NOTE TO USERS

This reproduction is the best copy available.

UMI®

DISSERTATION

MOLECULAR EPIDEMIOLOGY OF LEPROSY IN CEBU, PHILLIPPINES AND ESSENTIALITY AND CHARACTERIZATION OF AN ARABINOSYLTRANSFERASE INVOLVED IN THE CELL WALL SYNTHESIS OF MYCOBACTERIA

Submitted by

Rama Murthy Sakamuri

Department of Microbiology, Immunology and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall, 2009

UMI Number: 3401026

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI 3401026
Copyright 2010 by ProQuest LLC.
All rights reserved. This edition of the work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

COLORADO STATE UNIVERSITY

NOVEMBER 05, 2009

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY RAMA MURTHY SAKAMURI ENTITLED MOLECULAR EPIDEMIOLOGY OF LEPROSY IN CEBU, PHILLIPPINES AND ESSENTIALITY AND CHARACTERIZATION OF AN ARABINOSYLTRANSFERASE INVOLVED IN THE CELL WALL SYNTHESIS OF MYCOBACTERIA BE ACCEPTED AS FULFILLING IN PART THE REQURIEMENTS FOR THE DEGREE OF DOCTOR OF PHILOSPOHY.

Committee for Graduate Work

Norman P. Cuthoys

Herbert P. Schweizer

Delphi Chatterjee

Advisor: Varalakshmi D.Vissa

Department Head: Edward A. Hoover

ABSTRACT OF DISSERTATION

MOLECULAR EPIDEMIOLOGY OF LEPROSY IN CEBU, PHILLIPPINES AND ESSENTIALITY AND CHARACTERIZATION OF AN ARABINOSYLTRANSFERASE INVOLVED IN THE CELL WALL SYNTHESIS OF MYCOBACTERIA

Leprosy is a chronic granulomatous infectious disease caused by *Mycobacterium leprae,* which affects the skin, peripheral nerves and mucous membranes. Approximately a quarter million new cases are detected annually throughout the world. Since *M. leprae* is an obligate intracellular bacterium, not cultivable under *in vitro* conditions, and with a long incubation time, the causes for continued prevalence in endemic countries are still unknown. In this regard, molecular tools for differentiation of isolates of *M. leprae* are needed to track and control transmission of leprosy. Genome sequencing of *M. leprae* facilitated the identification of genetic markers such as variable number of tandem repeat (VNTR) and single nucleotide polymorphic (SNP) loci. We standardized methods for DNA extraction from the clinical samples and also developed rapid, high throughput and inexpensive methods like multiplex-PCR and fragment length analysis (FLA) for VNTR analysis; polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) for SNPs to facilitate molecular strain typing of clinical samples. We have applied these methods on the DNA extracts of stored and new samples obtained from leprosy patients from Leonard Wood memorial (LWM), Cebu, Philippines. We compiled a database containing epidemiological information from over 200 leprosy patients.

VNTR alleles were found to be highly concordant when obtained from biopsies and slit skin smears (SSS) from the same patients indicating that either can be used for strain typing. Based on parsimony cluster analysis of 207 clinical isolates, five major groups were observed. VNTR markers alone were able to differentiate the SNP type 3 isolates from the majority of the SNP type 1 isolates. VNTR profiles of the *M. leprae* isolates within the multicase families (MCFs) were similar indicating common source of transmission.

In a continuation study, VNTR data was obtained for another seventy samples from newly diagnosed leprosy patients from LWM, Cebu, Philippines. Multiplex PCR and FLA methods were found to be more efficient and accurate compared to short tandem repeat (STR) amplification and DNA sequencing. Intra-patient VNTR variability using different SSS was found to be minimal. Population structure of *M. leprae* isolates from Cebu remained similar with stable clusters in the phylogentic tree, even with the inclusion of 70 additional samples.

Drug resistance is a cause for concern in the treatment and control of any infectious disease. Drug resistance in leprosy was known to be caused during the dapsone monotherapy control program and patients' irregular treatment. WHO MDT which includes anti-leprosy drugs such as dapsone, rifampicin and clofazimine was introduced in 1982. However drug resistance was reported even in the areas where the MDT was implemented successfully. So to evaluate the primary drug resistance in newly diagnosed leprosy patients who are being treated with the standardized MDT regimen, we developed a multiplex PCR for amplifying the drug resistance determining regions (DRDR) in genes *rpoB, folP1*, *gyrA* and *gyrB*, which are targets of the drugs rifampicin, dapsone and

oflaxacin, respectively. No mutations were detected in *rpoB* and *folP1* DRDRs for 77 samples and in the *gyrA* DRDR for 15 samples tested thus far.

A separate chapter describes the studies on the cell wall biosynthesis. The cell envelope of mycobacteria are comprised of a core macromolecule known as mycolyl arabinogalactan (mAG) which is covalently attached to the peptidoglycan layer forming the mAGP complex and other macromolecules such as lipomannan and lipoarabinomannan. The mAGP complex is essential for survival of the bacteria and is a permeability barrier. The synthesis and assembly of this mAGP complex is still not fully understood. In this regard, through comparative genome approaches we identified *Rv3805c* as a putative glycosyltransferase located in the cell wall synthesis gene cluster. Rv3805c predicted to have ten transmembrane domains, contains a conserved glycosyltransferase and proline rich motif also present in Emb proteins which are involved in cell wall biosynthesis. In this study we identified that, *Rv3805c* encodes an arabinosyltransferase through an *in vitro* arabinosyltransferase assay. Furthermore, a knock out mutant of an ortholog gene *NCgl2780* was generated in *C. glutamicum*. The structural analysis of AG of *NCgl2780* mutant in *C. glutamicum* revealed the loss of β(1→2) Ara*f* residues at the non-reducing ends. We also indentified that the ortholog *MSMEG_6400* is essential for *M. smegmatis*, as we were unable to obtain a genetic knock out mutant in the absence of an additional functional copy of the gene.

Rama Murthy Sakamuri
Department of Microbiology,
Immunology and Pathology
Colorado State University
Fort Collins, Colorado 80523
Fall 2009

ACKNOWLEDGEMENTS

First and foremost, I would like to thank the Almighty for all his blessings!

I am heartily thankful to my advisor Dr. Varalakshmi Vissa, whose guidance and support from the initial to the final level enabled me to develop an understanding of the subject and make progress in my academic career. I would also like to gratefully acknowledge Dr. Patrick Brennan for giving me an opportunity to initially work in leprosy contract. I am truly thankful to my committee members – Drs. Delphi Chatterjee, Herbert Schweizer, and Norman Curthoys who always gave me valuable input over the years.

It is my pleasure to thank Miyako Kimura, Nathan Groathouse, Dr. Wei Li, and Jordon Harrison who initiated and developed some methods for the molecular epidemiology project. Special acknowledgements to past and present members of Macromolecular Resources Facility, CSU- David Gingrich, Jason Rivest, Sue Krueger-Koplin for sequencing and Fragment Length Analysis (FLA). I would like to thank members of Chatterjee lab- Shampa Khasnobis, Jian Zhang, Anita Amin, Shiva Angala and Libin Shi for their help in arabinosyltransferase assays. I would like to thank Drs. Mary Jackson, Dean Crick, Mike McNeil and their lab members for their support and valuable suggestions.

I would also like to take this opportunity to acknowledge our international collaborators in the project-

Dr. Xiaoman Weng – Beijing Tropical Medicine Research Institute, China

Dr. Indira Nath- Blue Peter Research Centre, LEPRA Society, India

Dr. Rupendra Jadhav – Stanley Browne Laboratories, India

Dr. Nora Cardona- Castro – Instituto Colombiano de Medicina Tropical, Colombia

Dr. Ray Cho – Yonsei University, Republic of Korea

Dr. Benjawan Phetsuksiri – National Institute of Health, Ministry of Public Health, Thailand

Dr. Philip Noel Suffys- Oswaldo Cruz, Rio de Janeiro, Brazil

It was truly wonderful to have company of Vissa lab members- Dr. Wei Li and Reem Al-Mubarak; Leprosy lab members- Dr. John Spencer, Hee Jin Kim, Becky Rivoire, Angela and Renato. I would like to thank our coffee club members- Shiva, Drs. Sai Prasad,

Suresh and Venugopal Pujari for their interesting discussions on Andhra politics to movies. Thanks to Drs. Hyung Jin, Devinder and Rakesh for their friendship and witty humor. Special thanks to Dr. Vissa's family for inviting us to their family celebrations. I would like to express my heartfelt gratitude to my dear friends- Drs. Janaiah, Hema Prasad, Prithviraj, Kiran, and Kalyani for their constant encouragement, support and good will. I am indebted to Lorene Martinez for her support and her time to proof read my thesis draft.

