715..3827650	CDS	169 (57)	Citrate lyase beta subunit; CitE_2 (C → G)
	3909328..3910338	CDS	850 (284)	Diphosphomevalonate decarboxylase (C → T)
	4238681..4239736	CDS	799 (267)	Mce family protein; Mce5C (G → T)
	4238681..4239736	CDS	800 (267)	Mce family protein; Mce5C (A → T)
	5038484..5039350	CDS	479 (160)	*3-hydroxybutyryl-CoA dehydrogenase FadB2 (A → G)
	5611592..5613037	CDS syn.	1053 (351)	*Poly(a) polymerase PcnA (C → T)
Mu_NM14.01	34213..34860	CDS syn.	138 (46)	ABC transporter permease (T → C)
Mu_NM54.02	5382158..5386408	CDS syn.	21 (7)	Non-ribosomal peptide synthetase (C → T)

Mu_NM49.02				
Mu_M3	111857..112735	CDS syn.	796 (266)	Membrane oxidoreductase (A → G)
	1787076..1788920	CDS syn.	1497 (499)	PE-PGRS family protein (T → C)
	2087969..2089690	CDS	917 (306)	Cytochrome C oxidase polypeptide I CtaD (T → C)
	2566355..2566786	CDS	20 (7)	Conserved membrane transport protein; MmpS4_1 (C → T)
	4639027..4640340	CDS	472 (158)	Conserved hypothetical lipoprotein LpqF (A → G)
	5606955..5607227	CDS syn.	165 (55)	EsaT-6 like protein EsxE (C → T)
Mu_06-3846	365985..366284	CDS syn.	276 (92)	PE-PGRS family protein family protein (C → T)
	837503..838177	CDS syn.	97 (33)	50S ribosomal protein L4; RplD (C → T)
	1608253..1608831	CDS	49 (17)	Conserved membrane protein (C → T)
	1681113..1682876	CDS syn.	1086 (362)	Maltooligosyltrehalose trehalohydrolase; TreZ (C → T)
	1881796..1883049	CDS	453 (151)	Acyl-CoA dehydrogenase (C → G)
	1898764..1902288	CDS	2950 (984)	Fatty-acid-CoA ligase FadD9 (C → T)
	2645861..2646982	CDS syn.	183 (61)	FAD-dependent oxidoreductase (A → G)
	3040182..3040781	CDS	551 (184)	Conserved protein (T → C)
	3700513..3701718	CDS	1190 (397)	SAM-dependent methyltransferase (G → A)
	3857061..3857549	CDS	133 (45)	Conserved hypothetical secreted protein (A → C)
	4248045..4249784	CDS	518 (173)	Conserved hypothetical membrane protein (G → A)
	4682330..4684114	CDS syn.	1137 (379)	Conserved hypothetical membrane protein (C → T)
	5094417..5095199	CDS	514 (172)	conserved hypothetical alanine and proline rich protein (C → T)
Mu_07-1082 and	153112..154194	CDS syn.	1002 (334)	6- phosphogluconate dehydrogenase; decarboxylating Gnd2 (A → G)
Mu_06-3845	163531..163977	CDS	238 (80)	Conserved protein (T → C)

	343471..344274	CDS	464 (155	Short chain dehydrogenase (T→ C)
	602637..604967	CDS syn.	828 (276)	Aerobic-type carbon monoxide dehydrogenase; CoxL_2 (C → T)
	642133..642480	CDS	232 (78)	Conserved hypothetical protein (T → C)
	1048174..1048725	CDS	388 (130)	Conserved hypothetical protein (T → C)
	1238239..1239669	CDS	1048 (350)	Conserved hypothetical membrane protein (C → T)
	1529017..1531545	CDS	2008 (670)	Conserved transmembrane ATP-binding protein ABC transporter (C → T)
	1673353..1675179	CDS	1508 (503)	Fatty-acid-CoA ligase; FadD11 (T → C)
	1685174..1687351	CDS	2072 (691)	Maltooligosyltrehalose synthase TreX (T → A)
	1706229..1706861	CDS syn.	54 (18)	Imidazoleglycerol-phosphate dehydratase (C → T)
	2121370..2122614	CDS	370 (124)	Transferase (C → T)
	2245103..2250523	CDS	62 (21)	Phenolpthiocerol synthesis type-I polyketide synthase PpsD (T → C)
	2261702..2266510	CDS	4315 (1439)	Phenolpthiocerol synthesis type-I polyketide synthase PpsA (C → T)
	2367598..2368641	CDS syn.	814 (272)	N utilization substance protein a NusA (C → T)
	2592553..2595147	CDS	933 (311)	Conserved alanine and proline rich protein (C → G)
	3094563..3095195	CDS syn.	531 (177)	Conserved hypothetical protein (C → A)
	3228218..3229168	CDS	779 (260)	Transcriptional regulatory protein (G → A)
	3401365..3402327	CDS	247 (83)	Membrane-bound C-5 sterol desaturase Erg3 (T → G)
	3443393..3443824	CDS	346 (116)	Conserved hypothetical protein (G → A)
	3826715..3827650	CDS	169 (57)	Citrate lyase beta subunit; CitE_2 (C → G)
	3845991..3847356		1360	Repeat region
	3909328..3910338	CDS	86 (29)	Diphosphomevalonate decarboxylase (G → A)
	4027368..4027955	CDS	202 (68)	Conserved hypothetical protein

				(C → A)
	4102054..4102830	CDS	748 (250)	Conserved hypothetical membrane protein (A → G)
	4338797..4339441	CDS syn.	156 (52)	Conserved hypothetical protein (G → A)
	4398723..4399904	CDS	471 (157)	Acetyl-CoA acetyltransferase FadA4 (C → T)
	4498374..4498799	CDS syn.	40 (80)	Conserved hypothetical transcriptional regulator (Lrp family) (C → T)
	4665975..4667657	CDS syn.	42 (14)	Monooxygenase (A → G)
	4721438..4722139	CDS syn.	243 (81)	endonuclease III Nth (A → G)
	4842660..4843931	CDS	1182 (394)	Conserved hypothetical protein (G → T)
	5038484..5039350	CDS	479 160	3-hydroxybutyryl-CoA dehydrogenase FadB2 (A → G)
	5155581..5156147	CDS	361 (121)	Serine acetyltransferase CysE_1 (G → A)
	5536907..5538826	CDS	1064 (355)	Conserved transmembrane protein (G → A)
	5611592..5613037	CDS syn.	1053 (351)	Poly(a) polymerase PcnA (C → T)
Mu_07-1082	1059311..1061197	CDS	1718 (573)	GTP-binding translation elongation factor TypA (T → C)
	1294486..1296087	CDS	1166 (389)	Conserved membrane protein (C → T)
	1907691..1909403	CDS	1585 (529)	ABC-type dipeptide transport system; periplasmic component (A → G)
	2691479..2692168	CDS syn.	234 (78)	Cell division ATP-binding protein FtsE (C → T)
	3297431..3298039	CDS	583 (195)	Conserved secreted protein (T→ C)
	3321506..3322858	CDS	632 (211)	Alpha amylase (A→ G)
	3638719..3639144	CDS syn.	33 (11)	Conserved hypothetical membrane protein (G → T)
	4514311..4515501	CDS syn.	615 (205)	8-amino-7-oxononanoate synthase BioF2 (G → A)
	5323534..5325282	CDS syn.	405 (135)	Divalent cation-transport integral membrane protein (G → A)
Mu_06-3845	43716..44120	CDS	257 (86)	Conserved hypothetical protein (T → C)
	482341..482760	CDS	86 (29)	Conserved hypothetical protein

				(A → C)
	685764..686771	CDS	406 (136)	Polyprenyl diphosphate synthetase; GrcC1 (G → A)
	1623755..1624498	CDS syn.	213 (71)	Two component regulator - receiver domain (C → G)
	1904493..1906340	CDS	1615 (539)	Protein-export membrane protein SecD (T → C)
	2152214..2154070	CDS syn.	1602 (534)	Acetolactate synthase (A → G)
	2303799..2305157	CDS syn.	888 (296)	Conserved hypothetical protein (A → G)
	2443441..2445378	CDS syn.	1926 (642)	Glycine betaine transport integral membrane protein BetP (C → T)
	2707280..2707930	CDS syn.	273 (91)	Two component transcriptional regulatory protein DevR (C → G)
	4045818..4048892	CDS syn.	2352 (784)	Polyketide synthase MbtD (T → C)
	5244936..5245406	CDS	179 (60)	Conserved protein (T → C)
	5459347..5460855	CDS syn.	342 (114)	Glutamate decarboxylase (A → G)
	5616444..5619986	CDS syn.	1716 (572)	Conserved transmembrane protein (T → C)
Mu_S43, Mu_F74, Mu_S77, Mu_F65, Mu_F85	34213..34860	CDS syn.	615 (205)	ABC transporter permease (C → T)
	179542..180699	CDS syn.	448 (150)	Conserved hypothetical transmembrane protein (C → T)
	187386..188261	CDS	415 (139)	Glycosyl hydrolase (T → C)
Mu_S43	286293..288245	CDS syn.	1554 (518)	Acyl-CoA dehydrogenase FadE10 (C → T)
	508010..509218	CDS	640 (214)	PPE family protein (C → T)
	642133..642480	CDS	232 (78)	Conserved hypothetical protein (T → C)
	1048174..1048725	CDS	388 (130)	Conserved hypothetical protein (T → C)
	1238239..1239669	CDS	1048 (350)	Conserved hypothetical membrane protein (C → T)
	1529017..1531545	CDS	2008 (670)	Conserved transmembrane ATP-binding protein ABC transporter (C → T)
	1559007..1560356	CDS	base 1318 codon 440	Sugar-transport integral membrane protein SugI (G → A)
	1673353..1675179	CDS	1508 (503)	Fatty-acid-CoA ligase; FadD11

				(T → C)
	2121370..2122614	CDS	370 (124)	Transferase (C → T)
	2245103..2250523	CDS	62 (21)	Phenolpthiocerol synthesis type-I polyketide synthase PpsD (T → C)
	2261702..2266510	CDS	4315 (1439)	Phenolpthiocerol synthesis type-I polyketide synthase PpsA (C → T)
	2496581..2497867	CDS syn.	282 (94)	PPE family protein (G → A)
	2535896..2537713	CDS	1118 (373)	N-term polyprenol-monophosphomannose synthase Ppm1A (C → T)
	2592553..2595147	CDS	933 (311)	Conserved alanine and proline rich protein (C → T)
	3228218..3229168	CDS	779 (260)	Transcriptional regulatory protein (G → A)
	3249943..3250926	CDS	836 (279)	D-amino acid oxidase Aao (A→ G)
	3401365..3402327	CDS	247 (83)	Membrane-bound C-5 sterol desaturase Erg3 (T → G)
	3492744..3493568	CDS	388 (130)	Ribulose-5-phosphate 4-epimerase; AraD (G → A)
	3826715..3827650	CDS	169 (57)	Citrate lyase beta subunit; CitE_2 (C → G)
	3909328..3910338	CDS	86 (29)	Diphosphomevalonate decarboxylase (G → A)
	4098355..4099455	CDS	664 (222)	Conserved hypothetical membrane protein (A → C)
	4102054..4102830	CDS	748 (250)	Conserved hypothetical membrane protein (A → G)
	4423569..4424081	CDS syn.	171 (57)	H(+)-transporting two-sector ATPase; B subunit (A → G)
	4656271..4657218	CDS	479 (160)	Pantoate--beta-alanine ligase PanC (G → A)
	5038484..5039350	CDS	479 (160)	3-hydroxybutyryl-CoA dehydrogenase FadB2 (A → G)
	5118177..5119706	CDS syn.	609 (203)	Conserved hypothetical transmembrane protein (G → A)
	5419621..5420646	CDS syn.	753 (251)	Conserved hypothetical protein (G → A)
	5611592..5613037	CDS syn.	1053 (351)	Poly(a) polymerase PcnA

				(C → T)
Mu_F75, Mu_F13,	446048..446860	CDS	329 (110)	Conserved protein (T → C)
Mu_S38, Mu_F64,	1238239..1239669	CDS	1048 (350)	Conserved hypothetical membrane protein (C → T)
Mu_F70	1592995..1597203	CDS	272 (91)	Formate dehydrogenase H FdhF (A → G)
Mu_F75 only	1660244..1660816	CDS syn.	207 (69)	Conserved hypothetical protein (C → T)
	1673353..1675179	CDS	1508 (503)	Fatty-acid-CoA ligase; FadD11 (T → C)
	2245103..2250523	CDS	62 (21)	Phenolpthiocerol synthesis type-I polyketide synthase PpsD (T → C)
	3826715..3827650	CDS	169 (57)	Citrate lyase beta subunit; CitE_2 (C → G)
	5038484..5039350	CDS	479 (160)	3-hydroxybutyryl-CoA dehydrogenase FadB2 (A → G)
	5346188..5346394	CDS syn.	33 (11)	Conserved protein (G → T)
	5611592..5613037	CDS syn.	1053 (351)	Poly(a) polymerase PcnA (C → T)
Mu_F70, Mu_S38	982609..983778	CDS	1078 (360)	Keto acyl-CoA thiolase; Ltp2 (A → G)
	995097..996380	CDS syn.	15 (5)	Conserved integral membrane transport protein (C → G)
	2073553..2074737	CDS	1063 (355)	Sulfate adenylyltransferase (A → G)
Mu_F13	1623755..1624498	CDS	41 (14)	Two component regulator - receiver domain(A → G)
	5040578..5041441	CDS	744 (248)	Mycolic acid synthase PcaA (T → G)
Mu_F64	1634621..1634833	CDS	8 (3)	Conserved hypothetical protein (T → C)
	1643250..1644687		1314	repeat_region (T → C)
	2750671..2751426	CDS	68 (23)	NADH dehydrogenase I (chain E) NuoE (NADH-ubiquinone oxidoreductase chain E) (T → C)
	5286901..5287281	CDS	374 (125)	Conserved protein (G → A)

4.20 Description of results for VNTR study

Out of the twenty five (25) newly-described VNTR loci tested in this study, DNA amplification was observed in twenty three (23, 92%) of them. There was no DNA amplification at two of the loci including locus 16. Fifteen (60%) of the VNTR loci were amplified in 14 (82%) of the isolates. The DNA band sizes observed for all loci appear to range between 300 bp and 750 bp. However, almost all isolates for each locus have similar band sizes with few exceptions (S8 and S43 at locus 5; F41and F74 at locus 9; S8, S21 and F45 at locus 15) that appear to differ in band sizes. Figures 4.70 to 4.76 show VNTR PCR results for *M. ulcerans* isolates from the three Buruli ulcer endemic areas studied. Variation of electrophoretic conditions for the separation of samples that appeared to differ in band sizes produced no differences for loci 5, 9 and 15 with estimated band sizes of 600 bp, 500 bp and 350 bp respectively (Fig. 4.77).