It is an honor to pay tributes to all the teachers and friends from my school, Bapatla Agricultural College (Bapatla), Mahatma Phule Krishi Vidhya Peeth (Maharashtra) and, Indian Agricultural Research Institute (New Delhi). I owe special acknowledgements to my previous mentor, Dr. P. Ananda Kumar for giving me an opportunity to get hands on experience with basic molecular biology techniques.

This thesis would not have been possible without the funding from NIH RO1 AI063457, HEISER Program for Research in Leprosy and Tuberculosis of the New York Community Trust and NIH/ NIAID leprosy contract NO1-AI-25469.

Special thanks to aunts, uncles, cousins and beloved grandparents. I am also grateful to my in-laws for their love and support.

I owe my deepest gratitude to my beloved parents and brother who have always cherished their dreams in my academic success. Words fail me to express my appreciation to my wife Lakshmi whose dedication, love and persistent confidence in me, kept me motivated for graduate school.

Finally,-I would like to thank everybody who was important in realization of this thesis, as well as my apology that I accidentally have forgotten any one of you.

DEDICATION

I dedicate this dissertation to my parents Late Mr. Veerappa Naidu and Mrs. Parvathi, my wife Lakshmi, and my brother Srinivasa Rao for their continuing support and love. They have given me such tremendous inspiration and confidence to complete this degree.

Rama Murthy Sakamuri

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

Ac	acylated (Ac_n where n is number of acyl chains)
ADML	armadillo-derived *M. leprae*
AG	arabinogalactan
Amp	ampicillin
Ara*f*	arabinofuranose
AraLAM	uncapped lipoarabinomannan
BB	borderline borderline
BI	bacteriological index
BL	borderline lepromatous
BT	borderline tuberculoid
CMI	cell mediated immunity
CSC	Cebu skin clinic
DAP	diaminopimelic acid
DCO	double crossover
DDS	4,4'-diamino-diphenyl sulfone, dapsone
DNA	deoxyribonucleic acid
DPA	decaprenyl-monophosphoryl-D-arabinose (C_{50}-P-Ara*f*)
DPM	decaprenyl-monophosphoryl-D-mannose (C_{50}-P-Man*p*)
DRDR	drug resistance determining region
EMB	ethambutol
FLA	fragment length analysis
Gal*f*	D-galactofuranose
GDP	guanidine-5'-diphosphate
Gen	gentamicin
GlcNAc	*N*-acetyl-glucosamine (2-acetamido-2-deoxy-a-D-glucopyranose)
GTF	glycosyltransferase
HLA	human leukocyte antigen
Hyg	hygromycin
IFN-γ	Interferon-γ
IL	interleukin
Kan	kanamycin
LAM	lipoarabinomannan
LB	Luria-Bertani
LL	lepromatous leprosy
LM	lipomannan
LWM	Leonard Wood Memorial
mAGP	mycolyl-arabinogalactan-peptidoglycan
ManLAM	mannose-capped lipoarabinomannan

Man*p*	D-mannopyranose
MB	multibacillary
MCF	multicase family
MDT	multidrug therapy
MFP	mouse foot-pad
MHC	major histocampatibility complex
MLVA	multi locus variable number of tandem repeat analysis
MP	maximum parsimony
M-PCR	multiplex PCR
MurNAc	*N*-acetylmuramic acid
MurNGlyc	*N*-glycolylmuramic acid
NCDR	new case detection rate
NGO	non-governmental organization
NJ	neighbor joining
OD	optical density
PB	paucibacillary
PCR	polymerase chain reaction
PG	peptidoglycan
PI	phosphatidyl-*myo*-inositol
PIM	phosphatidyl-*myo*-inositol mannoside
PPM	polyprenyl phosphomannose
PRPP	5'-phosphoribosyl-1-phosphate
RFLP	restriction fragment length polymorphism
Rha	rhamnose
SCO	single crossover
SNP	single-nucleotide polymorphism
STR	short tandem repeat
Strep	streptomycin
TB	tuberculosis
Th1	T helper cell type 1
Th2	T helper cell type 2
TLC	thin layer chromatography
TLR	Toll-like receptor
TN	sequenced *M. leprae* Tamil Nadu strain
TNF-α	tumor necrosis factor-α
TT	tuberculoid leprosy
UDP	uridine-5'-diphosphate
UPGMA	unweighted pair group method with arithmetic mean
VNTR	variable number of tandem repeat
WHO	World Health Organization

Chapter 1

Molecular epidemiology of leprosy in Cebu, Philippines

Chapter 1.1

Literature review of leprosy

1.1.1 History of leprosy

Leprosy is the oldest disease known to humans. The biblical word 'Lepra' is derived from two groups root words with opposite meaning, the first one means 'rough or harsh like a scaly surface' and the second 'thin, delicate, or fine like a scale' probably represent the symptoms of the disease (37). It was also noted that the word 'tsara'ath which means 'blemish' appears twenty eight times in the book of Leviticus and in Hebrew's ritualistic law. Leprosy was believed to be an 'unclean' condition and person with these symptoms were typically quarantined for seven to fourteen days before judging the condition (37). In interpreting the ancient literature, which was left behind by ancient physicians in the form of medical texts describing various illness and treatments, it was found that the word 'Lepra' was found in 'Ebers Papyrus', the ancient Egyptian medical document from 1600 B.C. However, these writings were often misinterpreted with other skin diseases like fungal, scoriasis, dermatitis, gas gangrene and others, as the translators of these ancient writings were criticized for lack of proper knowledge about the symptoms of various diseases (37). Other ancient medical texts written by famous physicians such as Imhotep and others during the Egyptian Old kingdom (~2000 B.C.), describe no sign of leprosy or its symptoms (33). If leprosy was a major disease during that period it would not have been ignored by these ancient physicians since all the common ailments were described in great detail. Furthermore, no evidence of leprosy was found in ancient Egypt from hundreds of mummies and skeletal material that has been identified from the ancient Egyptian period (42, 43). So it was hypothesized that the word 'lepra', was often used for different ailments not necessary leprosy (37). It was also

found that the leprosy disease and its symptoms were also not mentioned in Hippocrates Corpus, which was believed to be written between 430-330 B.C. (32).

The historical text by Thucydides on the peloponnesia war of around 400 B.C. which was translated by Jones and Powell described a contagious disease which began in Ethiopia, spread to Egypt and Libya (74). It was found that Aretaeus of first century B.C. gave a detailed description of leprosy, and from then onwards; leprosy was typically mentioned in the medical texts (26). It is assumed that leprosy was introduced in the Mediterranean region and European countries sometime between 467 B.C. and 150 A.D. (6). From paleontological studies of leprosy in Europe, it was found that the disease might have existed between 400 B.C. (63, 66) and in Britain and France the disease was believed to have reached its peak during $1^{st} - 4^{th}$ century A.D.

According to the ancient writings from Asia, the earliest reference to leprosy was found to be written in the Chinese medical classic entitled Nei Ching in the Chou Dynasty in 6th century B.C. (35). The ancient treatment with chaulmoogra oil along with a description of the symptoms of the disease was given in *Sushruta samhita* which is a compilation of ancient Indian medical writings from 600 B.C. and was also found in *Arthasastra* by Kautilya from 320 B.C. from India (17). In agreement with these historical writings, Robbins et. al., recently identified the oldest skeletal evidence of leprosy in India from 2000 B.C. and raised questioning the theory of an African origin (65). Anderson proposed that, it was the army of the Alexander the great who might have spread leprosy to the Mediterranean region after their return from Indian Campaign (~320 B.C.) (6). His contentions were based on the evidence found in India and lack of the evidence of leprosy before Alexander's times in Egypt, Mesopotamia and Persia. Also

the earliest archeological evidence for leprosy was from 2nd century B.C. in Egypt (42, 43). However Mark (47) disagrees with this hypothesis due to the fact that a Persian King named Darius undertook a similar expedition to India in 500 B.C. before Alexander. Mark suggested that if the army was able to acquire and spread the disease so easily, then leprosy should have spread throughout the Mediterranean region far earlier than suggested by the evidence. It was found from the ancient sources that leprosy first appeared in Egypt then spread throughout the Mediterranean region and then slowly to Italy and then Greece. Susceptibility factors within a population influence immunity to this disease (3, 5, 49). However, Mark (47) proposes an alternative way of disease transmission. During 3rd and 4th century B.C., trade between India and Egypt was common where Indian slaves, women, spices and animals were imported into Egypt. He hypothesized that this slave trade might be the cause of the spread of the disease. However, further research needs to be done to determine the geographical origin of leprosy combining both archeological and molecular approaches. During the twelfth and the thirteenth centuries leprosy was assumed to be terrible disease and believed to have affected a quarter of the North European population (18). It was found to have hit hard in the filthy and congested towns. Even though physicians were available it was believed that leprosy was an expression of wrath of God and was left to priests, so quarantine was found to be the best solution to combat the contagious nature of the disease. The sick people were outcasted from the towns and cities and strict rules were made on their movement. They were forbidden to mingle with healthy people (18).

It is believed that leprosy was introduced onto the American continent first by Columbus' soldiers and later by the slave trade from West Africa. It is also believed that

the Chinese introduced the disease into the Hawaiian islands, who visited these islands in various ships before 1830 and also from Chinese labor from Hong Kong in Honolulu around 1852 (25). By the end of the 18th century, leprosy had died out in Britain largely due to the restrictions placed on the leprosy patients and their contact with the healthy community and to general improvement in the conditions of society. Leprosy continued to be a problem in Norway in the 19th century. Gerhard Henrik Armauer Hansen a Norwegian, detected rod shaped bodies, either at rest or with slight movement in 1874 while examining the pathologic anatomy of the leprous nodules under a microscope (28). The significance of this was not immediately appreciated, not until Albert Neisser who stained the bacteria from 100 leprosy patients' samples given to him by Hansen and later confirmed the valuable findings of Hansen. Because of this leprosy is also known as Hansen's disease.

1.1.2 Epidemiology of leprosy

Leprosy was initially found to be present in 122 countries with a prevalence of >one case/10, 000 people in 1985 and this prevalence gradually declined over the years and was seen in 24 countries in 2000 (81), and 12 in 2002 (80). According to the recent report on the global leprosy situation by WHO , the registered prevalence was 212, 802 and the number of new cases detected during 2007 was 254,525 (79). It was also reported that the number of newly diagnosed cases declined by 4% (~11,000 cases) from 2006 to 2007. The annual new case detection had declined from >763 000 in 2001 to 254, 525 in 2007, which is attributed to active leprosy control programs. However, Brazil, Nepal and

Timor-Leste are the three countries with a population more than one million that still need to achieve the elimination goal and which account for 17% of the total number of new cases detected and 23% of registered cases in 2007. Even though leprosy in India has reached its elimination goal, India still accounts for the majority of new cases (137, 685) detected in 2007.

Multidrug therapy (MDT) was introduced by WHO in 1982 (78). The forty fourth World Health Assembly in 1991 proposed to eliminate leprosy by the year 2000, while recognizing leprosy as a "public health problem" (77). Since then, the leprosy elimination campaign has been active throughout the world mobilizing both people and resources. Both governmental and non-governmental organization (NGOs) worked together to achieve this decline in case numbers. A multisectoral partnership Global Alliance to Eliminate Leprosy (GAEL) was established in 1991 to supply free medicines for leprosy control. The foundation is comprised of WHO, Nippon foundation, The International Federation of Anti-leprosy association (ILEP) and Novartis foundation. However several activists argue against the elimination campaign concept because in many countries leprosy control activities were eliminated on reaching the goal of elimination (1 case per 10, 000 populations) (39, 41, 76).

1.1.3 Etiology, Symptoms, and Classifications of leprosy

Leprosy is caused by *Mycobacterium leprae*, which is an obligate intracellular, non-spore forming, microaerophilic bacterium. *M. leprae* is found in the cells singly or in clumps (globi) and replicates primarly in macrophages and Schwann cells. Bacilli are

Gram-positive, acid-fast, alcohol-fast, non-motile rods 1-8 μm in length and 0.2-0.5 μm in width. *M. leprae* has a doubling time of approximately two weeks and grows best at 30°C. Taxonomically *M. leprae* belongs to the order Actinomycetales, suborder Corynebacterineae, and family Mycobacteriaceae.

Leprosy is a chronic, infectious disease that affects mainly skin, mucosa, upper respiratory track and peripheral nerves. Characteristic rashes and bumps develop. When the peripheral nerves are infected and enlarged, the skin becomes numb and causes deterioration of the sense of touch and then the ablity to feel pain and temperature. Muscles become weak in the infected areas which results in the deformities. Sometimes the fingers become so weak that they curve inward and form what looks like a claw. Sores can develop on the soles of the feet, making walking painful and damage to the eyes may lead to glaucoma or blindness. The immune system may produce an inflammatory response causing fever and inflammation of the skin, peripheral nerves and sometimes the lymph nodes, joints, testes, kidneys, liver, and eyes. The skin around bumps swells and becomes red and painful, and the bumps may eventually form open sores.

Diagnosis of leprosy is most commonly based on the clinical signs and symptoms. Skin lesions with sensory loss, with or without thickened nerves and the presence of acid-fast bacilli in the slit skin scrapings are the diagnostic features of leprosy. Lepromin test is sometimes used as a diagnostic test in some countries. Mitsuda Lepromin, which is well characterized and given intradermally, is prepared from a whole cell suspension of autoclaved leprosy bacilli (71). The test response is measured as an induration at the site

of injection after four weeks and is a measure of the individual's ability to have a granulomatous response. However, the lepromin test is not reliable in the diagnosis of leprosy, because false positives are frequently seen. New studies are underway to identify the skin test antigen for diagnostic purposes (10).

For the purpose of global leprosy control, a simple diagnostic scheme suitable for field condition has been developed and implemented. According to WHO, leprosy can be classified based on the type and number of skin areas affected. When few skin lesions are observed with no bacteria in any of the lesions it is known as paucibacillary (PB) condition. When 6 or more skin lesions are present and any one site presents a positive smear it is known as multibacillary (MB) condition.

Ridley-Jopling classification of leprosy is a useful one in categorizing the disease based on the immunopathological features of the disease (64). According to this, leprosy can also be classified as tuberculoid, lepromatous, or borderline. People with tuberculoid leprosy (TT) typically have few localized skin or nerve lesions (PB), with a strong cellular immune response and a weak humoral immune response and the disease is milder. It was also found that granulomas with multi-nucleated giant cells, a large number of lymphocytes, and few bacteria ($<10^5$ bacilli/gram of tissue) were observed in the TT lesions. On the other extreme, people with lepromatous leprosy (LL) typically have more skin lesions (MB), with a weak or no cellular immune response and the disease is more severe. It was also found that the LL patients have high anti-PGL antibody titers with few lymphocytes in the lesions and the macrophages contains numerous bacteria ($>10^8$ bacilli/gram of tissue). The other conditions of leprosy in this classification includes: borderline lepromatous (BL) condition with numerous skin

lesions and bacilli, borderline borderline (BB) condition with numerous skin lesions and few bacilli, and a borderline tuberculoid (BT) condition with few to numerous lesions and few bacilli.

1.1.4 Social stigma of leprosy

Leprosy disease is one of the most socially stigmatized disease ever known. According to Goffman's definition "Stigma is an attribute that is deeply discrediting, and the stigmatized individual is one who is not accepted and not accorded the respect and regard of his peers: one who is disqualified from full social acceptance" (24). During the middle ages, leprosy patients were forced to wear special clothing and ring bells to warn others as they walked in the streets. In many countries leprosy is considered to be a shameful disease, where leprosy patients are often forced to live separately from rest of the community. False beliefs about the causation of leprosy like 'leprosy is a judgment of God for wrongdoings of a person either in his current or past life', victims of witchcraft, uncleanness etc., were found to be the some of the contributing factors (60). The pathological conditions like visible deformities, perceived incurability and the chronic nature of the disease was also found to contribute to this stigma.