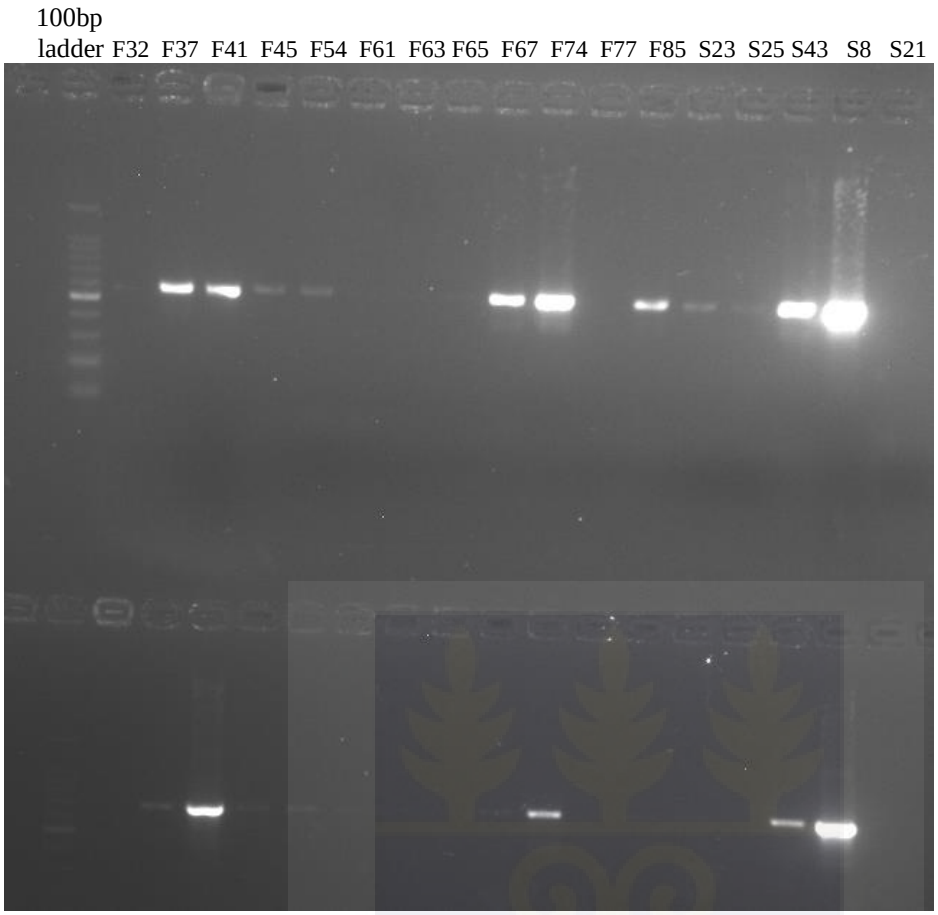


Fig. 4.70: PCR analysis of VNTR loci 3 and 5 for *M. ulcerans* isolates (F32 to S21)

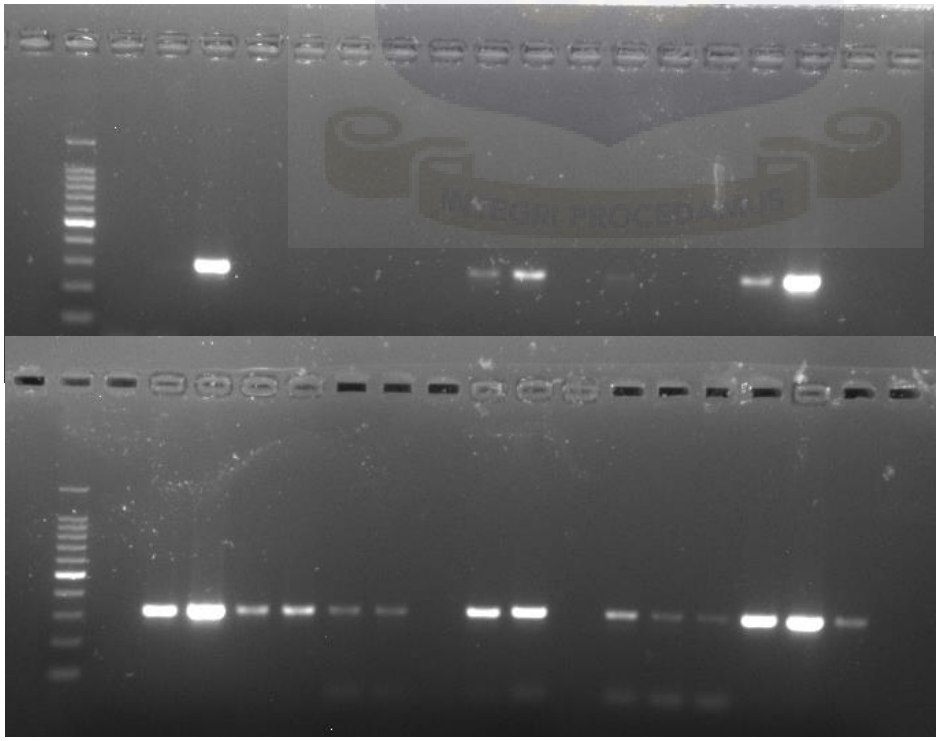


Fig. 4.71: PCR analysis of VNTR loci 7 and 8 for *M. ulcerans* isolates (F32 to S21)

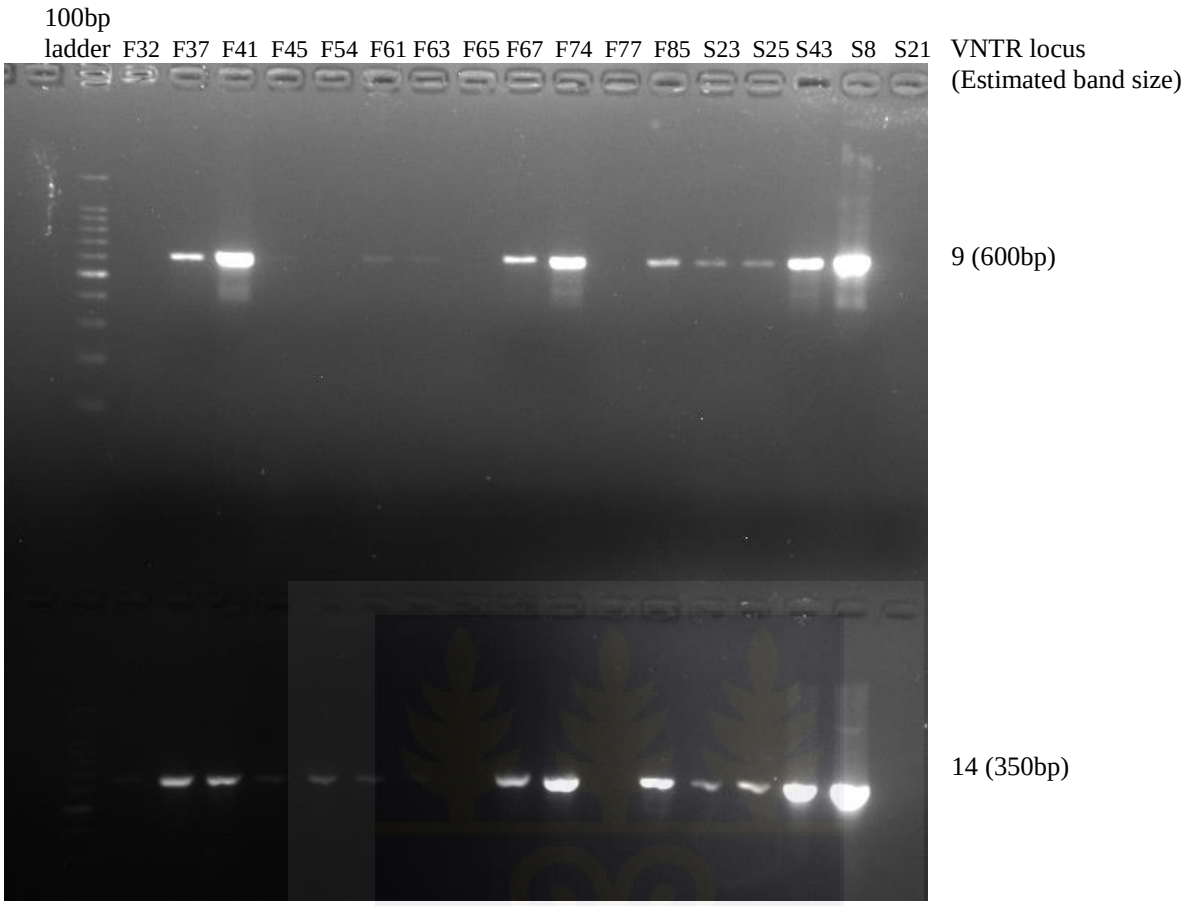


Fig. 4.72: PCR analysis of VNTR loci 9 and 14 for *M. ulcerans* isolates (F32 to S21)

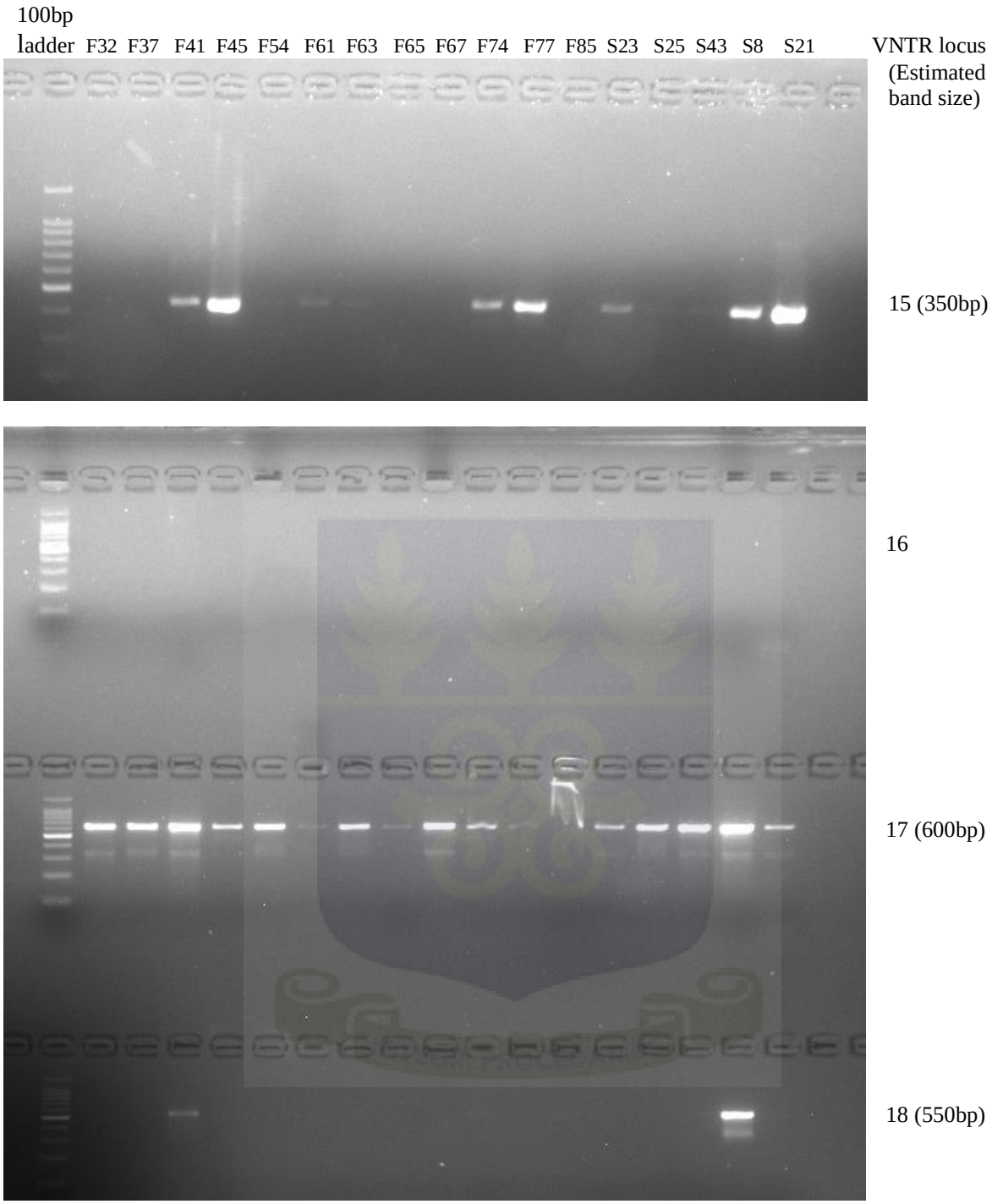


Fig. 4.73: PCR analysis of VNTR loci 15, 16, 17 and 18 for *M. ulcerans* isolates (F32 to S21)

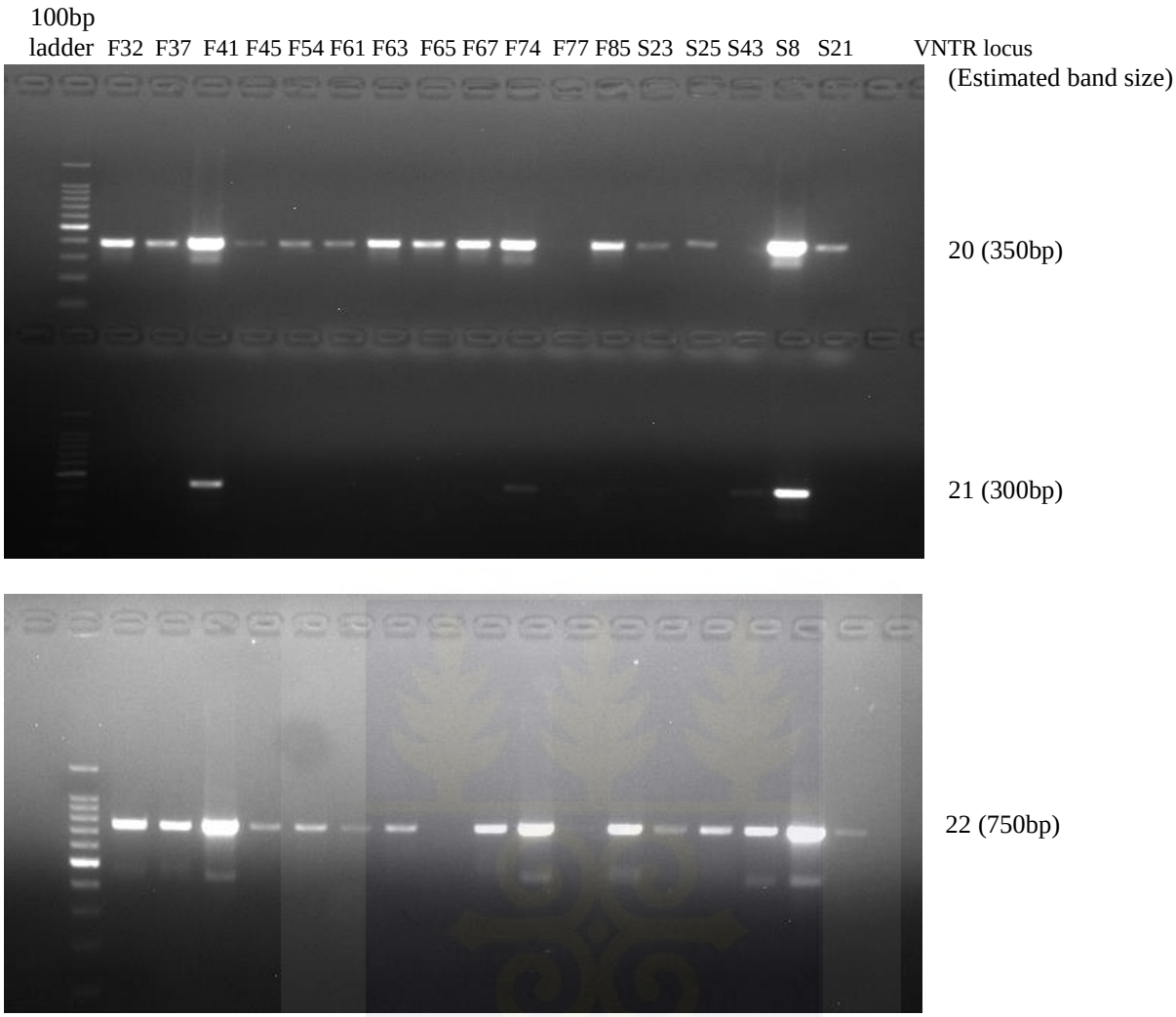


Fig. 4.74: PCR analysis of VNTR loci 20, 21 and 22 for *M. ulcerans* isolates (F32 to S21)

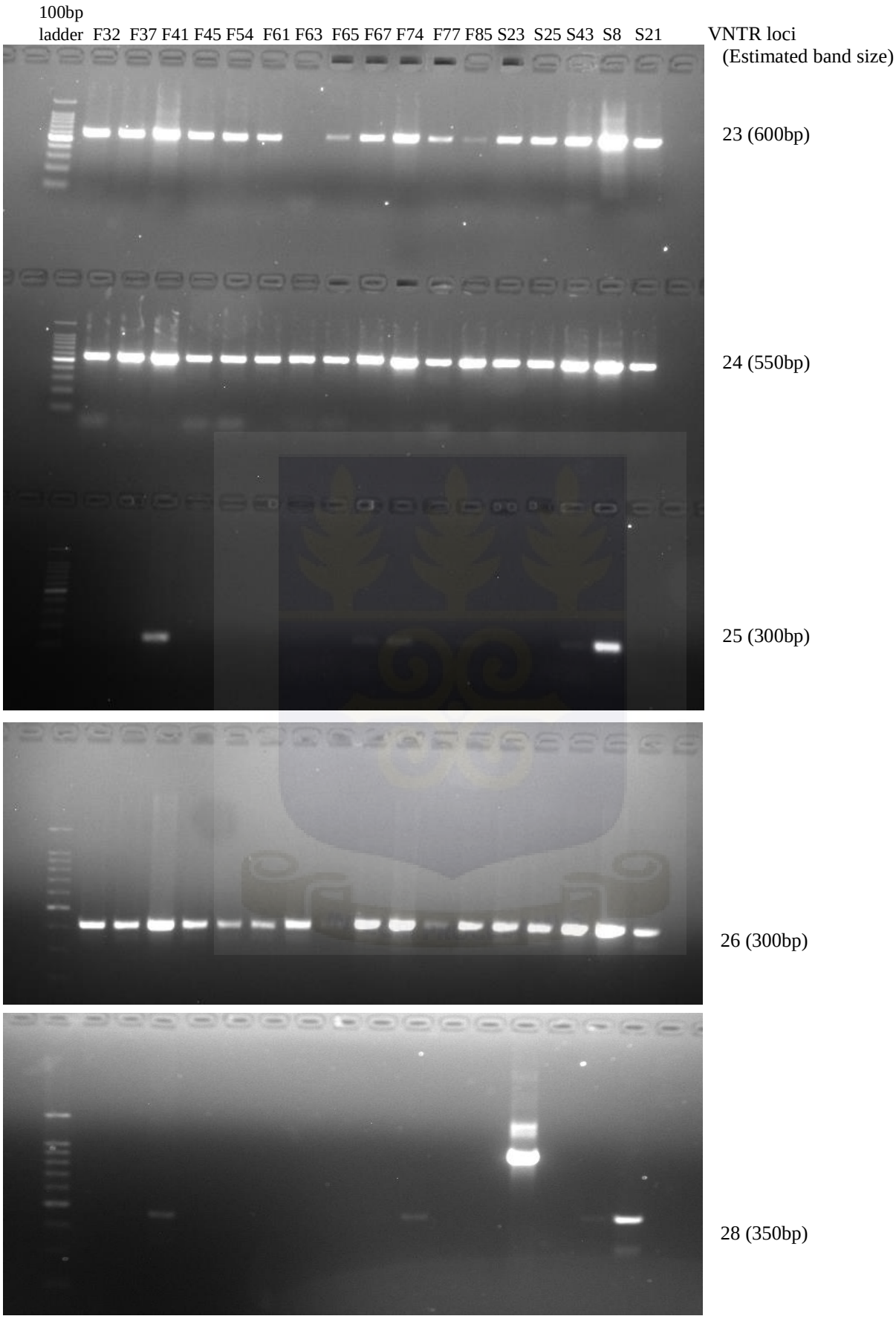


Fig. 4.75: PCR analysis of VNTR loci 23, 24, 25, 26 and 28 for *M. ulcerans* isolates (F32 to S21)

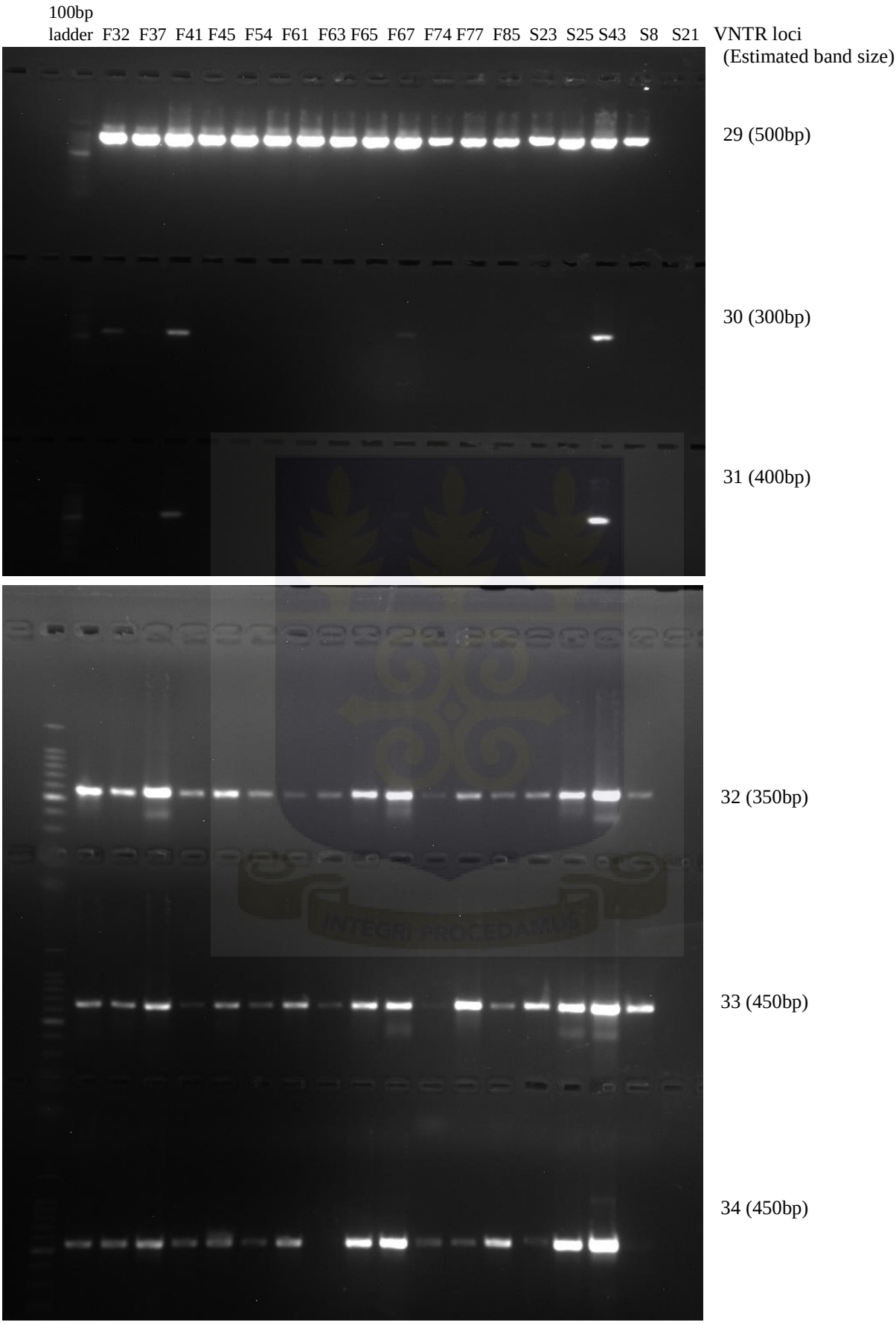


Fig. 4.76: PCR analysis of VNTR loci 29, 30, 31, 32, 33 and 34 for *M. ulcerans* isolates (F32 to S21)

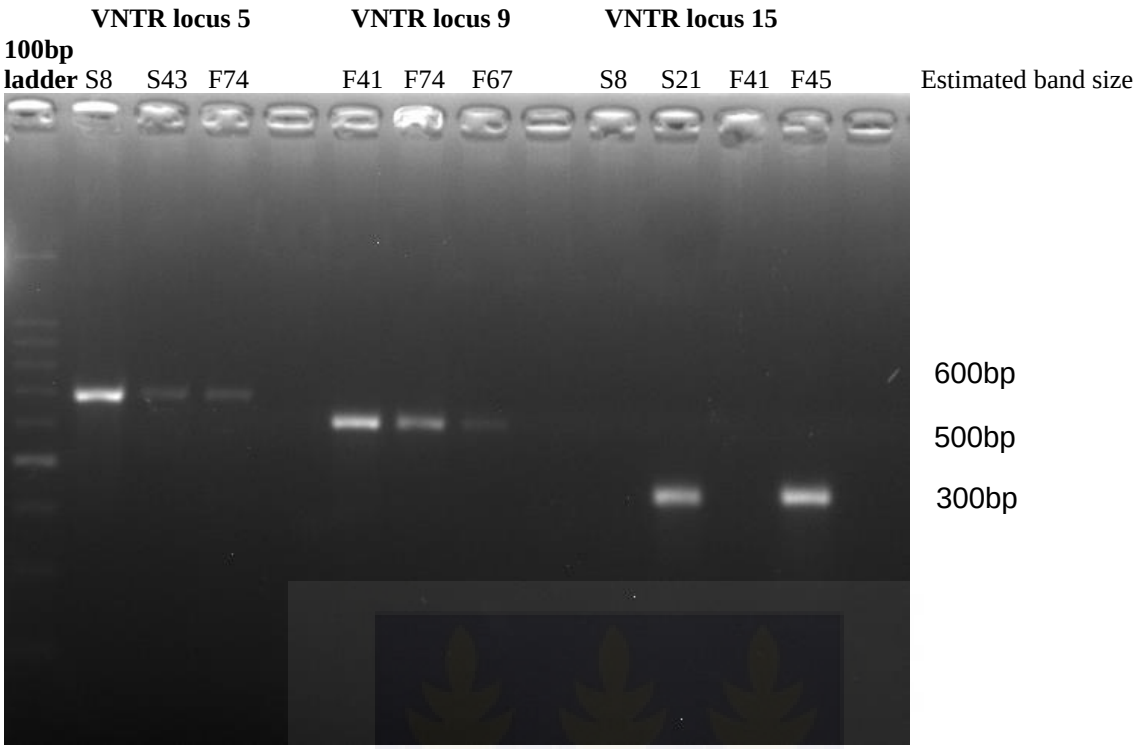


Fig. 4.77: PCR analysis of VNTR loci 5, 9 and 15 for *M. ulcerans* isolates

4.30 Description of results for the LAMP study

4.31 Detection limit of LAMP for *M. ulcerans*

The three pocket warmers (of a pack of 30 hand warmers) tested in this study achieved a temperature of 60^{o}C after 60 min and maintained this temperature for about 90 min. The pocket warmer thus provided a suitable temperature (60^{o}C) and time range (60 min) for amplification. Both pocket warmer loop-mediated isothermal amplification (pwLAMP) and the conventional LAMP assays were able to detect to the limit of 300 copies of the target sequence after 60 min of amplification. This limit improved to 30 copies when the conventional LAMP was carried out at 65^{o}C. The pocket warmer was not able to attain this temperature and was therefore not investigated.

4.32 Specificity of pwLAMP for *M. ulcerans*

Observation under ambient as well as UV illumination demonstrated clearly that the LAMP reaction produced positive signal specifically in DNA from *M. ulcerans*, but not in DNA extracts of *M. marinum*, *M. tuberculosis*, *M. avium*, *M. intracellularie*, *M. kansasii*, *M. abscessus*, *M. chelonae* and Jurkart, a human T cell line (Figure 4.8).

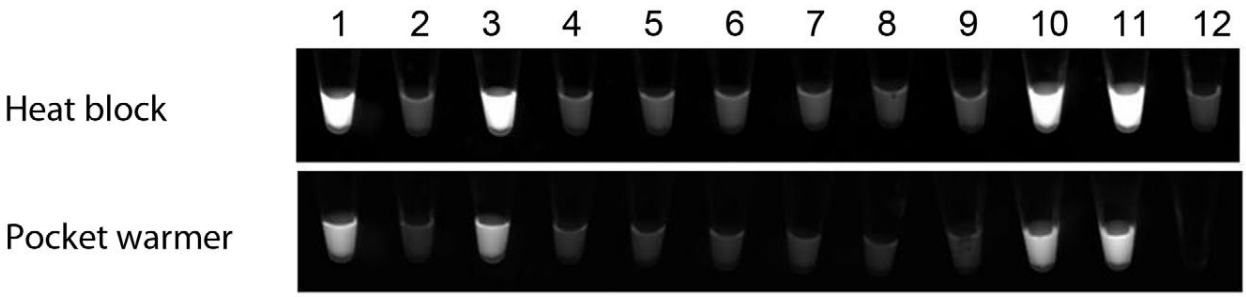


Figure 4.8: Specificity of Loop mediated isothermal amplification (LAMP) for *Mycobacterium ulcerans*. Conventional (upper, heat block) and pw-LAMP (lower, pocket warmer). Fluorescence image under the UV light are shown. Lanes; 1; *Mycobacterium ulcerans*, 2; *Mycobacterium marinum*, 3; *Mycobacterium shinsuense*, 4; *Mycobacterium tuberculosis*, 5; *Mycobacterium avium*, 6; *Mycobacterium intracellularie*, 7; *Mycobacterium kansasii*, 8; *Mycobacterium abscessus*, 9; *Mycobacterium chelonae*, 10; *Mycobacterium ulcerans*, 11; *Mycobacterium ulcerans* and 12; Jurkart cell line.

4.33 Comparison of LAMP with IS*2404* PCR

Under ambient illumination, positive specimens (specimens containing *M. ulcerans* DNA) in the LAMP assay produced greenish colouration (Figure 4.9).

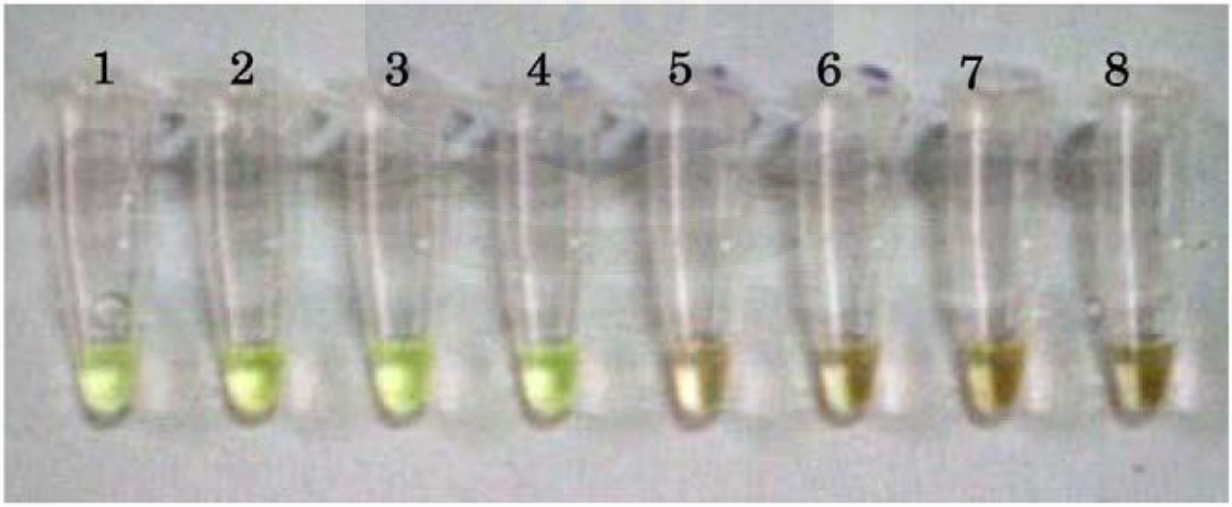


Figure 4.9: Detection of *Mycobacterium ulcerans* under ambient illumination. Tubes (1–4) containing *M. ulcerans* DNA produced greenish fluorescence.

The sensitivity and specificity of the LAMP assays for the detection of *M. ulcerans* is shown in Tables 4.5 and 4.6.

Table 4.5: Comparison of IS*2404* PCR with Pocket warmer LAMP for *M. ulcerans* detection

	Pocket warmer LAMP								
	Unboiled extract			Boiled extract			Purified extract		
	(+)	(-)	Total	(+)	(-)	Total	(+)	(-)	Total
IS*2404* PCR (+)	12	9	21	9	12	21	19	2	21
IS*2404* PCR (-)	0	9	9	0	9	9	0	9	9
Total	12	18	30	9	21	30	19	11	30
Positivity	40%			30%			63.3%		
*Sensitivity	100%			100%			90.5%		

*Sensitivity as compared with IS*2404* PCR

Table 4.6: Comparison of IS*2404* PCR with conventional LAMP for the detection of *M. ulcerans*

	Conventional LAMP								
	Unboiled extract			Boiled extract			Purified extract		
	(+)	(-)	Total	(+)	(-)	Total	(+)	(-)	Total
IS*2404* PCR (+)	12	9	21	9	12	21	21	0	21
IS*2404* PCR (-)	0	9	9	0	9	9	0	9	9
Total	12	18	30	9	21	30	21	9	30
Positivity	40%			30%			70%		
*Sensitivity	100%			100%			100%		

*Sensitivity as compared with IS*2404* PCR

When purified DNA extracts were used, 21 (16 swabs, 5 fine needle aspirates) (70%) of 30 clinical specimens were positive by IS*2404* PCR as well as by the conventional LAMP as shown in Table 4.2. None of the PCR positive specimens were negative by conventional LAMP. However 19 samples of purified DNA extracts were positive with the pwLAMP, but the 90.5% sensitivity (19/21) of the pwLAMP (Table 4.5) compared to that of conventional LAMP and IS*2404* PCR (100%, 21/21) was not statistically significant (χ^2 (1) = 0.40, p = 0.580).