This stigma towards leprosy patients and their families affected their quality of life due to its impact on mobility, interpersonal relationships, social, physical and psychological activities. Even the cured patients are not treated as normal patients. The stigma to leprosy was found to affect leprosy control because some patients would rather

conceal their illness than suffer from social rejection. In a study in South Africa, it was found that one third of the black patients in the study population have contemplated suicide after their diagnosis of leprosy (70). Recently due to the change of leprosy care through integrated approaches from both governmental and NGO's the stigma towards leprosy is found to be slightly diminishing.

1.1.5 Immunology of leprosy

As described above, the host immune response determines the spectrum of the disease within a patient. The tuberculoid form of leprosy has high cell mediated immunity (CMI) with a large Th1 type response and in contrary the lepromatous form has low CMI with a Th2 immune response. Once *M. leprae* (intracellular pathogen) invades the host either by invasion of the nasal mucosa, or through skin aberration or other means, dendritic cells are the first cells to modulate early innate immune response to the bacteria (16) and are also found to be effective presenters of the antigens to the T cells (44, 51). Accordingly the Langerhans cells are found to be fewer in LL patients compared to that of TT patients or healthy individuals (23). However it was found that the major histocampatibility complex (MHC) class I and II of monocyte derived dendritic cells are downregulated during *M. leprae* infection and phenolic glycolipid 1 (PGL-1, an abundant bacterial antigen) was found to be the cause for this immuno-suppression (29). Toll-like receptors (TLR) were found to play a important role in the immune responses to *M. leprae*, particularly TLR2 was found to recognize *M. leprae* lipoproteins (19 kDa and 33 kDa) (36, 45, 57) which leads to activation of monocytes and dendritic cells. The

expression of TLR 1 and 2 was found to be greater in TT lesions compared to that of LL lesions (36). The activated dendritic cells present antigen and cause the activation of naïve T cells by IL-12 secretion (16). IL-2 is a cytokine which promotes cell mediated immunity (CMI), and is expressed more on Th1 lymphocytes preferentially shifting the immune response further towards a Th1 response. IL-2 also induces TNF-α (cytokine responsible for cellular activation and granuloma formation). The large quantities of TNF-α in tuberculoid lesions leads to induction of apoptosis in *M. leprae*-infected cells thereby decreasing the bacterial load (30) leading to granuloma formation in tuberculoid lesions where the $CD4^+$ cells are present in the center of the granuloma surrounded by $CD8^+$ cells (52). Cytokines Interferon gamma (IFN-γ), IL-2 and lymphotoxin-α are secreted by T cells in the tuberculoid lesions which results in intense phagocytic activity. The bacteria generally replicate very slowly under the pressure of this effective immune response (TT patients). On the other hand, in lepromatous leprosy, an effective cell-mediated immune response is lacking, with no or poor granuloma formation. The cytokines IL-4 (which down regulates TLR2 in monocytes), IL-5 and IL-10 (which suppress the production of IL-12) are largely found in lepromatous lesions (83). This leads to the differentiation of T cells into a Th2 type of immune response where the B cells are activated to make neutralizing antibodies, predominantly of the IgM, IgA, and IgE classes. These antibodies are not able to reach intracellular bacteria and do not constitute an effective immune response. *M. leprae* is able to grow abundantly in macrophages and cause abundant tissue destruction. Infection is widespread throughout the body, resulting in extensive damage to bones, cartilage, and nerves.

1.1.6 Leprosy reactions and its immunology

Leprosy reactions are the acute inflammatory complications due to adverse immune response to bacterial antigens. Leprosy reactions can occur in patients before treatment, during treatment, and even after treatment. Reactions frequently cause severe nerve damage which leads to disability and deformity. There are two types of reactions: Type 1 and Type 2.

Type 1 reactions: These reactions are often seen in BB, BL and BT patients usually after starting the chemotherapy for leprosy. The clinical symptoms, that appear mostly within the existing lesions, include induration, erythema, tenderness, neuritis with pain, and loss of function. These reactions contribute significantly to nerve damage, particularly if not treated early. Because the immune response is increased, these are termed reversal reactions. Type 1 reactions occur gradually and last many weeks. Patients with Type 1 reactions are treated with corticosteroids.

Type 1 reactions are found to be mainly due to increase in cellular immunity and delayed hypersensitivity (DTH) to *M. leprae* antigens. The *M. leprae* antigens were found to localize in Schwann cells and macrophages (40). Increase of proinflammatory cytokines like IL-12, IFN-γ and TNF-α were observed in the skin (38, 55, 84) and nerves (7) during Type 1 reaction compared with non-reaction controls. From these studies it was found that when the bacteria is killed by anti-leprosy drugs, large amounts of *M. leprae* antigens are freely available which induces a increase in Th1 immune response.

Type 2 reactions which are also called erythema nodosum leprosum (ENL) mainly occurs in BL and LL patients with high bacterial loads, and may also occur spontaneously before therapy. ENL involves erythematous and painful nodules that may sometimes ulcerate and produce fever, neuritis, lymphadenitis, orchitis, arthritis and glomerulonephritis. Type 2 reactions occur spontaneously, sometimes recurrences over several months have been observed and these can last days to weeks. Patients with Type 2 reactions are treated with corticosteroids and/or thalidomide.

Type 2 reactions were found to occur mainly in patients with poor cellular immunity. High levels of IL-12, IFN-γ and TNF-α have been observed in the plasma of some individuals with Type 2 reactions (55, 84). Peripheral blood mononuclear cells (PBMCs) from individuals with ENL secrete increased amounts of TNF-α following stimulation with *M. leprae* whole cells or lipoarabinomannan (9). Increase in these cytokines expression in the skin lesions indicates local cellular immune activation in these lepromatous patients. However the cause and the mechanism of these reactions are poorly understood and need to be studied further.

1.1.7 Host susceptibility and resistance factors in leprosy

The unique features of leprosy are the clinical symptoms of patients. These range from a single to few lesions with no detectable bacilli (PB) to multiple lesions containing large number of bacilli (MB). From the immunological point of view, PB patients have Th1 response whereas MB patients have a Th2 response and this is mainly due to the

immune responses of the host. The spectrum of disease manifestation within the host is not yet clearly known, however many researchers have focused on the hypothesis that the leprosy patients have some immunological defects because in an endemic setting, even though almost all the population is exposed to the bacteria, only few develop the disease and in varying degrees. A two step model for the human immune response to leprosy was first proposed in the 1970s (15) in which in the first level (innate), some genes are responsible for overall susceptibility/resistance to the infection. If innate resistance is insufficient and infection becomes established, then in the second level (acquired), the degree of specific immune responses (cellular/humoral) generated by the infected individual is controlled by other genes (48).

1.1.7.1 Innate response

One of the most extraordinary advances in the understanding of leprosy has been the identification of a locus by Mira and colleagues (50), within the *PARK2/PACRG* locus that is associated with overall susceptibility of human populations to *M. leprae*. This is the first example where a human gene was associated with susceptibility to an infectious disease. In a genome scan of a Vietnamese patient population, the investigators identified a locus within this gene that was highly associated with leprosy. These results were confirmed by a second analysis of Brazilian families with one or more persons affected by leprosy. It was found that two single nucleotide polymorphism (SNPs), PARK2_e01(-2599) and rsl040079 are enough to capture the association between the SNPs and leprosy susceptibility. It was found that haplotypes with the common allele 'T' of PARK2_e01(-2599) and a rare allele 'C' of rsl040079 have an increased risk of leprosy. In addition T-C is dominant over T-T and C-T haplotypes and these are the only

haplotypes present in Vietnamese population. However, the T allele of SNP PARK2_e01 (-2599) showed a significant recessive effect in susceptibility to leprosy in Indian population (46) as opposed tothe dominant effect of haplotype T-C in Brazilian and Vietnamese populations. Haplotype analysis also showed a lack of significant association of any haplotype with leprosy patients or controls. The non-involvement of major risk SNPs in the regulatory region of the PARK2 and PACRG loci with leprosy susceptibility in Indian population highlights the differential effect of these SNPs in regulating genetic susceptibility to leprosy in different populations. Adel et al. (1) reported that the *NRAMP1* gene (2q35), is associated with the overall susceptibility of leprosy in a multicase family study which was confirmed in subsequent studies (4, 22, 62). Interleukin-10 is a cytokine found to be associated with leprosy for both susceptibility to early stages of leprosy infection and disease progression (54). Recently it was identified that there is an association of the presence of a SNP, LTA+80 A allele of the low producing lymphotoxin-α with a high risk of leprosy and found to be more prominent in the younger patients (2). It was observed that SNP, LTA+293 has a significant association in Vietnamese populations but not in North Indian populations and in the case of SNP, LTA+294 it was the opposite.