All negative specimens in IS*2404* PCR were negative in both LAMP assays, indicating specificities of both LAMP assays to the reference method were 100%. Twelve unboiled (9 swabs and 3 fine needle aspirates) and 9 boiled (6 swabs and 3 fine needle aspirates) extracts were positive by all 3 detection assays with sensitivities of 57.1% (unboiled, 12/21) and 42.9% (boiled, 9/21) compared to results using purified DNA extracts respectively for both LAMP and IS*2404* PCR assays. The positivity of swabs was found to be in the range of 30% to 80% compared to 30% to 50% for fine needle aspirates.

When the positivities in crude DNA specimens were compared with those in purified DNA, the differences were statistically significant by chi-square test (unboiled vs purified DNA, [χ^2 (1) = 6.04, p = 0.0195], and boiled vs. purified DNA, [χ^2 (1) = 10.71, p = 0.0019]). None of the IS*2404* PCR negatives was positive in the LAMP assays irrespective of the DNA extracts type used. These data suggest that sensitivity of LAMP and PCR assays for the detection of *M. ulcerans* in clinical specimens is enhanced when purified DNA extracts are used.

CHAPTER FIVE

5.00 Discussion, conclusions and recommendations

5.10 Discussion on whole genome sequencing study

5.11 Overview of *M. ulcerans* diversity assessed by typing methods

Different molecular typing assays employing several independent markers have been used to study genetic diversity in *M. ulcerans*. Invariably, these investigations (Ablordey *et al*., 2005a, b; Chemlal *et al*., 2001; Hilty *et al*., 2006; Jackson *et al*., 1995; Portaels *et al*., 1996; Stragier *et al*., 2006; Stinear *et al*., 2000) give indication of limited genetic diversity in *M. ulcerans* especially among isolates within a given geographic region. This together with congruent phyologenetic trees topologies derived from independent genetic target analysis gives indication of a clonal population structure for *M. ulcerans*. Lack of genetic diversity in *M. ulcerans* populations is an important factor hindering the understanding of some fundamental issues of Buruli ulcer (BU) epidemiology such as identifying the reservoir and transmission route(s) of this pathogen. These knowledge gaps have led to difficulties in formulating primary preventive measures for the control of the disease. For Mycobacteria generally, mobile genetic elements such as insertion sequences are major drivers of genetic diversity, causing genome reshuffling through transition and attendant deletions and insertions and also recombination mechanisms.

M. ulcerans is known to possess high copies of two insertion sequences- IS*2404* and IS*2606*. However, molecular typing assays based on these targets have yielded limited genetic diversity among *M. ulcerans* isolates originating from a given geographic area. Genomic regions containing tandem repeat (TR) sequences are also hotspots for genetic differences among species of Mycobacteria. PCR amplification of tandem repeat loci is used to index

length polymorphism in *M. tuberculosis* and provides a useful molecular tool for studying epidemiology of tuberculosis. Although tandem repeat profiling has resulted in limited discrimination among *M. ulcerans* populations, the level of discrimination is not permissive for detailing epidemiological events. Since only a limited fraction of the entire tandem repeat repertoire of *M. ulcerans* have been investigated, researchers including Ablordey and colleagues (2005b) and Stragier and colleagues (2005) have recommended the need for further expansion of the current TR loci used for typing *M. ulcerans*.

Given the extremely clonal structure of *M. ulcerans* populations, single nucleotide polymorphism (SNP) analysis is now used for studying genetic differences among isolates of this pathogen. SNP analysis using limited sections of the genome has already led to the discrimination of previously unresolved types, and more significantly, improvement in discrimination among isolates from a given locale. This study extends the SNP discovery to cover the entire genome of isolates using the resequencing approach.

5.12 *M. ulcerans* diversity based on whole genome sequencing

The application of Ion Torrent and Illumina sequencing technologies to whole genome sequencing of MU isolates of different endemic areas in Ghana and Benin in this study revealed a total of 299 single nucleotide polymorphism (SNP) positions in comparison with the reference isolate, Mu_Agy99. This finding imply that African *M. ulcerans* populations are not as diverse as other geographic variants particularly a Japanese strain in which 26,564 SNPs were identified based on genome sequencing approach (Qi *et al*., 2009). Genetic relationships based on the 299 SNP positions inferred by neighbor-joining method using uncorrected *p* distances produced five distinct *M. ulcerans* clusters with high bootstrap values (Table 4.6), implying high confidence in clustal branches. Clustering of the isolates largely

reflected the origin of isolates. Members of each cluster represent a unique SNP type except the tight cluster that had four isolates with identical SNP type together with one other isolate with a unique SNP type.

It was generally observed that *M. ulcerans* isolates from a given area had comparable SNP numbers as well as similar pattern of distribution of SNPs. The only isolate from the Amansie West District (Mu_NM33.04) differed from the reference isolate (Mu_Agy99) by only 25 SNPs across the entire genome. The two isolates clustered when phylogenetic relationship was inferred (Figure 4.6). Similar cluster was observed by Roltgen and colleagues (2010) when neighbor-joining analysis sub-grouped these isolates into a clade together with few other *M. ulcerans* isolates on the basis of SNP typing. The close genetic relationship between these isolates suggests that they originate from the same endemic area. Although Mu_Agy99 has previously been described as a patient isolate from the Ga District, Ghana (Stinear *et al*., 2007), additional information on previous residence or travel history of the patient, perhaps from hospital records, may help clarify the place of origin of Mu_Agy99. This isolate also displays variant distribution pattern that is completely different from those of isolates from the Ga and Asante Akim North Districts as well as isolates from Benin (Figures 4.3 and 4.4), reinforcing its distinct place of origin.

Given the common place of origin described previously (Roltgen *et al*., 2010), one would expect fewer than the observed SNP differences between the Ga District isolates and the reference strain, Mu_Agy99. Additionally, the reference strain clustered with the strain from the Amansie West District and, therefore, seems more likely to originate from the Amansie West District and not from the Ga District.

Variant distribution is across the entire genome of these isolates with regions of conservation generally interspersed with more variable regions as shown in Figure 4.4. However, the pattern of variant distribution appears to be similar for Mu_NM49.02 and Mu_NM54.02 across the entire genome compared to isolates Mu_NM43.02 and Mu_NM14.01. These *M. ulcerans* isolates are members of ten different haplotypes identified by Roltgen and colleagues (2010) in a relatively small BU endemic area within the Densu river basin in the Ga District of Ghana. The isolate NM14/01 (haplotype 5) was observed to be the most common isolate, colocalizing with all other haplotypes within the endemic area, and has been suggested by Roltgen and colleagues (2010) to represent the founder haplotype in the endemic area. Using real-time PCR SNP typing method and sequencing technique to investigate MU diversity, Roltgen and colleagues (2010) identified smaller SNP variation (14 SNPs) within 74 MU isolates originating from the Ga District. On the contrary, 20 SNPs were identified within only four Ga District isolates in this study using whole genome resequencing approach. It is, therefore, conceivable that whole genome sequencing of more isolates from the Ga District will discover more SNP numbers.

The three Benin *M. ulcerans* isolates had interesting SNP numbers. The insect isolate (Mu_06-3845) from Gerridae, *Gerris* sp., and one of the human isolates (Mu_07-1082) had comparable SNP numbers, 258 and 254 respectively, differing by 4 single nucleotide polymorphisms (SNPs). These two isolates originated from the Zou/Ouémé valley in Benin and clustered together in this study in conformity with the findings of Doig and colleagues (2012). The comparable SNP numbers and the common place of origin may suggest insect transmission of the bacterium. However, it is quite unlikely that insects might have transmitted the bacterium to humans, as there are no reports of *Gerris* spp. biting humans (Portaels *et al*., 2008).

The observation that the Benin isolate (Mu_06-3846) was much more closely related to the reference strain (with 24 SNP differences between them) than the other Benin isolates also represents a departure from the usual clustering of isolates from the same endemic area. The close similarity between Mu 06-3846 and the Ga district isolates has also been previously documented (Doig *et al*., 2012). *M. ulcerans* isolates from the Couffo river basin, from where Mu_06-3846 was obtained, and Zou/Ouémé river basin have been observed to have genetic difference (Portaels *et al*., 2008). As suggested previously (Doig *et al*., 2010), sequencing more isolates from the Couffo valley, Togo and Ghana may offer explanation for this pattern of clustering. In addition, information on previous residence or travel history of the patient from whom Mu_06-3846 was isolated may be useful in this regard.

The isolates that form the tight cluster in the AAND showed similar pattern of variation across the entire genome and comparable SNP numbers, about two times the SNP numbers recorded for other isolates within the endemic area. The pattern of SNP variation among AAN isolates that form the tight cluster suggests a recent introduction and rapid spread of this genotype in the endemic communities of the AAN district. The origin of this genotype is yet to be established.

5.121 *M. ulcerans* distribution and diversity in the Asante Akim North District (AAND)

Farming is the predominant occupation for most of the BU victims in the AAND. The victims include children who often help their parents on the farm or fish in nearby rivers. It is interesting to note that most of the *M. ulcerans* isolates that formed the tight cluster were recovered from patients who reside or farm close to the Ongwam River and its tributaries (Figure 3.1). This suggests a common source of infection somewhere around this river. One

of the victims from whom an isolate of the tight cluster was recovered lived in Nsonyameye for four years and, at the time of the study, was a resident of Ananekrom. She visits endemic areas near Agogo to buy and sell fresh fish from the Afram River. Although this victim does not live in the vicinity of the Ongwam River system, she might have had contact with the bacterium whiles doing her business.

Similarly, victims of the other isolates live and work in villages close to water bodies such as the Egyan Stream and the Afram River. For example isolate Mu_F64 was obtained from a three-year-old boy whose parents farmed close to the Egyan Stream and resided in Nsonyameye (Figure 3.1). The boy might have been exposed to *M. ulcerans* whiles the parents were farming. Interestingly, the sister of the young boy had Buruli ulcer in the past. However, the SNP numbers of this group of isolates observed are smaller and form a cluster different from the tight cluster. It is revealing that this study discovered two unique MU clusters in the AAND with larger, but different SNP numbers compared with SNP numbers discovered in previous studies (Qi *et al*., 2009; Roltgen *et al*., 2010) focusing on isolates from other endemic areas in Ghana. The Buruli ulcer victim from whom Mu_S38 was obtained is a farmer who lives at Serebouso and goes fishing in Ongwam River. The Mu_S38 isolate had SNP numbers comparable to isolates in this group, suggesting mixed distribution of isolates in the district. Studying more *M. ulcerans* isolates from villages drained by the water bodies in the AAND including the Ongwam, Afram and Kowire Rivers and Egyan Stream will help throw more light on the diversity and the extent of distribution of isolates in the two clusters observed in this study. Such information will also help track transmission pathways.

Undoubtedly, residents and Buruli ulcer victims of these villages have the freedom to move to other villages within the endemic area for various reasons including work and business.

Although it is probable to pick up infection from villages of residence by virtue of frequent exposure to potential sources of infection, the possibility of getting infection from sources outside villages of residence cannot be ruled out. It is, therefore, imperative for Buruli ulcer clinical sample collection to be accompanied with detailed information on activities that may expose people to *M. ulcerans* infection to help clarify the mode of infection.

5.122 Protein consequences of SNPs detected

The study revealed that about one-third (58) of the single nucleotide polymorphisms (SNPs) observed in coding regions in *M. ulcerans* isolates (Table 4.3) are synonymous. However, the rest may have consequences on proteins produced and metabolic processes in *M. ulcerans*. Further investigation on functional analysis of coding sequences with non-synonymous SNPs may help assess the importance of the SNPs in relation to niche expansion and epidemic spread of *M. ulcerans* as suggested previously (Huber *et al*., 2008; Kaser & Pluschke, 2008). Protein coding genes in *M. ulcerans* African populations have been suggested to harbour extremely low levels of polymorphisms (Qi *et al*., 2009). Among strains originating from Africa, no SNPs were detected in a few chromosomal and plasmid genes based on multi-locus sequence typing (Stinear *et al*., 2000b; Stinear *et al*., 2005a; Yip *et al*., 2007). However, whole genome sequencing analysis in this study revealed SNPs in coding sequences of *M. ulcerans* (Table 4.3), implying variable distribution of SNPs in *M. ulcerans* coding sequences.

The isolate Mu_NM43.02 from the Ga District was observed to have more SNPs in coding sequences that are unique to it and may be a reflection of the cluster for the Ga District isolates. Generally, the isolates in the same cluster have more common SNPs in their coding sequences compared with those in different clusters, emphasizing their close genetic

relatedness. It is interesting to note that the insect isolate, Mu_06-3845, obtained from *Gerris* sp. (Water Strider) in a BU endemic region in Benin (Portaels *et al*., 2008), has a lot of SNPs in its coding sequences that are common to the human isolate from Benin, Mu_07-1082. On the contrary, Mu_M3, a recent Belostomatid isolate from the Asante Akim North District shared more similarities with the Ga District isolates than AAND isolates (Table 4.3 and Figure 4.6).

5.20 Conclusions on whole genome sequencing study

The application of Ion Torrent and Illumina sequencing technologies to whole genome sequencing of *M. ulcerans* isolates of different endemic areas in Ghana and Benin in this study revealed a total of 299 single nucleotide polymorphism (SNP) positions in comparison with the reference isolate, Mu_Agy99. Phylogenetic analysis produced five distinct *M. ulcerans* (MU) clusters with high confidence in clustal branches: two for the isolates from the Asante Akim North District (AAND) and one each for the isolates from other endemic areas studied. Generally, clustering of isolates largely reflected the origin of isolates. However, there was a departure from the usual clustering of isolates from the same endemic area, as observed in the Benin isolate (Mu_06-3846) and the reference strain.

This study discovered two unique MU clusters in the AAND with larger, but different SNP numbers compared with SNP numbers discovered in previous studies (Qi *et al*., 2009; Roltgen *et al*., 2010) focusing on isolates from other endemic areas in Ghana. Interestingly, victims of the isolates that form the tight cluster in the AAND live in and work around villages close to the Ongwam River and its tributaries. The pattern of SNP variation among these isolates suggests a recent introduction, rapid spread of this genotype and a common

source of infection in the endemic communities of the AAN district. Studying more isolates from villages drained by water bodies in AAND will be useful in clarifying the mode of infection and throwing more light on the diversity and the extent of distribution of *M. ulcerans* isolates.

Given the common place of origin described previously (Roltgen *et al*., 2010), one would expect fewer than the observed SNP differences between the Ga District isolates and the reference strain. Additionally, the reference strain clustered with the strain from the Amansie West District and, therefore, seems more likely to originate from the Amansie West District and not from Ga district. Additional information on previous residence or travel history of the patient from whom the reference strain was isolated will be necessary to be definitive on its place of origin. Although the insect isolate from Gerridae and one human isolate originated from the Zou/Ouémé valley in Benin had comparable SNP numbers, it is quite unlikely that insects might have transmitted the bacterium to humans, as there are no reports of *Gerris* spp. biting humans (Portaels *et al*., 2008).

Clearly, standard molecular typing methods have proven to be useful in differentiating *M. ulcerans* strains from different geographic areas. However, they have generally made modest contribution to differentiating strains from the same geographic areas. The availability and utility of high-throughput sequencing technologies and their vast superiority to traditional sequencing methods on account of the high volume of data (Hall, 2007; Church, 2006) and the relatively short time it takes to sequence a whole genome will make a huge impact on Buruli ulcer control. Whole genome sequencing of *M. ulcerans* isolates will, no doubt, provide more information on genetic diversity which could aid in the identification of the reservoir and transmission routes of *M. ulcerans*. It could also further illuminate our

understanding on the relative contribution of exogenous reinfection versus reactivation in disease recurrence. These will go a long way to improve our ability to control Buruli ulcer disease.

5.30 Recommendations

- Whole genome analysis of a collection of isolates has the potential of yielding larger number of SNPs, which will form the basis for comprehensive assessment of *M. ulcerans* genomic variability at the regional and local levels of Buruli ulcer endemicity. Whole genome single nucleotide polymorphism (SNP) analysis should, therefore, be encouraged.

- Whole genome SNP analysis complemented with previous residence or travel history of patients from whom *M. ulcerans* isolates are collected will help bring clarity to geographical origins of isolates.

- Development of molecular typing methods based on genomic regions of high variability, for all the isolates, that have not been explored will be useful in differentiating isolates with unknown SNP characteristics.

- It is recommended that more *M. ulcerans* isolates from villages drained by the water bodies in the Asante Akim North District including the Ongwam, Afram and Kowire Rivers and Egyan Stream be studied to throw more light on the diversity and the extent of distribution of isolates in the two clusters observed in this study.

- Collecting detailed information on activities that may expose people to *M. ulcerans* infection during Buruli ulcer clinical sample collection will be useful in clarifying the mode of infection and tracking transmission pathways.

- Focussing attention on other genetic changes, such as multiple nucleotide variant and insertion-deletions, in these isolates will give clearer picture of overall genetic diversity in *M. ulcerans*.

- Further investigation on functional analysis of coding sequences with non-synonymous SNPs may help assess the importance of the SNPs in relation to niche expansion and epidemic spread of *M. ulcerans*.

5.40 Discussion, conclusion and recommendation of the VNTR study

Variable number of tandem repeat (VNTR) typing methods have made important contribution to the understanding of genetic diversity in *M. ulcerans*, differentiating in particular isolates from different geographic regions (Stragier *et al*., 2006; Ablordey *et al*., 2005a, b; Hilty *et al*., 2006). In addition, some gains in strain discrimination within the same geographic region have been recorded (Ablordey *et al*., 2005a; Hilty *et al*., 2006), demonstrating their importance in genetic diversity studies in isolates from the same geographic region. With this background, this study tested twenty five (25) newly-described VNTR loci in seventeen (17) isolates from three different Buruli ulcer endemic areas in Dormaa, Suhum-Kraboa-Coaltar and Asante Akim North Districts in Ghana and assessed their importance in typing *M. ulcerans* for the purpose of genetic differentiation.

Twenty three (23, 92%) of the VNTR loci were observed to be useful in typing *M. ulcerans*, amplifying VNTR sequences in isolates studied. More importantly, 15 (60%) of the loci were amplified in 14 (82%) of the isolates, highlighting them as targets for VNTR genotyping for the purpose of genetic differentiation. The DNA band sizes for all loci were estimated to range between 300 bp and 750 bp. Almost all isolates for each locus have similar band sizes with few exceptions (S8 and S43 at locus 5; F41and F74 at locus 9; S8, S21 and F45 at locus 15) that appear to differ in band sizes (Figures 4.70 – 4.76), necessitating variation of electrophoretic conditions for better resolution of the DNA band sizes. However, variation of electrophoretic conditions produced no differences in band sizes for loci 5, 9 and 15 with estimated DNA band sizes of 600 bp, 500 bp and 350 bp respectively (Fig. 4.77). Interestingly, isolates S43, F41, F74 and F45 come from the AAND endemic area different from those of S8 and S21, which differ in districts of origin. Sequencing of these VNTR loci will give clearer picture of the extent to which the isolates may differ genetically.

Clearly, the newly-described VNTR loci have been demonstrated to be useful for VNTR typing with 60% of them being targets for DNA amplification in 14 of the isolates. Further study based on these loci in combination with sequencing is recommended. This will provide useful information on the diversity of *M. ulcerans* and the epidemiology of Buruli ulcer.

5.50 Discussion, conclusion and recommendation of the LAMP study

5.51 Discussion of the LAMP study

The IS*2404* PCR is considered to be the most reliable technique for the detection of *M. ulcerans* in human diagnostic samples. Despite its reliability due to its high sensitivity (Beissner *et al*., 2010; Herbinger *et al*., 2009; Stienstra *et al*., 2003), very high specificity (Ross *et al*., 1997; Rondini *et al*., 2003; Siegmund *et al*., 2005) and fairly good positivity ratios (Siegmund *et al*., 2007; Mensah-Quainoo *et al*., 2008), it is rarely available for use especially in countries where there are inadequate laboratory facilities and trained personnel, and financial constraints. This calls for the development of rapid and reliable point of care diagnostic assays to facilitate effective control of Buruli ulcer.

Isothermal DNA amplification coupled with fluorescence signal detection provides unique opportunity for the development of point of care test. In this study, the LAMP assay was developed as a point of care test for BU diagnosis. Initially, few challenges to the use of LAMP assay in the field including the difficulty in maintaining isothermal condition for the reaction as well as specimen purification were addressed.

The pocket warmers used provided the requisite temperature (60^{o}C) for the LAMP reaction in accordance with a previous study (Hatano *et al*., 2010). Moreover, the pwLAMP had 100% specificity in clinical specimens of Buruli ulcer and was found to have comparable sensitivity with the conventional LAMP at 60^{o}C, as both assays were able to detect 300 copies of IS*2404* element (equivalent of 1.5 genomes of *M. ulcerans*). At 65^{o}C, the conventional LAMP improved its detection limit to 30 copies of IS*2404*. It is possible that this level of sensitivity will be achieved with a pocket warmer capable of generating a temperature of 65^{o}C and maintaining a holding time of 60 min.

It is important to note that the results for the pw-LAMP, conventional LAMP and IS*2404* PCR were in agreement (Tables 1 and 2), yielding 90.5% sensitivity for the pw-LAMP compared with 100% sensitivity for both the conventional LAMP and IS*2404* PCR when purified DNA extracts were used with no significant difference ($p > 0.58$, Chi-square test). All IS*2404* PCR negative samples were also negative for both LAMP assays, indicating 100% specificities for them compared to the reference method. However, there was a drop in detection of positive samples from 63–70% to 30–40% when crude extracts of clinical specimens were used (Tables 1 and 2), with significant differences by chi-square test (unboiled vs. purified DNA, [χ^2 (1) = 6.04, p = 0.0195], and boiled vs. purified DNA, [χ^2 (1) = 10.71, p = 0.0019]). These results imply that using crude DNA extracts as template may not be appropriate for detection of *M. ulcerans* by the LAMP method. This observation contradicts the findings of Kaneko and colleagues (2007) that omission of DNA extraction has no effect on sensitivity of the LAMP assay. Interestingly, the detection of positive samples for the LAMP assay was significantly higher for the unboiled extracts than for the boiled extracts, suggesting that the LAMP assay was not inhibited especially for the unboiled specimens. This observation is consistent with previous study (Kaneko *et al*., 2007) that has shown LAMP to be tolerant to culture medium and to certain biological substances including phosphate buffered saline, serum, plasma, urine and vitreous.

5.52 Conclusion and recommendation of the LAMP study

The study demonstrates that the LAMP assay yields results comparable to that of IS*2404* PCR when it is performed at 60°–65°C for 60 minutes on purified DNA extracts, supporting the hypothesis that the pw-LAMP method for *M. ulcerans* DNA amplification is as efficient as that for IS*2404* PCR. It further supports the use of the pocket warmer as a device for providing isothermal amplification condition for the LAMP assay and is obviously a potential

boost to the application of pwLAMP in resource poor settings. However, challenges of obtaining pure DNA extracts of clinical specimen as well as the use of a pocket warmer capable of maintaining 65ºC for one hour need to be addressed in order to improve the performance of the pwLAMP assay. Further development and testing in larger numbers of specimens is recommended to further assess the potential use of pwLAMP as a simple and rapid point of care diagnostic test for Buruli ulcer.

REFERENCES

Ablordey, A., Hilty, M., Stragier, P., Swings, J. & Portaels, F. (2005a). Comparative nucleotide sequence analysis of polymorphic variable-number tandem-repeat loci in *Mycobacterium ulcerans*. *J Clin Microbiol* **43**: 5281–5284.

Ablordey, A., Swings, J., Hubans, C., Chemlal, K., Locht, C., Portaels, F. & Supply, P. (2005b). Multilocus variable-number tandem repeat typing of *Mycobacterium ulcerans*. *J Clin Microbiol*, **43**: 1546–1551.

Ablordey, A., Kotlowski, R. Swings, J. & Portaels, F. (2005c). PCR amplification with primers based on IS*2404* and GC-rich repeated sequence reveals polymorphism in *Mycobacterium ulcerans*. *J. Clin. Microbiol*., **43:**448–450.

Ablordey, A., Fonteyne, P.A., Stragier, P., Vandamme, P. & Portaels, F. (2007). Identification of a new variable number tandem repeat locus in *Mycobacterium ulcerans* for potential strain discrimination among African isolates. *Clin Microbiol Infect* **13**: 734–736.

Achtman, M. (2008). Evolution, Population Structure, and Phylogeography of Genetically Monomorphic Bacterial Pathogens. *Annu. Rev. Microbiol*. **62**:53–70.

Adessi, C., Matton, G., Ayala, G., Turcatti, G., Mermod, J-J., Mayer. P. & Kawashima, E. (2000). Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res.* **28**, e87.

Adusumilli, S., Mve-Obiang, A., Sparer, T., Meyers, W., Hayman, J. & Small, P.L.C. (2005). *Mycobacterium ulcerans* toxic macrolide, mycolactone modulates the host immune response and cellular location of *M. ulcerans in vitro* and *in vivo*. *Cell Microbiol* **7**: 1295–1304.

Aguiar, J. & Stenou, C. (1997). Buruli ulcers in rural areas of Bénin: management of 635 cases. *Med Trop* **57** (1): 83–90.

Aguiar, J., Domingo, M.C., Guédénon, A., Meyers, W., Steunou, C. & Portaels, F. **(**1997). L'ulce`re de Buruli, une maladie mycobacte´rienne importante et en recrudescence au Be´nin. *Bull. Se´ances Acad. R. Sci. Outre Mer*. **43:**325–358.

Ahoua, L., Guetta, A. N., Ekaza, E., Bouzid, S., N'Guessan, R. & Dosso, M. (2009). Risk factors for Buruli ulcer in Côte d'Ivoire: Results of a case-control study. *African Journal of Biotechnology* Vol. 8 (4), pp. 536-546.

Amofah, G., Bonsu, F., Tetteh, C., Okrah, J., Asamoa, K., Asiedu, K. & Addy, J. (2002). Buruli ulcer in Ghana: results of a national case search. *Emerg Infect Dis*; **8**:167-70.

Amofah, G. K., Sagoe-Moses, C. & Frimpong, E. H. (1993). Epidemiology of Buruli ulcer in Amansie West District, Ghana. *Trans R Soc Trop Med Hyg.*, **87**:644–5.

Amonsin, A., Li, L.L., Zhang, Q., Bannantine, J.P., Motiwala, A.S., Sreevatsan, S. & Kapur, V. (2004). Multilocus short sequence repeat sequencing approach for differentiating among *Mycobacteriun avium* subsp *paratuberculosis* strains., *J. Clin. Microbiol*., **42:**1694–1702.

Asiedu, K. & Etuaful, S. (1998). Socioeconomic implications of Buruli ulcer in Ghana: a three-year review. *Am J Trop Med Hyg* **59**:1015–1022.

Asiedu, K. & Wansbrough-Jones, M. (2007). Mycobacterium ulcerans infection (Buruli or Bairnsdale ulcer): challenges in developing management strategies. *Med J Aust* **186**: 55–6.

Aujoulat, I., Johnson, C., Zinsou, C., Guédénon, A. & Portaels, F. (2003). Psychosocial aspects of health seeking behaviours of patients with Buruli ulcer in southern Benin. *Trop Med Int Health* **8**: 750–759.

Bafende, A.E., Phanzu, M.D. & Imposo, B. B. (2004). Buruli ulcer in the Democratic Republic of Congo: epidemiology, presentation and outcome. *Trop Doct*; **34**:82-4.

Bär, W., Rüsh-Gerdes, S., Richter, E., Marquez de Bär, G., Dittmer, C., Papsdorf, H., ... Portaels, F. (1998). *M. ulcerans* infection in a child from Angola: diagnosis by direct detection and culture. *Trop Med Int Health*., **3**:189–96.

Barker, D.J. (1971). Buruli disease in a district of Uganda. *J Trop Med Hyg*; 74:260-4.

Barogui, Y., Johnson, R. C., van der Werf, T. S., Sopoh, G., Dossou, A., Dijkstra, P. U. & Ymkje Stienstra, Y. (2009). Functional Limitations after Surgical or Antibiotic Treatment for Buruli Ulcer in Benin. *Am. J. Trop. Med. Hyg.*, **81**(1), pp. 82–87.

Bayley, A.C. (1971). Buruli ulcer in Ghana. *BMJ*; 2: 401-2.

Beissner, M., Herbinger, K. & Bretzel, G. (2010). Laboratory diagnosis of Buruli ulcer disease. *Future Microbiology*, 5(3): 363-370(8).

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., ... Smith, A.J. (2008). "Accurate whole human genome sequencing using reversible terminator chemistry". *Nature* **456** (7218): 53–59.

Bhamidi, S. (2009). "Mycobacterial Cell Wall Arabinogalactan". *Bacterial Polysaccharides: Current Innovations and Future Trends*. Caister Academic Press.

Bretzel, G., Huber, K.L., Kobara, B., Beissner, M., Piten, E., Herbinger, K.H., ... Nitschke, J. (2011). Laboratory Confirmation of Buruli ulcer Disease in Togo, 2007–2010. *PLoS Negl Trop Dis* **5**(7): e1228.

Buntine, J. & Crofts, K. (2001). Buruli ulcer: management of *Mycobacterium ulcerans* disease: a manual for health care providers. Geneva: World Health Organization.

Burchard, G.D. & Bierther, M. (1986). Buruli ulcer: clinical pathological study of 23 patients in Lambarene, Gabon. *Trop. Med. Parasitol*; 37(1): 1–8.

Chauty, A., Ardant, M. F., Adeye, A., Euverte, H., Guedenon, A., Johnson, C., ... Grosset, J. (2007). Promising clinical efficacy of the combination streptomycin-rifampin for the treatment of Buruli ulcer (*Mycobacterium ulcerans* disease). *Antimicrob Agents Chemother* **51***:* 4029–4035.

Chemlal, K., de Ridder, K., Fonteyne, P. A., Meyers, W. M., Swings, J. & Portaels, F. (2001). The use of *IS2404* restriction fragment length polymorphisms suggests the diversity of *Mycobacterium ulcerans* from different geographic areas. *Am. J. Trop. Med. Hyg*., **64:**270–273.

Chemlal, K., Huys, G., Fonteyne, P. A., Vincent, V., Lopez, A. G., Rigouts, L., ... Portaels, F. **(**2001b). Evaluation of PCR-restriction profile analysis and IS*2404* restriction fragment length polymorphism and amplified fragment length polymorphism fingerprinting for identification and typing of *Mycobacterium ulcerans* and *Mycobacterium marinum*. *J. Clin. Microbiol*. **39:**3272–3278.

Church, G.M. (2006). "Genomes for all". *Sci. Am.* **294** (1): 46–54.

Clancey, J.K., Dodge, O.G., Lunn, H.F. & Oduori, M.L. (1961). Mycobacterial skin ulcers in Uganda. *Lancet*; ii: 951-4.

Costerton, J. W., Stewart, P. S. & Greenberg, E. P. (1999). Bacterial biofilms: A common cause of persistent infections. *Science* **284**: 1318–1322.

Cousins, D. V., Williams, S.N., Ross, B.C. & Ellis, T.M. (1993). Use of a repetitive element isolated from *Mycobacterium tuberculosis* in hybridization studies with *Mycobacterium bovis*: a new tool for epidemiological studies of bovine tuberculosis. *Vet. Microbiol*. **37:**1–17.

Coutanceau, E., Marsollier, L., Brosch, R., Perret, E., Goossens, P., Tanguy, M., ...Demangel, C. (2005). Modulation of the host immune response by a transient intracellular stage of *Mycobacterium ulcerans*: the contribution of endogenous mycolactone toxin. *Cell Microbiol*, **7**: 1187–1196.

David, M., Dzamba, M., Lister, D., Ilie, L. & Brudno, M. (2011). SHRiMP2: Sensitive yet Practical Short Read Mapping. *Bioinformatics*, 27(7):1011–1012.

Davies, K. (2011). Powering Preventative Medicine. Bio-IT World 2011.

Davies, K. (2010). It's "Watson Meets Moore" as Ion Torrent Introduces Semiconductor Sequencing. Bio-IT World 2010.