1.1.7.2 Acquired response

Several studies indicate that there is an association of human leukocyte antigen (HLA) genes, HLA-DR2 and -DR3 with PB leprosy. Early molecular genetic studies reported that the HLA region does play a role in determining the response to *M. leprae* (58). The *TAP2* gene (transporter associated protein with antigen processing) has also been found to be associated with tuberculoid leprosy (61), but because this gene is located

so close to other HLA genes, interpretation of the results was difficult and its association is unclear. In a study in South India Siddiqui et al. (73) identified a series of microsatellite markers at the 10p13 genetic locus, that has a significant linkage with susceptibility to leprosy. TNF-α was found to be generally associated with resistance to *M. leprae*, since the TNF-α serum levels are elevated in PB patients (resistant form of disease). Several polymorphisms were found in the promoter region of TNF-α which may influence the immune responses (53). In another study Roy et al. (68) found that there is a strong association of a variant allele (-308A) of the promoter region of TNF2 of TNF-α. However, in two Brazil studies it was identified that the association exists in the common allele TNF1 (69, 72). Roy et al (67) also found that the different alleles of the human vitamin D receptor gene (*VDR*) are associated with different forms of leprosy i.e. tuberculoid and lepromatous leprosy.

These studies indicate that the risk factors to leprosy may be variable within different populations, however the bacterial genotypes prevailing in each subcontinent are also different. In a study with two different strains of *Mycobacterium tuberculosis*, it was found that individuals with the C allele of Toll Like Receptor, TLR-2 T597C allele were more likely to have tuberculosis caused by the East-Asian/Beijing genotype than other individuals (13). A significant protective association between the Euro-American lineage of *M. tuberculosis* and pulmonary rather than meningeal tuberculosis was also found, suggesting the spectrum of the disease could be variable depending on the bacterial genotype. So it is important to know if there is any association between the bacterial genotype and the host risk factors in the susceptibility of leprosy which needs to be proven.

1.1.8 Chemotherapy of leprosy

Treatment of leprosy with Chaulmoogra oil is mentioned in ancient writings in Ayurvedic medicine from India and dates back more than 2000 years. This was introduced into the western world by a British surgeon working in the Bengal Medical Service (56). Chaulmoogra oil is obtained from the fruit seeds of *Hydnocarpus kurzii,* a plant which belongs to the Flacourtiaceae family. The oil is believed to stimulate the activity of serum lipases on the bacterial cell wall, thereby facilitating the lysis of the bacteria. After its introduction into the west in the late 19th century, it was extensively used until the 1940s. In the 1930s a number of chemical compounds such as sulphonamides and chemically related sulphones were under development for streptococcal and other bacterial infections. Faget (21) was one who first tried sulphanilamide in a group of patients, but without success, but his later use of Promin (glucosulphone sodium) yielded impressive results. Dapsone (diaminodiphenylsulfone, DDS) a compound synthesized by the German scientists, Fromm and Whittmann in 1908 was originally used for the treatment of streptococcal mastitis in veterinary medicine (82). In 1945, Robert Cochrane was attracted to this parent compound, and tried it on leprosy patients. Later, this compound was widely used in leprosy control. However, bacterial resistance to dapsone did develop and is widespread throughout the world. In 1981 WHO recommended the treatment of leprosy with multi-drug therapy (MDT) and these initial recommendations suggested the use of drugs for two years or until the bacillary index is negative in the skin smears and one year for MB and PB patients, respectively. However these recommendations were changed and now WHO reduced the

duration of drug therapy for both MB and PB patients to one year and six months respectively to reduce the resources allocated to leprosy. Several leprosy control activists argue about this change, because relapse rates may increase (39, 41) with this short-course MDT. Current WHO MDT regimens for MB and PB leprosy patients are shown in TABLE 1.1.1.

1.1.9 Genomic features of *M. leprae*

The genomic analysis of leprosy began in 1982 when Imaeda et al., isolated DNA from armadillo grown *M. leprae*, where they studied the genome size, G + C content and homology with other related bacteria using DNA-DNA re-association studies (34). Later DNA hybridization analysis with *M. leprae* and other cultivable mycobacteria indicated that the *M. leprae* genome is unique (8, 27, 59). Whole genome analysis of *M. leprae* by Cole et al. in 2001 of a clinical isolate from Tamil Nadu, India which was propagated in nine banded armadillo at the National Institute for Medical Research, at Mill Hill, London, revealed that the genome size is 3,268,203 bp, with a G+C content of 57.79%, with 1,604 protein coding genes and 1,116 pseudogenes (14). DNA for this whole gemone sequencing was prepared from bacteria isolated from the liver and used either to construct a cosmid library in Lorist6 or a whole-genome shotgun library in pUC18 (19). Comparative genomic analysis of the *M. leprae* genome with *Mycobacterium tuberculosis* genome revealed that *M. leprae* has undergone an extensive reductive evolution (12, 14, 20). Several excellent reviews on the downsizing of genome of *M. leprae* have been published (11, 12, 20, 31, 75). The reduction of the *M. leprae* genome coding capacity was found to be mainly due to deletion of chromosomal regions and the

	Rifampicin	Dapsone	Clofazimine	Ofloxacin	Minocycline	Duration
Multibacillary (MB) patients						
Adult 50-70kg	600mg/m	100mg/d	50mg/d &300mg/m	-	-	12 m
Child 10-14 years	450mg/m	50mg/d	50mg/d &150mg/m	-	-	
Less than 10 years	300mg/m	25mg/d	50mg twice/d &100mg/m	-	-	
Paucibacillary (PB) patients						
Adult 50-70kg	600mg/m	100mg/d	-	-	-	6 m
Child 10-14 years	450mg/m	50mg/d	-	-	-	
Less than 10 years	300mg/m	25mg/d	-	-	-	
Single skin lesion Paucibacillary patients						
Adult	600mg	-	-	400mg	100mg	One dose

Table 1.1.1 Current WHO recommended MDT regimens for MB, PB and single skin lesion leprosy patients.

accumulation of mutations in genes. This reductive evolution is believed to be because of its nature of survival as an intracellular pathogen. It was found that fewer genes were present in every functional category of genes and the gene duplication found in *M. tuberculosis* was absent in the *M. leprae* genome (20, 75). Mutations exist in genes involved in regulation, detoxification, DNA repair, and transport (75). Genes for the synthesis of most small molecules, macro-molecules and cell wall are reasonably conserved.

1.1.10 References:

1. **Abel, L., F. O. Sanchez, J. Oberti, N. V. Thuc, L. V. Hoa, V. D. Lap, E. Skamene, P. H. Lagrange, and E. Schurr.** 1998. Susceptibility to leprosy is linked to the human NRAMP1 gene. J Infect Dis **177:**133-45.

2. **Alcais, A., A. Alter, G. Antoni, M. Orlova, V. T. Nguyen, M. Singh, P. R. Vanderborght, K. Katoch, M. T. Mira, H. T. Vu, T. H. Ngyuen, N. B. Nguyen, M. Moraes, N. Mehra, E. Schurr, and L. Abel.** 2007. Stepwise replication identifies a low-producing lymphotoxin-alpha allele as a major risk factor for early-onset leprosy. Nat Genet **39:**517-22.

3. **Alcais, A., M. Mira, J. L. Casanova, E. Schurr, and L. Abel.** 2005. Genetic dissection of immunity in leprosy. Curr Opin Immunol **17:**44-8.

4. **Alcais, A., F. O. Sanchez, N. V. Thuc, V. D. Lap, J. Oberti, P. H. Lagrange, E. Schurr, and L. Abel.** 2000. Granulomatous reaction to intradermal injection of lepromin (Mitsuda reaction) is linked to the human NRAMP1 gene in Vietnamese leprosy sibships. J Infect Dis **181:**302-8.

5. **Alter, A., A. Alcais, L. Abel, and E. Schurr.** 2008. Leprosy as a genetic model for susceptibility to common infectious diseases. Hum Genet **123:**227-35.

6. **Anderson, J. G.** 1969. Studies in the Medical Diagnosis of Leprosy in Denmark. Dan. Med. Bull. **16:**10-45.