Debacker, M., Portaels, F., Aguiar, J., Steunou, C., Zinsou, C., Meyers, W. & Dramaix, M. (2006). Risk factors for Buruli ulcer, Benin. *Emerg Infect Dis* **12**: 1325–1331.

Debacker, M., Aguiar, J., Steunou, C., Zinsou, C., Meyers, W.M. & Portaels, F. (2005). Buruli ulcer recurrence, Benin. *Emerg Infect Dis*.; **11**: 584–9.

Debacker, M., Aguiar, J., Steunou, C., Zinsou, C., Meyers, W.M., Guedenon, A. ..... Portaels, F. (2004a). *Mycobacterium ulcerans* disease (Buruli ulcer) in a rural hospital, southern Benin, 1997–2001. *Emerg Infect Dis.*, **10**: 1391–8.

Debacker, M., Aguiar, J., Steunou, C., Zinsou, C., Meyers, W. M., Scott, J. T., ... Portaels, F. (2004b). *Mycobacterium ulcerans* disease: role of age and gender in incidence and morbidity. *Trop Med Int Health*., **9**:1297–304.

Dega, H., Bentoucha, A., Robert, J., Jarlier, V. & Grosset, J. (2002). Bactericidal activity of rifampin-amikacin against *Mycobacterium ulcerans* in mice. *Antimicrob. Agents Chemother*. **46:** 3193–3196.

Dobos, K. M., Small, P. L., Deslauriers, M., Quinn, F. D. & King, C. H. (2001). *Mycobacterium ulcerans* cytotoxicity in an adipose cell model. *Infect Immun*; **69:** 7182–86.

Drancourt, M., Jarlier, V. & Raoult, D. (2002). The environmental pathogen *Mycobacterium ulcerans* grows in amphibian cells at low temperatures. *Appl Environ Microbiol* **68**, 6403–6404.

Dressman, D., Yan, H., Traverso, G., Kinzler, K.W. & Vogelstein, B. (2003). Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc. Natl. Acad. Sci. USA* **100**, 8817–8822.

Dukes, J.P., King, D.P. & Alexandersen, S. (2006). Novel reverse transcription loop-mediated isothermal amplification for rapid detection of foot-and-mouth disease virus. *Arch. Virol*. **151**:1093-1106.

Durnez, L., Stragier, P., Roebben, K., Ablordey, A., Leirs, H. & Portaels, F. (2009). A comparison of DNA extraction procedures for the detection of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer, in clinical and environmental specimens. *J Microbiol Meth* **76**: 152–158.

Erlich, H.A., Gelfandand, D.H. & Saiki, R.K. (1988). Specific DNA amplification. *Nature* **331**, 461-462.

Etuaful, S., Carbonnelle, B., Grosset, J., Lucas, S., Horsfield, C., Phillips, ... Wansbrough-Jones, M. (2005). Efficacy of the combination rifampin-streptomycin in preventing growth of *Mycobacterium ulcerans* in early lesions of Buruli ulcer in humans. *Antimicrob Agents Chemother* **49***:* 3182–3186.

Evans, M. R., Phillips, R., Etuaful, S. N., Amofah, G., Adomako, J., Adjei, O., ... Wansbrough-Jones, M. H. **(**2003). An outreach education and treatment project in Ghana for the early stage of *Mycobacterium ulcerans* disease. *Trans. R. Soc. Trop. Med. Hyg*. **97:**159–160.

Ewing, B. & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* **8**, 186–194.

Ewing, B., Hillier, L., Wendl, M.C. & Green, P. (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185.

Fedurco, M., Romieu, A., Williams, S., Lawrence, I. & Turcatti, G. (2006). BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. *Nucleic Acids Res.* **34**, e22.

Fine, P.E. (2001). BCG: the challenge continues. *Scand J Infect Dis.,* **33**:243–5.

Fine, P.E., Ponnighaus, J.M. & Maine, N. (1989). The distribution and implications of BCG scars in northern Malawi. *Bull World Health Organ*., **67**:35–42.

Floyd, S., Ponnighaus, J. M., Bliss, L., Warndorff, D. K., Kasunga, A., Mogha, P. & Fine, P.E.M. (2000). BCG scar in northern Malawi: sensitivity and repeatability of scar reading, and factors affecting scar size. *Int J Tuberc Lung Dis*., **4**:1133–42.

Fyfe, J., Lavender, C., Johnson, P., Globan, M., Sievers, A., Azuolas, J. & Stinear, T. P. (2007). Development and application of two multiplex real-time PCR assays for the detection of *Mycobacterium ulcerans* in clinical and environmental samples. *Appl Environ Microbiol* **73**: 4733–4740.

Frothingham, R. **(**1995). Differentiation of strains in *Mycobacterium tuberculosis* complex by DNA sequence polymorphisms, including rapid identification of *M. bovis* BCG. *J. Clin. Microbiol*., **33:**840–844.

Frothingham, R. & Meeker-O'Connell, W.A. (1998). Genetic diversity in *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. *Microbiology* **144**:1189–1196.

GenomeWeb DNA Electronics Licenses IP to Ion Torrent. August 2010.

George, K. M., Chatterjee, D., Gunawardana, G., Welty, D., Hayman, J., Lee, R. & Small, P. L. C. (1999). Mycolactone: a polyketide toxin from *Mycobacterium ulcerans* required for virulence. *Science*; **283:** 854–57.

George, K. M., Pascopella, L., Welty, D. M. & Small, P. L. (2000). A *Mycobacterium ulcerans* toxin, mycolactone, causes apoptosis in guinea pig ulcers and tissue culture cells. *Infect Immun*; **68:** 877–83.

Gibson, J. (1975). Buruli ulcers in Bo. *Bulletin of the Sierra Leone Medical and Dental Association*. **2**: 64-6.

Goto, M., Honda, E., Ogura, A., Nomoto, A. & Hanaki, K. (2009). Colorimetric detection of loopmediated isothermal amplification reaction by using hydroxy naphthol blue. *BioTechniques* **46**:167-172.

Goto, M., Hayashidani, H. Takatori, K. & Kudo, Y.H. (2007). Rapid detection of enterotoxigenic *Staphylococcus aureus* harbouring genes for four classical enterotoxins, SEA, SEB, SEC and SED, by loop-mediated isothermal amplification assay. *Lett. Appl. Microbiol*. **45**:100-107.

Guarner, J., Bartlett, J., Whitney, E.A., Raghunathan, P.L., Stienstra, Y., Asamoa, K. ….. Ashford, D.A. (2003). Histopathologic features of *Mycobacterium ulcerans* infection. *Emerging Infectious Diseases*, **9**: 651–656.

Guedenon, A., Zinsou, C., Josse, R., Andele, K., Pritze, S., Portaels, F. & Meyers, W.M. (1995). Traditional treatment of Buruli ulcer in Benin. *Arch Dermatol 131*: 741–742.

Guimaraes-Peres, A., Portaels, F., de Rijk, P., Fissette, K., Pattyn, S.R., van Vooren, J.-P. & P.-A. Fonteyne. (1999). Comparison of two PCRs for detection of *Mycobacterium ulcerans. J Clin Microbiol*; **37**:206–8.

Hall, N. (2007). "Advanced sequencing technologies and their wider impact in microbiology". *J. Exp. Biol.* **210** (Pt 9): 1518–25.

Hara-Kudo, Y., Nemoto, J., Ohtsuka, K., Segawa, Y., Takatori, K., Kojima, T. & Ikedo, M. (2007). Sensitive and rapid detection of Vero toxinproducing *Escherichia coli* using loop-mediated isothermal amplification. *J. Med. Microbiol*. **56**:398-406.

Hatano, B., Maki, T., Obara, T., Fukumoto, H., Hagisawa, K., Matsushita, Y., ... Katano, H. (2010). LAMP Using a Disposable Pocket Warmer for Anthrax Detection, a Highly Mobile and Reliable Method for Anti-bioterrorism. *Jpn. J. Infect. Dis*., **63**, 36-40.

Hayman, J. (1991). Postulated epidemiology of *Mycobacterium ulcerans* infection. *Int J Epidemiol*, **20**: 1093-8.

Herbinger, K., Adjei, O., Awua-Boateng, N., Nienhuis, W.A., Kunaa, L., Siegmund, V., ... Bretzel, G. (2009). Comparative Study of the Sensitivity of Different Diagnostic Methods for the Laboratory Diagnosis of Buruli Ulcer Disease. *Clin Infect Dis*, **48**:1055–1064.

Hilty, M., Yeboah-Manu, D., Boakye, D., Mensah-Quainoo, E., Rondini, S., Schelling, E., ... Pluschke, G. (2006). Genetic diversity in *Mycobacterium ulcerans* isolates from Ghana revealed by a newly identified locus containing a variable number of tandem repeats. *J Bacteriol*.; **188**:1462–5.

Huber, C.A., Ruf, M-T., Pluschke, G. & Kaser, M. (2008). Independent loss of immunogenic proteins in *Mycobacterium ulcerans* suggests immune evasion. *Clinical and Vaccine Immunology*: CVI 15: 598–606.

Hunkapiller, T., Kaiser, R.J., Koop, B.F. & Hood, L. (1991). Large-scale and automated DNA sequence determination. *Science* **254**, 59–67.

Hunter, R. C. & Beveridge, T. J. (2005). High-resolution visualization of *Pseudomonas aeruginosa* PAO1 biofilms by freeze-substitution transmission electron microscopy. *J Bacteriol* **187**: 7619–7630.

Huson, D.H. (1998). SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics*, 14(1):68–73.

Ion Torrent. Schematic cross-section of a single well of an Ion Torrent sequencing chip. Retrieved from http://www.iontorrent.com/lib/images/PDFs/amplicon_application_note_040411.pdf

Janssens, P.G., Pattyn, S.R., Meyers, W.M. & Portaels, F. (2005). Buruli ulcer: an historical overview with updating to 2005. *Bull. Seances Acad. R. Sci. Outre Mer*. 2005; 51:165–99.

Johnson, P.D.R., Azuolas, J., Lavender, C.J., Wishart, E., Stinear, T.P., Hayman, ... Fyfe, J.A.M. (2007). *Mycobacterium ulcerans* in Mosquitoes Captured during Outbreak of Buruli Ulcer, Southeastern Australia. *Emerging Infectious Diseases*, **13** (11): 1653-1660.

Johnson, R. C., Makoutode, M., Sopoh, G. E., Elsen, P., Gbovi, J., Pouteau, ... Portaels, F. (2005). Buruli ulcer distribution in Benin. *Emerg Infect Dis*., **11**:500–1.

Kaneko, H., Kawana, T.& Fukushima, E. (2007). Tolerance of loop-mediated isothermal amplification to a culture medium and biological substances. *J Biochem Biophys Methods* **70**: 499–501.

Karl, V. V., Shale, A. D. & Jacob, D. D. (2009). "Next Generation Sequencing: From Basic Research to Diagnostics". *Clinical Chemistry* **55** (4): 41–47.

Kaser, M., Gutmann, O., Hauser, J., Stinear, T., Cole, S., Yeboah-Manu, D., ... Pluschke, G. (2009a). Lack of Insertional-Deletional Polymorphism in a Collection of *Mycobacterium ulcerans* Isolates from Ghanaian Buruli Ulcer Patients. *J Clin. Microbiol*, **47** (11): 3640–3646.

Kaser, M & Pluschke, G. (2008). Differential Gene Repertoire in *Mycobacterium ulcerans* Identifies Candidate Genes for Patho-Adaptation. *PLoS Negl Trop Dis* 2: e353. doi:10.1371/journal.pntd.0000353.

Kaser, M., Rondini, S., Naegeli, M., Stinear, T., Portaels, F, Certa, U. & Pluschke, G. (2007). Evolution of two distinct phylogenetic lineages of the emerging human pathogen *Mycobacterium ulcerans*. *BMC Evolutionary Biology* **7**: 177–177.

Keim, P., Price, L. B., Klevytska, A. M., Smith, K. L., Schupp, J. M., Okinaka, R., ... Hugh-Jones, M. E. (2000). Multiple-locus-variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J. Bacteriol.* **182**:2928–2936.

Kibadi, K., Panda, M., Tamfum, J.J.M., Fraga, A.G., Filho, A.L., Anyo, G., ... Portaels, F. (2008). New Foci of Buruli Ulcer, Angola and Democratic Republic of Congo. *Emerg Infect Dis.* **14** (11): 1790-1792.

Klevytska, A., Price, L. B., Schupp, J. M., Worsham, P. L., Wong, J. & Keim, P. (2001). Identification and characterization of variable-number tandem repeats in *Yersinia pestis* genome. *J. Clin. Microbiol.,* **39**:3179–3185.

Kotlowski, R., Martin, A., Ablordey, A., Chemlal, K., Fonteyne, P. A., Fonteyne, P. & Portaels, P. (2004). One-tube cell lysis and DNA extraction procedure for PCR-based detection of *Mycobacterium ulcerans* in aquatic insects, mollusks and fish. *J Med Microbiol* **53**: 927–933.

Kuboki, N., Inoue, N., Sakurai, T., Cello, F.D., Grab, D. J., Suzuki, H., ... Igarashi, I. (2003). Loop-Mediated Isothermal Amplification for Detection of African Trypanosomes. *J. CLIN. MICROBIOL*., Vol. 41, No. 12p. 5517–5524.

Levy-Frebault, V.V. & Portaels, F. (1992). Proposed Minimal Standards for the genus *Mycobacterium* and for description of new slowly growing *Mycobacterium* Species. *Int. J. Syst. Bacteriol*.; **42**:315-23.

Lehmann, K.B. & Neumann, R. Atlas und Grundriss der Bakteriologie und Lehrbuch der speziellen bakteriologischen Diagnostik, First Edition, J.F. Lehmann, München, 1896, p. 1-448.

Liu, L., Li, Y., Li, S., Hu, N., He, Y., Pong, R., ... Law, M. (2012). "Comparison of Next-Generation Sequencing Systems". *Journal of Biomedicine and Biotechnology* **2012**: 1–11.

Livanainen, E. (1995). Isolation of mycobacteria from acidic forest soil samples: comparison of culture methods. *J Appl Bacteriol.*, **78**(6):663-8.

MacCallum, P., Tolhurst, J. C., Buckle, G. & Sissons, H. A. (1948). A new mycobacterial infection in man. *J Path Bacteriol*, **60**:93-122.

Marsollier, L., Brodin, P., Jackson, M., Kordula´kova, J., Tafelmeyer, P., Carbonnelle, E. & Cole, S. T. (2007). Impact of *Mycobacterium ulcerans* biofilm on transmissibility to ecological niches and Buruli ulcer pathogenesis. *PloS Pathog*, 3: 0582 – 0594.

Marsollier, L., Stinear, T., Aubry, J., Saint-Andre´, J. P., Robert, R., Legras, P., ... Carbonnelle, B. (2004). Aquatic plants stimulate the growth of and biofilm formation by Mycobacterium ulcerans in axenic culture and harbor these bacteria in the environment. *Appl Environ Microbiol* **70**: 1097–1103.

Marsollier, L., Robert, R., Aubry, J., Saint Andre, J. P., Kouakou, H., Legras, ……. Carbonnelle, B. (2002). Aquatic insects as a vector for *Mycobacterium ulcerans*. *Appl Environ Microbiol*, **68**, 4623–4628.

Marston, B.J., Diallo, M.O., Horsburgh, C.R. Jr, Diomande, I., Saki, M.Z., Kanga, J.M., ... Good, R.C. (1995). Emergence of Buruli ulcer disease in the Daloa region of Côte d'Ivoire. *Am J Trop Med Hyg*, **52**:219-24.

Maxam, A.M. & Gilbert, W. (1977). "A new method for sequencing DNA". *Proc. Natl. Acad. Sci. U.S.A.* **74** (2): 560–4.

Mazars, E., Lesjean, S. Banuls, A., Gilbert, M., Vincent, V., Gicquel, B. ..... Supply, P. (2001). High resolution minisatellitebased typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology. *Proc. Natl. Acad. Sci*. USA **98**:1901–1906.

Mensah-Quainoo, E., Yeboah-Manu, D., Asebi, C., Patafuor, F., Ofori-Adjei, D., Junghanss, T. & Pluschke, G. (2008). Diagnosis of *Mycobacterium ulcerans* infection (Buruli ulcer) at a treatment centre in Ghana: a retrospective analysis of laboratory results of clinically diagnosed cases. *Trop Med Int Health*; **13**:191–8.

Metzker, M. L. (2010). "Sequencing technologies - the next generation". *Nat Rev Genet* **11** (1): 31–46.

Meyers, W.M., Tignokpa, N., Priuli, G.B. & Portaels, F. (1996). *M. ulcerans* infection (Buruli ulcer): First reported patients in Togo. *Br. J. Dermatol*.; **134**: 1116-1121.

Meyers, W.M., Connor, D.H., McCullough, B., Bourland, J., Moris, R. & Proos, L. (1974a). Distribution of *M. ulcerans* infections in Zaïre, including the report of new foci. *Ann Soc Belg Med Trop*, **54**:147–57.

Meyers, W. M., Shelly, W. M., Connor, D. H. & Meyers, E. K. (1974b). Human *Mycobacterium ulcerans* infections developing at sites of trauma to skin. *Am J Trop Med Hyg*, **23:** 919 – 23.

Mitra, R.D., Shendure, J., Olejnik, J., Edyta Krzymanska, O. & Church, G.M. (2003). Fluorescent in situ sequencing on polymerase colonies. *Anal. Biochem.* **320**, 55–65.

Mitra, R.D. & Church, G.M. (1999). In situ localized amplification and contact replication of many individual DNA molecules. *Nucleic Acids Res.* **27**, e34.

Mori, Y., Nagamine, K., Tomita, N. & Notomi, T. (2001). Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. *Biochem Biophys Res Commun* **289**, 150–154.

Municipal Chief Executive (2012). Asante Akim North Municipal Assembly. Available at: http://www.asanteakimnorth.ghanadistricts.gov.gh/

National Library of Medicine (2009). "Medical Subject Headings (MeSH) Fact sheet" on Tandem repeats. Available at: http://www.nlm.nih.gov/pubs/factsheets/mesh.html.

Nagamine, K., Hase, T. & Notomi, T. (2002). Accelerated reaction by loop-mediated isothermal amplification using loop primers. *Molecular and Cellular Probes*, Volume 16, Issue 3, pp 223-229.

N'Guessan, K., Kouassi, Y., Bouzid, S., Ehuie, P., Koffi, K., Oniangue, C., ... Dosso, M. (2001). [Value and limits of microscopy of exudates in *Mycobacterium ulcerans* cutaneous infection in Cote d'Ivoire]. *Bulletin de la Societe de Pathologie Exotique* **94**, 9–10.

Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N. & Hase, T. (2000). Loop-mediated isothermal amplification of DNA. *Nucleic Acids Res* **28**, E63. E63.

Oluwasanmi, J.O., Solanke, T.F., Itayemi, S.O. & Alabi, G.O. (1975). Buruli (Mycobacterial) skin ulcers in Caucasians in Nigeria. *Br. J. of Plast. Surg.*: **28**: 111-3.

Ouattara, D., Meningaud, J. P., Kaba, L., Sica, A. & Asse, H. (2004). Treatment of Buruli ulcer desease by excision and skin graft. *Ann Chir Plast Esthet* **49**: 11–16.

Palomino, J. C., Obiang, A. M., Realini, L., Meyers, W.M. & Portaels, F. (1998). Effect of oxygen on growth of Mycobacterium ulcerans in the BACTEC system. *J Clin Microbiol*; **36**(11): 3420-2.

Pattyn, S. R. **(**1965). Bacteriology and human and experimental pathology of ulcers caused by *Mycobacterium ulcerans*. *Ann. Soc. Belg. Med. Trop*. **45:** 121–129.

Pennisi, E. (2010). "Semiconductors inspire new sequencing technologies". Science **327**(5970): 1190.

Pettersson, E., Lundeberge, J. & Ahmadian, A. (2008).Generation of sequencing technologies. *Genomics*, pp. 105-111.

Perkel, J. (2011). "Making contact with sequencing's fourth generation". Biotechniques, 2011.

Picardeau, M. & Vincent, V. (1996). Typing of *Mycobacterium avium* isolates by PCR. *J. Clin. Microbiol*. **34**:389–392.

Pidot, S.J., Asiedu, K., Kaser, M., Fyfe, J.A.M. & Stinear, T.P. (2010). *Mycobacterium ulcerans* and Other Mycolactone-Producing Mycobacteria Should Be Considered a Single Species. *Plos Negl Trop Dis*, **4**: 1-3.

Portaels, F., Meyers, W.M., Ablordey, A., Castro, A.G., Chemlal, K., de Rijk, P., ... Pedrosa, J. (2008). First Cultivation and Characterization of *Mycobacterium ulcerans* from the Environment. *PLoS Negl Trop Dis*, 2 (3): e178.

Portaels, F. & Meyers, W.M. (2006). Buruli ulcer. In: Faber, W.R., Hay, R.J., Naafs, B., editors. Imported skin diseases. Maarssen, The Netherlands: Elsevier Gezondheidszorg. pp. 117–129.

Portaels, F., Chemlal, K., Elsen, P., Johnson, P. D., Hayman, J. A., Hibble, J. & Meyers, W. M. (2001). *Mycobacterium ulcerans* in wild animals. *Rev Sci Tech*, **20**, 252–264.

Portaels, F., Elsen, P., Guimaraes-Peres, A., Fonteyne, P. A. & Meyers, W.M. (1999). Insects in the transmission of Mycobacterium ulcerans infection (Buruli ulcer). *Lancet* **353**: 986.

Portaels, F., Fonteyne, P.A., de Beenhouwer, H., de Rijk, P., Guedenon, A., Hayman, J. & Meyers, M.W. (1996). Variability in 39 end of 16S Rrna sequence of *Mycobacterium ulcerans* is related to geographic origin of isolates. *J. Clin. Microbiol*. **34**:962–965.

Portaels, F. (1995). Epidemiology of mycobacterial diseases. *Clin Derm*, **13**: 207-22.

Portaels, F., De Muynck, A. & Sylla, M. P. (1988). Selective Isolation of Mycobacteria from Soil: a Statistical Analysis Approach, *Journal of General Microbiology*, **134**, 849-855.

Qi, W., Kaser, M., Roltgen, K., Yeboah-Manu, D. & Pluschke, G. (2009). Genomic diversity and evolution of *Mycobacterium ulcerans* revealed by next-generation sequencing. *PLoS Pathog* 5: e1000580.

Quail, M., Smith, M. E., Coupland, P., Otto, T. D., Harris, S. R, Connor, T. R., & Gu, Y. (2012). "A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers". *BMC Genomics* **13** (1): 341.

Raghunathan, P.L., Whitney, E.A., Asamoa, S., Stienstra, Y., Taylor Jr, T.H., Amofah, G.K., ... Ashford, D.A. (2005). Risk factors for Buruli ulcer disease (*Mycobacterium ulcerans* infection): Results from a Case–Control study in Ghana. *Clin Infect Dis*, **40**:1445-53.

Ravisse, P. (1977). L'ulcère cutane à Mycobacterium ulcerans au Cameroun. (Skin ulcers caused by *Mycobacterium ulcerans* in Cameroon. I. Clinical, epidemiological and historical study). Bull. Soc. Pathol. Exot. Filiales; **70**: 109–124.

Roltgen, K., Qi, W., Ruf, M., Mensah-Quainoo, E., Pidot, S. J., Seemann, T., ... Pluschke, G. (2010). Single Nucleotide Polymorphism Typing of *Mycobacterium ulcerans* Reveals Focal Transmission of Buruli Ulcer in a Highly Endemic Region of Ghana. *PLoS Negl Trop Dis*, 4 (7): e751.

Roring, S., Scott, A., Brittain, D., Walker, I., Hewinson, G., Neill, S. & Skuse, R. (2002). Development of variable-number tandem repeat typing of *Mycobacterium bovis*: comparison of results with those obtained by using existing exact tandem repeats and spoligotyping. *J. Clin. Microbiol*. **40**:2126–2133.

Rondini, S., Mensah-Quainoo, E., Troll, H., Bodmer, T. & Pluschke, G. (2003). Development and application of real-time PCR assay for quantification of *Mycobacterium ulcerans* DNA. *J Clin Microbiol*; **41**:4231–7.

Ross, B. C., Johnson, P. D. R., Oppedisano, F., Marino, L., Sievers, A., Stinear, T., ... Robins-Browne, R. M. (1997). Detection of *Mycobacterium ulcerans* in environmental samples during an outbreak of ulcerative disease. *Appl Environ Microbiol* **63**: 4135–4138.

Ross, B. C., Raios, K., Jackson, K. & Dwyer, B. (1992a). Molecular cloning of a highly repeated DNA element from *Mycobacterium tuberculosis* and its use as an epidemiological tool. *J. Clin. Microbiol*. **30:**942–946.

Ross, B. C., Jackson, K., Yang, M., Sievers, A. & Dwyer, B. (1992b). Identification of a genetically distinct subspecies of *Mycobacterium kansasii*. *J. Clin. Microbiol*. **30:**2930–2933.

Ross, B.C., Marino, L., Oppedisano, F., Edwards, R., Robins-Browne, R.M. & Johnson, P.D. (1997). Development of a PCR assay for rapid diagnosis of *Mycobacterium ulcerans* infection. *J Clin Microbiol*; **35**:1696–700.

Roberts, B. & Hirst, R. (1997). Immunomagnetic separation and PCR for detection of *Mycobacterium ulcerans*. *J Clin Microbiol* **35**: 2709–2711.

Rusk, N. (2011). "Torrents of sequence". *Nat Meth* 8(1): 44-44.

Ryan, K.J. & Ray, C.G. (2004). *Sherris Medical Microbiology* (4th ed.). McGraw Hill.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, ... Erlich, H.A. (1988*).* Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487–491.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A. & Arnheim, N. (1985). Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230(4732):1350-4.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977). "DNA sequencing with chain-terminating inhibitors". *Proc. Natl. Acad. Sci. U.S.A.* **74** (12): 5463–7.

Sanger, F. & Coulson, A.R. (1975). "A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase". *J. Mol. Biol.* **94** (3): 441–8.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., ... Smith, M. (1977 b). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature* **265**, 687–695.

Shendure, J. & Ji, H. (2008). Next-generation DNA sequencing. *Nature biotechnology*, **26** (10): 1135 – 1145.

Swerdlow, H., Wu, S.L., Harke, H. & Dovichi, N.J. (1990). Capillary gel electrophoresis for DNA sequencing. Laser-induced fluorescence detection with the sheath flow cuvette. *J. Chromatogr.* **516**, 61–67.

Shiau, A. L. & Wu, C. L. (1998). The inhibitory effect of *Staphylococcus epidermidis* slime on the phagocytosis of murine peritoneal macrophages is interferon independent. *Microbiol Immunol* **42**: 33–40.

Shinnick, T.M. & Good, R.C. (1994). Mycobacterial Taxonomy. *Eur. J. Clin. Microbiol. Infect.Dis.*, p. 884-901.

Siegmund, V., Adjei, O., Nitschke, J., Thompson, W., Klutse, E., Herbinger, K.H. ..... Bretzel, G. (2007). Dry reagent–based polymerase chain reaction compared with other laboratory methods available for the diagnosis of Buruli ulcer disease. *Clin Infect Dis*; **45**:68–75.

Siegmund V, Adjei O, Racz P, Berberich, C., Klutse, E., van Vloten, F., ... Bretzel, G. (2005). Dry reagent-based PCR as a novel tool for laboratory confirmation of clinically diagnosed *Mycobacterium ulcerans*–associated disease in areas in the tropics where *M. ulcerans* is endemic. *J Clin Microbiol*; **43**:271–6.

Sizaire, V., Nackers, F., Comte, E. & Portaels, F. (2006). *Mycobacterium ulcerans* infection: control, diagnosis, and treatment. *Lancet Infect. Dis*. **6:**288–296.

Skuce, R. A., MacCorry, T. P., McCarroll, J. F., Roring, S. M. M., Scott, A. N., Brittain, D., ... Neill, S. D. (2002). Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTRPCR targets. *Microbiology* **148**:519–528.

Smith, J. H. (1970). Epidemiologic observations on cases of Buruli ulcer seen in a hospital in the lower Congo. *Am J Trop Med Hyg*., **19**:657–63.

Smith, P. G., Revill, W. D., Lukawgo, E. & Rykushin, Y. P. (1997). The protective effect of BCG against *Mycobacterium ulcerans* disease: a controlled trial in an endemic area of Uganda. *Trans R Soc Trop Med Hyg*., **70**:449–57.

Smittipat, N., & Pallittapongarnpim, P. (2000). Identification of possible loci of variable number of tandem repeats in *Mycobacterium tuberculosis*. *Tuber. Lung Dis*. **80**:69–74.

Snyder, D. S. & Small, P. L. (2003). Uptake and cellular actions of mycolactone, a virulence determinant for *Mycobacterium ulcerans*. *Microb Pathog*; **34:** 91–101.

Sreevatsan, S., Pan, X., Stockbauer, K.E., Connell, N.D., Kreiswirth, B.N., Whittam, T.S. & Musser**,** J.M. (1997). Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination. *Proc. Natl. Acad. Sci*. USA **94:**9869–9874.

Stienstra, Y., van der Werf, T.S., Guarner, J., Raghunathan, P.L., Whitney, E.A.S., van der Graaf, W.T.A., ... King, C.H. (2003). Analysis of an IS2404- based nested PCR for diagnosis of Buruli ulcer disease in regions of Ghana where the disease is endemic. *J Clin Microbiol*; **41**:794–7.

Stienstra, Y., van der Graaf, W. T. A., Asamoa, K. & van der Werf, T. S. (2002). Beliefs and attitudes towards Buruli ulcer Ghana. *Am J Trop Med Hyg* **67**: 207–213.

Stinear, T.P., Seemann, T., Pidot, S., Frigui, W., Reysset, G., Garnier, T., ... Cole, S.T. (2007). Reductive evolution and niche adaptation inferred from the genome of *Mycobacterium ulcerans*, the causative agent of Buruli ulcer. *Genome Res* **17**: 192–200.

Stinear, T.P., Hong, H., Frigui, W., Pryor, M.J., Brosch, R., Garnier, T., ... Cole, S.T. (2005a). Common evolutionary origin for the unstable virulence plasmid pMUM found in geographically diverse strains of *Mycobacterium ulcerans*. *J. Bacteriol.*, **187**:1668–1676.

Stinear, T.P., Pryor, M.J., Porter, J.L. & Cole, S.T. (2005b). Functional analysis and annotation of the virulence plasmid pMUM001 from *Mycobacterium ulcerans*. *Microbiol.*, **151**: 683–692.

Stinear, T.P., Mve-Obiang, A., Small, P.L., Frigui, W., Pryor, M.J., Brosch, R., ... Cole, S.T. (2004). Giant plasmid-encoded polyketide synthases produce the macrolide toxin of *Mycobacterium ulcerans*. *Proc. Natl. Acad. Sci.*, **101**:1345–1349.

Stinear, T., Davies, J. K., Jenkin, G. A., Hayman, J. A., Oppedisano, F. & Johnson, P. D. R. (2000). Identification of *Mycobacterium ulcerans* in the environment from regions in Southeast Australia in which it is endemic with sequence capture-PCR. *Appl Environ Microbiol* **66**: 3206–3213.

Stinear, T. P., G. A. Jenkins, P. D. R J. Johnson, & J. K. Davis (2000). Comparative genetic analysis of *Mycobacterium ulcerans* and *Mycobacterium marinum* reveals evidence of recent divergence. *J. Bacteriol*., **182:**6322–6330.

Stinear, T. P., Davis, J. K., Jenkins, G. A., Portaels, F., Ross, B. C., Oppedisano, F., ... Johnson, P. D. R. (2000). A simple PCR method for rapid genotype analysis of *Mycobacterium ulcerans*. *J. Clin. Microbiol*., **38:**1482–1487.

Stinear, T., Ross, B.C., Davies, J.K., Marino, L., Robins-Browne, R.M., Oppedisano, F., ... Johnson, P.D.R. (1999). Identification and characterization of IS*2404* and IS*2606*: two distinct repeated sequences for detection of *Mycobacterium ulcerans* by PCR. J. Clin. Microbiol. **37:**1018–1023.

Stragier, P., Ablordey, A., Durnez, L. & Portaels, F. (2007). VNTR analysis differentiates *Mycobacterium ulcerans* and IS*2404* positive mycobacteria. *Syst Appl Microbiol* **30**: 525–530.

Stragier, P., Ablordey, A., Bayonne, L. M., Lugor, Y. L., Sindani, I. S., Suykerbuyk, P., ... Portaels, F. (2006). Heterogeneity among *Mycobacterium ulcerans* isolates from Africa. *Emerg Infect Dis* **12**, 844–847.