7. **Andersson, A. K., M. Chaduvula, S. E. Atkinson, S. Khanolkar-Young, S. Jain, L. Suneetha, S. Suneetha, and D. N. Lockwood.** 2005. Effects of

prednisolone treatment on cytokine expression in patients with leprosy type 1 reactions. Infect Immun **73:**3725-33.

8. **Athwal, R. S., S. S. Deo, and T. Imaeda.** 1984. Deoxyribonucleic-Acid Relatedness among Mycobacterium-Leprae, Mycobacterium-Lepraemurium, and Selected Bacteria by Dot Blot and Spectrophotometric Deoxyribonucleic-Acid Hybridization Assays. International Journal of Systematic Bacteriology **34:**371-375.

9. **Barnes, P. F., D. Chatterjee, P. J. Brennan, T. H. Rea, and R. L. Modlin.** 1992. Tumor necrosis factor production in patients with leprosy. Infect Immun **60:**1441-6.

10. **Brennan, P. J.** 2000. Skin test development in leprosy: progress with first-generation skin test antigens, and an approach to the second generation. Lepr Rev **71 Suppl:**S50-4.

11. **Brennan, P. J., and V. D. Vissa.** 2001. Genomic evidence for the retention of the essential mycobacterial cell wall in the otherwise defective Mycobacterium leprae. Lepr Rev **72:**415-28.

12. **Brosch, R., S. V. Gordon, K. Eiglmeier, T. Garnier, and S. T. Cole.** 2000. Comparative genomics of the leprosy and tubercle bacilli. Res Microbiol **151:**135-42.

13. **Caws, M., G. Thwaites, S. Dunstan, T. R. Hawn, N. T. Lan, N. T. Thuong, K. Stepniewska, M. N. Huyen, N. D. Bang, T. H. Loc, S. Gagneux, D. van Soolingen, K. Kremer, M. van der Sande, P. Small, P. T. Anh, N. T. Chinh, H. T. Quy, N. T. Duyen, D. Q. Tho, N. T. Hieu, E. Torok, T. T. Hien, N. H. Dung, N. T. Nhu, P. M. Duy, N. van Vinh Chau, and J. Farrar.** 2008. The influence of host and bacterial genotype on the development of disseminated disease with Mycobacterium tuberculosis. PLoS Pathog **4:**e1000034.

14. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007-11.

15. **de Vries, R. R., R. F. Fat, L. E. Nijenhuis, and J. J. van Rood.** 1976. HLA-linked genetic control of host response to Mycobacterium leprae. Lancet **2:**1328-30.

16. **Demangel, C., and W. J. Britton.** 2000. Interaction of dendritic cells with mycobacteria: where the action starts. Immunol Cell Biol **78:**318-24.

17. **Dharmendra.** 1947. Leprosy in ancient Indian medicine. Int J Lepr **15:**424-30.

18. **Edmond, R.** 2006. Leprosy and Empire: A Medical and Cultural History, vol. Cambridge University Press, Cambridge, UK.

19. **Eiglmeier, K., N. Honore, S. A. Woods, B. Caudron, and S. T. Cole.** 1993. Use of an ordered cosmid library to deduce the genomic organization of Mycobacterium leprae. Mol Microbiol **7:**197-206.

20. **Eiglmeier, K., J. Parkhill, N. Honore, T. Garnier, F. Tekaia, A. Telenti, P. Klatser, K. D. James, N. R. Thomson, P. R. Wheeler, C. Churcher, D. Harris, K. Mungall, B. G. Barrell, and S. T. Cole.** 2001. The decaying genome of Mycobacterium leprae. Lepr Rev **72:**387-98.

21. **Faget G.H., P. R. C., Johansen F.A., Dinan J.F., Prejean B.M., Eccles C.G.** 1943. The promin treatment of leprosy, A progress report. . United States Government Printing Office.

22. **Ferreira, F. R., L. R. Goulart, H. D. Silva, and I. M. Goulart.** 2004. Susceptibility to leprosy may be conditioned by an interaction between the NRAMP1 promoter polymorphisms and the lepromin response. Int J Lepr Other Mycobact Dis **72:**457-67.

23. **Gimenez, M. F., I. Gigli, and F. A. Tausk.** 1989. Differential expression of Langerhans cells in the epidermis of patients with leprosy. Br J Dermatol **121:**19-26.

24. **Goffman, E.** 1963. Stigma: Notes on the managment of spoiled identity., vol. Prantice-Hall, Eglewood Cliffs, NJ.

25. **Gould, T.** 2005. A disease apart : leprosy in the modern world, vol. St. Martin's Press, New York.

26. **Grmek, M. D.** 1989. Diseases in ancient greek world, XII ed, vol. John Hopkins University Press, Baltimore.

27. **Grosskinsky, C. M., W. R. Jacobs, J. E. Clarkcurtiss, and B. R. Bloom.** 1989. Genetic-Relationships among Mycobacterium-Leprae, Mycobacterium-Tuberculosis, and Candidate Leprosy Vaccine Strains Determined by DNA Hybridization - Identification of an Mycobacterium-Leprae-Specific Repetitive Sequence. Infection and Immunity **57:**1535-1541.

28. **Hansen, G. H. A.** 1874. Undersogelser angaende spedalskhedens aasager. Norsk Magazin for Laegervidenskaben **4 (Suppl.):**1-88.

29. **Hashimoto, K., Y. Maeda, H. Kimura, K. Suzuki, A. Masuda, M. Matsuoka, and M. Makino.** 2002. Mycobacterium leprae infection in monocyte-derived dendritic cells and its influence on antigen-presenting function. Infect Immun **70:**5167-76.

30. **Hernandez, M. O., I. Neves, J. S. Sales, D. S. Carvalho, E. N. Sarno, and E. P. Sampaio.** 2003. Induction of apoptosis in monocytes by Mycobacterium leprae in vitro: a possible role for tumour necrosis factor-alpha. Immunology **109:**156-64.

31. **Honore, N.** 2002. [The Mycobacterium leprae genome: from sequence analysis to therapeutic implications]. Med Trop (Mars) **62:**473-9.

32. **Howatson, M. C.** 1989. The Oxford Companion of Classical Literature, vol. Oxford University Press, New York.

33. **Ian Shaw, a. P. N.** 1995. The Dictionary of Ancient Egypt, vol. 1. **Princeton University Press**, Princeton, New Jersey

34. **Imaeda, T., W. F. Kirchheimer, and L. Barksdale.** 1982. DNA isolated from Mycobacterium leprae: genome size, base ratio, and homology with other related bacteria as determined by optical DNA-DNA reassociation. Lepr India **54:**801-8.

35. **Jamison, D. G.** 1969. Modern trends in leprosy. Postgrad Med J **45:**408-14.

36. **Krutzik, S. R., M. T. Ochoa, P. A. Sieling, S. Uematsu, Y. W. Ng, A. Legaspi, P. T. Liu, S. T. Cole, P. J. Godowski, Y. Maeda, E. N. Sarno, M. V. Norgard, P. J. Brennan, S. Akira, T. H. Rea, and R. L. Modlin.** 2003. Activation and regulation of Toll-like receptors 2 and 1 in human leprosy. Nat Med **9:**525-32.

37. **Lendrum, F. C.** 1952. The name leprosy. Am J Trop Med Hyg **1:**999-1008.

38. **Little, D., S. Khanolkar-Young, A. Coulthart, S. Suneetha, and D. N. Lockwood.** 2001. Immunohistochemical analysis of cellular infiltrate and gamma interferon, interleukin-12, and inducible nitric oxide synthase expression in leprosy type 1 (reversal) reactions before and during prednisolone treatment. Infect Immun **69:**3413-7.

39. **Lockwood, D. N.** 2002. Leprosy elimination-a virtual phenomenon or a reality? BMJ **324:**1516-8.

40. **Lockwood, D. N., M. J. Colston, and S. R. Khanolkar-Young.** 2002. The detection of Mycobacterium leprae protein and carbohydrate antigens in skin and nerve from leprosy patients with type 1 (reversal) reactions. Am J Trop Med Hyg **66:**409-15.

41. **Lockwood, D. N., and S. Suneetha.** 2005. Leprosy: too complex a disease for a simple elimination paradigm. Bull World Health Organ **83:**230-5.

42. **Lowenstein, E. J.** 2006. Paleodermatology update. Dermatol Nurs **18:**245-6, 255.

43. **Lowenstein, E. J.** 2004. Paleodermatoses: lessons learned from mummies. J Am Acad Dermatol **50:**919-36.

44. **Maeda, Y., M. Gidoh, N. Ishii, C. Mukai, and M. Makino.** 2003. Assessment of cell mediated immunogenicity of Mycobacterium leprae-derived antigens. Cell Immunol **222:**69-77.

45. **Maeda, Y., M. Makino, D. C. Crick, S. Mahapatra, S. Srisungnam, T. Takii, Y. Kashiwabara, and P. J. Brennan.** 2002. Novel 33-kilodalton lipoprotein from Mycobacterium leprae. Infect Immun **70:**4106-11.

46. **Malhotra, D., K. Darvishi, M. Lohra, H. Kumar, C. Grover, S. Sood, B. S. Reddy, and R. N. Bamezai.** 2006. Association study of major risk single nucleotide polymorphisms in the common regulatory region of PARK2 and PACRG genes with leprosy in an Indian population. Eur J Hum Genet **14:**438-42.

47. **Mark, S.** 2002. Alexander the Great, seafaring, and the spread of leprosy. J Hist Med Allied Sci **57:**285-311.

48. **Marquet, S., and E. Schurr.** 2001. Genetics of susceptibility to infectious diseases: tuberculosis and leprosy as examples. Drug Metab Dispos **29:**479-83.

49. **Mira, M. T.** 2006. Genetic host resistance and susceptibility to leprosy. Microbes Infect **8:**1124-31.

50. **Mira, M. T., A. Alcais, V. T. Nguyen, M. O. Moraes, C. Di Flumeri, H. T. Vu, C. P. Mai, T. H. Nguyen, N. B. Nguyen, X. K. Pham, E. N. Sarno, A. Alter, A. Montpetit, M. E. Moraes, J. R. Moraes, C. Dore, C. J. Gallant, P. Lepage, A. Verner, E. Van De Vosse, T. J. Hudson, L. Abel, and E. Schurr.** 2004. Susceptibility to leprosy is associated with PARK2 and PACRG. Nature **427:**636-40.

51. **Mittal, A., R. S. Mishra, and I. Nath.** 1989. Accessory cell heterogeneity in lepromatous leprosy; dendritic cells and not monocytes support T cell responses. Clin Exp Immunol **76:**233-9.

52. **Modlin, R. L.** 2002. Learning from leprosy: insights into contemporary immunology from an ancient disease. Skin Pharmacol Appl Skin Physiol **15:**1-6.

53. **Moraes, M. O., N. C. Duppre, P. N. Suffys, A. R. Santos, A. S. Almeida, J. A. Nery, E. P. Sampaio, and E. N. Sarno.** 2001. Tumor necrosis factor-alpha promoter polymorphism TNF2 is associated with a stronger delayed-type hypersensitivity reaction in the skin of borderline tuberculoid leprosy patients. Immunogenetics **53:**45-7.

54. **Moraes, M. O., A. G. Pacheco, J. J. Schonkeren, P. R. Vanderborght, J. A. Nery, A. R. Santos, M. E. Moraes, J. R. Moraes, T. H. Ottenhoff, E. P. Sampaio, T. W. Huizinga, and E. N. Sarno.** 2004. Interleukin-10 promoter single-nucleotide polymorphisms as markers for disease susceptibility and disease severity in leprosy. Genes Immun **5:**592-5.

55. **Moraes, M. O., E. N. Sarno, A. S. Almeida, B. C. Saraiva, J. A. Nery, R. C. Martins, and E. P. Sampaio.** 1999. Cytokine mRNA expression in leprosy: a possible role for interferon-gamma and interleukin-12 in reactions (RR and ENL). Scand J Immunol **50:**541-9.

56. **Mouat, F. J.** 1854. Notes on native remedies No 1: the chaulmoogra. *Ind Ann Med Sci.* **1:** 646-652.

57. **Oliveira, R. B., M. T. Ochoa, P. A. Sieling, T. H. Rea, A. Rambukkana, E. N. Sarno, and R. L. Modlin.** 2003. Expression of Toll-like receptor 2 on human Schwann cells: a mechanism of nerve damage in leprosy. Infect Immun **71:**1427-33.

58. **Ottenhoff, T. H., and R. R. de Vries.** 1987. HLA class II immune response and suppression genes in leprosy. Int J Lepr Other Mycobact Dis **55:**521-34.

59. **Portaels, F., C. Asselineau, I. Baess, M. Daffe, G. Dobson, P. Draper, D. Gregory, R. M. Hall, T. Imaeda, P. A. Jenkins, and et al.** 1986. A cooperative taxonomic study of mycobacteria isolated from armadillos infected with Mycobacterium leprae. J Gen Microbiol **132:**2693-707.

60. **Rafferty, J.** 2005. Curing the stigma of leprosy. Lepr Rev **76:**119-26.

61. **Rajalingam, R., D. P. Singal, and N. K. Mehra.** 1997. Transporter associated with antigen-processing (TAP) genes and susceptibility to tuberculoid leprosy and pulmonary tuberculosis. Tissue Antigens **49:**168-72.

62. **Ranque, B., A. Alter, M. Mira, N. V. Thuc, V. H. Thai, N. T. Huong, N. N. Ba, P. X. Khoa, E. Schurr, L. Abel, and A. Alcais.** 2007. Genomewide linkage analysis of the granulomatous mitsuda reaction implicates chromosomal regions 2q35 and 17q21. J Infect Dis **196:**1248-52.

63. **Rawcliffe, C.** 2006. leprosy in medieval england, vol. Boydell Press, Woodbridge.

64. **Ridley, D. S., and W. H. Jopling.** 1966. Classification of leprosy according to immunity. A five-group system. Int J Lepr Other Mycobact Dis **34:**255-73.

65. **Robbins, G., V. M. Tripathy, V. N. Misra, R. K. Mohanty, V. S. Shinde, K. M. Gray, and M. D. Schug.** 2009. Ancient skeletal evidence for leprosy in India (2000 B.C.). PLoS One **4:**e5669.

66. **Roberts, C. A.** 2002. The antiquity of leprosy in Britain: The skeletal evidence,, vol. Bradford University, Archaeopress, Oxford.

67. **Roy, S., A. Frodsham, B. Saha, S. K. Hazra, C. G. Mascie-Taylor, and A. V. Hill.** 1999. Association of vitamin D receptor genotype with leprosy type. J Infect Dis **179:**187-91.

68. **Roy, S., W. McGuire, C. G. Mascie-Taylor, B. Saha, S. K. Hazra, A. V. Hill, and D. Kwiatkowski.** 1997. Tumor necrosis factor promoter polymorphism and susceptibility to lepromatous leprosy. J Infect Dis **176:**530-2.

69. **Santos, A. R., P. N. Suffys, P. R. Vanderborght, M. O. Moraes, L. M. Vieira, P. H. Cabello, A. M. Bakker, H. J. Matos, T. W. Huizinga, T. H. Ottenhoff, E. P. Sampaio, and E. N. Sarno.** 2002. Role of tumor necrosis factor-alpha and interleukin-10 promoter gene polymorphisms in leprosy. J Infect Dis **186:**1687-91.

70. **Scott, J.** 2000. The psychosocial needs of leprosy patients. Lepr Rev **71:**486-91.

71. **Sengupta, U.** 1991. Studies on lepromin and soluble antigens of M.leprae: their classification standardization and use. Indian J Lepr **63:**457-65.

72. **Shaw, M. A., I. J. Donaldson, A. Collins, C. S. Peacock, Z. Lins-Lainson, J. J. Shaw, F. Ramos, F. Silveira, and J. M. Blackwell.** 2001. Association and linkage of leprosy phenotypes with HLA class II and tumour necrosis factor genes. Genes Immun **2:**196-204.

73. **Siddiqui, M. R., S. Meisner, K. Tosh, K. Balakrishnan, S. Ghei, S. E. Fisher, M. Golding, N. P. Shanker Narayan, T. Sitaraman, U. Sengupta, R. Pitchappan, and A. V. Hill.** 2001. A major susceptibility locus for leprosy in India maps to chromosome 10p13. Nat Genet **27:**439-41.

74. **Thucydides.** 1942. Historiae, vol. Oxford University Press, Oxford.