Stragier, P., Ablordey, A., Meyers, W. & Portaels, F. (2005). Genotyping *Mycobacterium ulcerans* and *Mycobacterium marinum* by using mycobacterial interspersed repetitive units. *J Bacteriol* **187**: 1639–1647.

Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B. & Locht, C. (2000). Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. *Mol. Microbiol*. **39**:3563–3571.

Thekisoe, O. M. M., Bazie, R. S. B., Coronel-Servian, A. M., Sugimoto, C. Kawazu, S. & Inoue, N. (2009). Stability of Loop-Mediated Isothermal Amplification (LAMP) Reagents

and its Amplification Efficiency on Crude Trypanosome DNA Templates. *J. Vet. Med. Sci.* 71(4): 471–475.

Turcatti, G., Romieu, A., Fedurco, M. & Tairi, A.P. (2008).A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis. *Nucleic Acids Res.* **36**, e25.

Uganda Buruli Group (1971). Epidemiology of *Mycobacterium ulcerans* infection (Buruli ulcer) at Kinyara, Uganda. *Trans. R. Soc. Trop. Med. Hyg.*; **65**: 763-75.

Uganda Buruli Group (1969). BCG vaccination against *Mycobacterium ulcerans* infection (Buruli ulcer). First results of a trial in Uganda. *Lancet*; 7586: 111-115.

Uganda Buruli Group (1970). Clinical features and treatment of preulcerative Buruli lesions (*Mycobacterium ulcerans* infection): Report II. *Br. Med. J.* **2**:390-393.

Vandamme, P., Pot, B., Gillis, M., de Vos, P., Kersters, K. & Swings, J (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* 60: 407–438.

van der Werf, T.S., Stienstra, Y., Johnson, R.C., Phillips, R., Adjei, O., Fleischer, B., Wansbrough-Jones, M.H., Johnson, P.D.R., Portaels, F., van der Graaf, W.T.A., Asiedu, K. (2005). *Mycobacterium ulcerans* disease. *Bull World Health Organ*, **83** (10):785 – 791.

van der Werf, T. S., van der Graaf, W. T. A., Tappero, J. W. & Asiedu, K. (1999). *Mycobacterium ulcerans* infection. *Lancet,* **354***:* 1013–1018.

van der Werf, T.S., van der Graaf, W.T.A., Groothuis, D.G. & Knell, A.J. (1989). Mycobacterium ulcerans infection in Ashanti region, Ghana. *Trans R Soc Trop Med Hyg*; **83**:410-3.

van Oye, E, & Ballion, M. (1951). A possible new condition due to acid-fast organisms in Africa. *Ann Soc Belg Med Trop*; 619-27.

Wang, L., Shi, L., Alam, M.J., Geng, Y. & Li, L. (2008). Specific and rapid detection of foodborne *Salmonella* by loop-mediated isothermal amplification method. *Food Research International* **41**: 69–74.

Ward, K. H., Olson, M. E., Lam, K. & Costerton, J. W. (1992). Mechanism of persistent infection associated with peritoneal implants. *J Med Microbiol* **36**: 406–413.

Werner, E., Roe, F., Bugnicourt, A., Franklin, M. J., Heydorn, A., Molin, S., ... Stewart, P.S. (2004). Stratified growth in *Pseudomonas aeruginosa* biofilms. *Appl Environ Microbiol*, **70**: 6188–6196.

Whitney, E.A.S., Phelan, M., Raghunathan, P.L., Stienstra, Y., Dobos, K., Guarner, J., ... Ashford, D.A. (2002) Latent class analysis (LCA) evaluation of four diagnostic tests for buruli ulcer disease. 5th WHO Advisory Group Meeting on Buruli Ulcer. WHO, Geneva.

Williamson, H.R., Benbow, M.E., Nguyen, K.D., Beachboard, D.C., Kimbirauskas, R.K., Merritt, R.W., ..., Small, P.L.C. (2008). Distribution of *Mycobacterium ulcerans* in

Buruli ulcer endemic and non-endemic aquatic sites in Ghana. *PLoS Negl Trop Dis* **2**: e205.

World Health Organization (2015). Buruli ulcer (*Mycobacterium ulcerans* infection). Available at: *http://www.who.int/mediacentre/factsheets/fs199/en/*

World Health Organization (2014). Laboratory Diagnosis of Buruli Ulcer: A Manual for Health-Care Providers. Available at: *http://www.who.int/buruli/laboratory_diagnosis/en/*

World Health Organisation (WHO, 2010). Guidance on sampling techniques for laboratory-confirmation of *Mycobacterium ulcerans* infection (Buruli ulcer disease). Available at: www.who.int/entity/buruli/Guidance_sampling_techniques_MU_infection.pdf

World Health Organisation (2009). Cotonou Declaration on Buruli ulcer, *Cotonou, Benin, 30 March 2009.*

World Health Organization (2008a) Buruli ulcer: progress report, 2004–2008. *Weekly epidemiological record* **17**: 145–154.

World Health Organization **(**2008b). Buruli ulcer: progress report, 2004–2008. *Wkly. Epidemiol. Rec*. **83:**145–156.

World Health Organisation Media Centre (2007). Buruli ulcer disease (*Mycobacterium ulcerans* infection). Revised March 2007. Fact sheet N°199. Available at: http://www.who.int/mediacentre/factsheets/fs199/en/

World Health Organization/ Global Buruli Ulcer Initiative (2005). Buruli ulcer: a pocket guide for community health workers. Available at: http://www.who.int/buruli/en

WHO/Global Buruli Ulcer Initiative – Map: after Public Health Mapping & GIS, WHO Communicable Diseases 2005.

World Health Organization (2004). Provisional guidance on the role of specific antibiotics in the management of *Mycobacterium ulcerans* disease (Buruli ulcer). Geneva: WHO, **2004.**

World Health Organisation (2004). Resolution WHA57.1 Surveillance and control of *Mycobacterium ulcerans* disease (Buruli ulcer). In: *Fifty-seventh World Health Assembly, Geneva, 17–22 May 2004. Resolutions and decisions.* Geneva, World Health Organization, 2004 (WHA57/2004/REC/1):1–2.

World Health Organisation (2001). Buruli Ulcer. Diagnosis of Mycobacterium ulcerans disease. WHO, Geneva.

World Health Organization (2000). Buruli ulcer-*Mycobacterium ulcerans* infection. In: Asiedu K, Scherpbier, R, Raviglione, M., (editors). WHO/CDS/CPE/GBUI/2000.1. Geneva: The Organization; 2000. p.9–14.

World Health Organization. Buruli ulcer Photos. Available at: *http://www.who.int/buruli/photos/en/index.html*

Wünsch-Filho, V., Moncau, J, E. & Nakao, N. (1993). Methodological considerations in case-control studies to evaluate BCG vaccine effectiveness. *Int J Epidemiol*., **22**:149–55.

Yeboah-Manu, D., Bodmer, T., Mensah-Quainoo, E., Owusu, S., Ofori-Adjei, D. & Pluschke, G. **(**2004). Evaluation of decontamination methods and growth media for the primary isolation of *Mycobacterium ulcerans* from surgical specimens. *J. Clin. Microbiol*. **42:**5875–5876.

Yip, M.J., Porter, J.L., Fyfe, J.A.M., Lavender, C.J., Portaels, F., Jenkin, G. A., ... Stinear, T. (2007). Evolution of *Mycobacterium ulcerans* and other mycolactone-producing mycobacteria from a common *Mycobacterium marinum* progenitor. *Journal of Bacteriology* 189: 2021–2029.

Yoder, S., Argueta, C., Holtzman, A., Aronson, T., Berlin, O.G.W., Tomasek, P., ... Stelma, G. Jr. (1999). PCR comparison of *Mycobacterium avium* isolates obtained from patients and foods. *Appl. Environ. Microbiol*. **65:**2650–2653.

APPENDICES

Appendix I: Preparation of the amplification solution

The reagents for the amplification solution were prepared, based on manufacturer's information provided in the Ion PGM™ Template OT2 200 Kit User Guide, as followed.The Ion PGM™ Template OT2 200 reagent mix was allowed to come to room temperature, vortexed for 30 seconds and centrifuged for 2 seconds. The reagent mix was kept at room temperature during use and stored at 2°C to 8°C. The Ion PGM™ Template OT2 200 PCR Reagent B was vortexed for 1 minute, centrifuged for 2 seconds and inspected to ensure that the solution was clear and kept at room temperature. The suspension of Ion PGM™ Template OT2 200 Ion Sphere™ Particles was placed at room temperature.

Appendix II: Washing and resuspension of Dynabeads® MyOne™ Streptavidin C1Beads

The tube containing the Dynabeads® MyOne™ Streptavidin C1 Beads was vortexed for 30 seconds to thoroughly resuspend the beads, centrifuged for 2 seconds and the dark pellet of beads pipetted up and down dispersed. Thirteen (13) µL of Dynabeads® MyOne™ Streptavidin C1 Beads was transferred to a new 1.5-mL Eppendorf LoBind® Tube. The tube was placed on a DynaMag™-2 magnet for 2 minutes and the supernatant carefully removed and discarded without disturbing the pellet of the beads followed by the addition of 130 µl of MyOne™ Beads Wash Solution to the Dynabeads® MyOne™ Streptavidin C1 Beads. The tube was then removed from the magnet, vortexed for 30 seconds and centrifuged for 2 seconds.

Appendix III: Ion 316™ chip check

The chip check tested the chip to ensure proper functioning prior to loading the sample. The chip was labelled to identify the experiment and placed on the PGM™ Sequencer grounding plate. Using a Rainin^{R} SR-L200F pipette tip, 100 µl of 100% isopropanol was slowly and steadily pipetted into the large port of the chip and any displaced solution from the other port aspirated. The chip was washed two times with 100 µl of Annealing Buffer by slowly and steadily pipetting it into the large port on the chip and aspirating any displaced solution from the other port. The chip for each sample was checked on the instrument following scanning of the chip barcode located on the chip package using the barcode scanner, as a chip cannot be run without scanning or entering the barcode. After successful chip check and calibration, the chip was removed and 100 µl of annealing buffer pipetted into the large port on the chip followed by aspiration of any displaced solution from the other port.

Appendix IV: Summary statistics for *Mycobacterium ulcerans* sequenced isolates

Isolate	Summary statistics					
	Sequence parameter	**Count**	**Percentage of reads (%)**	**Average length**	**Number of bases**	**Percentage of bases (%)**
Mu_06-3845	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	13,018,332	95.67	35.98	468,439,948	95.67
	Not mapped reads	589,294	4.33	35.96	21,189,325	4.33
	Reads in pairs	12,450,212	91.49	188.45	447,999,852	91.50
	Broken paired reads	568,120	4.18	35.98	20,440,096	4.17
	Total reads	13,607,626	100.00	35.98	489,629,273	100.00
Mu_NM33.04	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	24,837,745	97.25	31.70	787,305,172	96.96
	Not mapped reads	703,525	2.75	35.11	24,698,306	3.04
	Reads in pairs	20,967,478	82.09	537.69	669,173,213	82.41
	Broken paired reads	3,870,267	15.15	30.52	118,131,959	14.55
	Total reads	25,541,270	100.00	31.79	812,003,478	100.00
Mu_NM43.02	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	30,098,279	97.34	32.44	976,525,047	97.12
	Not mapped reads	821,883	2.66	35.29	29,007,700	2.88
	Reads in pairs	24,514,860	79.28	214.17	802,044,601	79.76
	Broken paired reads	5,583,419	18.06	31.25	174,480,446	17.35
	Total reads	30,920,162	100.00	32.52	1,005,532,747	100.00
Mu_NM14.01	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	30,162,918	97.19	32.22	971,893,819	96.93
	Not mapped reads	870,802	2.81	35.31	30,747,976	3.07

	Reads in pairs	24,803,240	79.92	227.96	805,673,960	80.36
	Broken paired reads	5,359,678	17.27	31.01	166,219,859	16.58
	Total reads	31,033,720	100.00	32.31	1,002,641,795	100.00
Mu_NM54.02	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	28,554,059	97.35	29.87	852,846,250	96.92
	Not mapped reads	777,117	2.65	34.93	27,140,919	3.08
	Reads in pairs	25,005,904	85.25	156.99	758,002,987	86.14
	Broken paired reads	3,548,155	12.10	26.73	94,843,263	10.78
	Total reads	29,331,176	100.00	30.00	879,987,169	100.00
Mu_06-3846	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	8,820,473	96.38	21.56	189,862,848	94.34
	Not mapped reads	331,179	3.62	34.39	11,389,584	5.66
	Reads in pairs	1,931,288	21.10	11,162,80	46,978,558	23.34
	Broken paired reads	6,889,185	75.28	20.74	142,884,290	71.00
	Total reads	9,151,652	100.00	21.99	201,252,432	100.00
Mu_07-1082	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	11,222,261	94.48	35.99	403,900,342	94.57
	Not mapped reads	655,931	5.52	35.33	23,172,219	5.43
	Reads in pairs	10,682,878	89.94	503.02	384,490,222	90.03
	Broken paired reads	539,383	4.54	35.99	19,410,120	4.54
	Total reads	11,878,192	100.00	35.95	427,072,561	100.00
Mu_NM49.02	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	28,606,620	96.98	32.96	942,863,737	97.03
	Not mapped reads	890,480	3.02	32.36	28,812,322	2.97
	Reads in pairs	23,395,918	79.32	804.17	778,006,576	80.07

	Broken paired reads	5,210,702	17.67	31.64	164,856,797	16.97
	Total reads	29,497,100	100.00	32.94	971,675,695	100.00
Mu_F64	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	3,304,625	95.73	191.70	633,483,852	95.72
	Not mapped reads	147,397	4.27	192.31	28,345,668	4.28
	Total reads	3,452,022	100.00	191.72	661,829,520	100.00
Mu_F65	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	3,281,977	95.98	191.13	627,278,983	95.70
	Not mapped reads	137,554	4.02	205.05	28,205,022	4.30
	Total reads	3,419,531	100.00	191.69	655,484,005	100.00
Mu_F70	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	3,165,967	96.80	192.13	608,291,220	96.75
	Not mapped reads	104,636	3.20	195.51	20,457,852	3.25
	Total reads	3,270,603	100.00	192.24	628,749,072	100.00
Mu_S77	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	2,714,298	95.58	194.43	527,727,799	95.61
	Not mapped reads	125,413	4.42	193.28	24,239,624	4.39
	Total reads	2,839,711	100.00	194.37	551,967,423	100.00
Mu_F74	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	2,010,482	95.25	200.00	402,104,611	95.30
	Not mapped reads	100,262	4.75	197.99	19,850,745	4.70
	Total reads	2,110,744	100.00	199.91	421,955,356	100.00
Mu_M3	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	3,995,559	59.79	151.00	603,329,409	59.79
	Not mapped reads	2,687,643	40.21	151.00	405,834,093	40.21

	Reads in pairs	3,767,676	56.38	325.03	568,919,076	56.38
	Broken paired reads	227,883	3.41	151.00	34,410,333	3.41
	Total reads	6,683,202	100.00	151.00	1,009,163,502	100.00
Mu_F85	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	3,354,736	95.56	209.19	701,773,226	95.59
	Not mapped reads	155,844	4.44	207.97	32,411,644	4.41
	Total reads	3,510,580	100.00	209.13	734,184,870	100.00
Mu_F13	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	3,000,555	95.74	196.47	589,507,553	95.74
	Not mapped reads	133,579	4.26	196.53	26,252,911	4.26
	Total reads	3,134,134	100.00	196.47	615,760,464	100.00
Mu_S38	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	3,162,095	94.52	212.75	672,736,392	94.50
	Not mapped reads	183,371	5.48	213.64	39,175,853	5.50
	Total reads	3,345,466	100.00	212.80	711,912,245	100.00
Mu_S43	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	3,891,101	95.96	194.75	757,796,440	95.94
	Not mapped reads	163,850	4.04	195.64	32,056,027	4.06
	Total reads	4,054,951	100.00	194.79	789,852,467	100.00
Mu_F75	References	1	-	5,631,604.00	5,631,604	-
	Mapped reads	2,696,054	96.10	162.91	439,217,416	95.95
	Not mapped reads	109,365	3.90	169.38	18,524,681	4.05
	Total reads	2,805,419	100.00	163.16	457,742,097	100.00