75. **Vissa, V. D., and P. J. Brennan.** 2001. The genome of Mycobacterium leprae: a minimal mycobacterial gene set. Genome Biol **2:**REVIEWS1023.

76. **Visschedijk, J., J. van de Broek, H. Eggens, P. Lever, S. van Beers, and P. Klatser.** 2000. Mycobacterium leprae--millennium resistant! Leprosy control on the threshold of a new era. Trop Med Int Health **5:**388-99.

77. **World Health Assembly, W. H. O.** 1991. Elimination of leprosy: resolution of the 44th World Health Assembly. Resolution WHA 44.9. World Health Organization, , Geneva, Switzerland.

78. **World Health Organization.** 1982. Chemotherapy of leprosy for control programme. . Geneva: WHO. Technical Report Series. **768.**

79. **World Health Organization.** 2008. Global leprosy situation, beginning of 2008. . WHO Wkly. Epidemiol. Rec. **83:**293-300.

80. **World Health Organization.** 2002. Global situation on Leprosy., p. 1-8. Wkly. Epidemiol. Rec., vol. 77. World Health Organization, Geneva.

81. **World Health Organization.** 2000. Global situation on Leprosy., p. 225-232. Wkly. Epidemiol. Rec., vol. 75. World Health Organization, Geneva.

82. **Wozel, G.** 1989. The story of sulfones in tropical medicine and dermatology. Int J Dermatol **28:**17-21.

83. **Yamamura, M., K. Uyemura, R. J. Deans, K. Weinberg, T. H. Rea, B. R. Bloom, and R. L. Modlin.** 1991. Defining protective responses to pathogens: cytokine profiles in leprosy lesions. Science **254:**277-9.

84. **Yamamura, M., X. H. Wang, J. D. Ohmen, K. Uyemura, T. H. Rea, B. R. Bloom, and R. L. Modlin.** 1992. Cytokine patterns of immunologically mediated tissue damage. J Immunol **149:**1470-5.

Chapter 1-2

Molecular strain typing of *M. leprae* isolates from Cebu Philippines

1.2.1 INTRODUCTION

The ability to quickly and reliably differentiate among related bacterial isolates is essential for epidemiological surveillance of a disease. Historically, bacterial strain typing started with phenotypic characters mainly by cellular morphology, bacteriophages, serotyping or biochemical typing methods. In the 1970s after the introduction of DNA hybridization techniques, isolates that showed >70% DNA-DNA homology were considered to be the same species (22). Later with the advancements of sequencing technologies, several typing methods with different genomic markers were developed. Since the discovery of *M. leprae* by Hansen, much of the focus was on the culturing of the bacteria (1, 26, 27). After the development of the mouse foot pad model for multiplication of *M. leprae* by Shepard in 1960s (31, 32), there was momentum in leprosy research, until then several attempts were made and failed to grow the bacilli in the culture media. Subsequently, infection had also been achieved in several other rodent species such as white rats (16), Syrian and Chinese hamsters (3) and Mystomys sps (2). During that period specimens from leprosy patients were collected from different countries and were passaged in mouse foot-pads. First observation of strain differences in *M. leprae* isolates was described by Shepard and McRae (33), where the growth rate and the number of bacilli obtained after the harvest were measured in the mouse foot-pad. They observed some of the isolates were 'fast growers' with a growth rate of less than 25 days per generation with yields above 10^6 bacilli/animal, while others 'slow growers,' had a growth rate of more than 30 days per generation with yields below 10^6 bacilli/animal (33).

Mycobacteriophages were used for detecting the strain variation in *M. leprae* isolates (9, 10). Some of the bacteriophages such as BK1, Clark, Sedge, Baits, Watson and D29 were found to be absorbed on *M. leprae* while others like Bg1, Legendre, Marshall, Panetti, Leo and Wiseman were not adsorbed on the leprosy bacilli. However this approach has failed to detect any strain variation in *M. leprae* isolates.

Leprosy in nine banded armadillos (*Dasypus novenmcinctus*) was first reported by Storrs et al. in 1974 (35), however attempts to infect armadillos were started in 1971 (18). Other than humans nine banded armadillos are the only endemic natural hosts of *M. leprae*. Leprosy was found to progress very slowly in armadillos, even in the laboratory the infected animals require 18–24 months of incubation before they were sacrificed to obtain bacilli. Large amounts of bacilli could be obtained from these armadillos. Several reviews on the aspect of leprosy in armadillos were published (15, 23, 29, 36).

The nucleic acid based analyses were started, when Imaeda et al. isolated DNA from armadillo grown *M. leprae*, to study the genome size, G + C content and homology with other mycobacterial sps and other related bacteria by DNA-DNA re-association studies (17). Genomic libraries of *M. leprae* representing more than 99% of its genome were prepared by cloning into a cosmid vector, pHC79, in Escherichia coli K-12 (5). DNA hybridization studies with a DNA fragment from a 65 kDa antigen revealed that *M. leprae* DNA obtained from a single patient or from a pool of patients and cultured in an armadillo, and from a leprosy vaccine strain (*Mycobacterium* strain w) were genetically similar (14). Subsequently the, restriction fragment length polymorphism (RFLP) method was extensively used with several DNA probes on DNA obtained from *M. leprae* isolates from different geographical locations. However, no genetic diversity was detected (6, 39,

41). Then the rRNA genes of *M. leprae* were targeted. These were subcloned from a cosmid library of *M. leprae* into pUC13 and tested as probes on the *M. leprae* strains from different origins, variability was not found (30). Clark-Curtiss et al.(4) identified that a 2.2kb *M. leprae* DNA probe from a genomic library that could hybridize to at least 19 fragments of the chromosomal DNA from *M. leprae* isolates and identical patterns were noticed in four different *M. leprae* isolates. This genomic probe also found to be specific to *M. leprae*, which was later named as RLEP. When the genome sequence was available 28 copies were identified. it was found that no two RLEP sequences are identical and the features of it resemble a bacterial insertion sequence (IS) elements (42).

With the advancements in amplification based techniques, these methods were tested for their potential use in *M. leprae* strain differentiation. Initially these methods were used for diagnostic purposes (28, 40, 43). To test for the genotypic variation, a 282 bp spacer region between 16S and 23S rRNA genes was amplified and subjected to single strand conformation polymorphism (SSCP) analysis from seventy five different sources of *M. leprae*, and the results obtained from this analysis shared no strain variation (11).

Analysis of the genome of the leprosy bacillus has revealed that the bacterium has undergone a reductive evolution characterized by extensive deletions, inactivation of genes, and repetitive DNA (2% of the genome) (7). Cole et al. indentified four families of dispersed repetitive elements which are of five copies or more: RLEP (37 copies), REPLEP (15 copies), LEPREP (8 copies) and LEPRPT (5 copies) in the genome of *M. leprae* (8). Minisatellite composed of variable number tandem repeats (VNTR) which is commonly known as mycobacterial interspersed repetitive units (MIRUs) in *Mycobacterium tuberculosis* were also found in the *M. leprae* genome. Twenty dispersed

loci were found in the *M. leprae* genome. When examined in 14 different *M. leprae* isolates, these dispersed MIRU like loci contained only a single repeat unit, therefore they are not promising loci for strain typing (8). Overall, no variation was seen in the pattern of insertion of these repetitive elements (8), except in one case (12).

VNTR based molecular typing of *M. leprae* isolates was initiated by the discovery of TTC repeats (microsatellite) (34) and a hexamer in the *rpoT* gene (minisatellite) in the genome (19). There are 21 TTC repeats in the sequenced TN *M. leprae* strain. The TTC repeats were found to be absent in the other mycobacterial sps tested, hence were specific to *M. leprae* and varied between 10 to 37 repeats in the *M. leprae* isolates obtained from 34 MB patients (34). Matsuoka et al. found that the six base pair tandem repeat in *rpoT* gene, has four copies in the isolates from East Asia, Korea and the main island of Japan, while three copies are found in the rest of the world (19). Then our group revealed other possible polymorphic genomic regions within *M. leprae* genome (13), where we shortlisted 44 loci (including *rpoT* and TTC loci) by *in silico* analyses of the *M. leprae* genome. We tested 11 short tandem repeat (STR) loci, 9 of which were found to be polymorphic in four *M. leprae* isolates derived from passage through armadillos (13). Matsuoka et al. (20) used the microsatellite (TTC)21 locus to strain type *M. leprae* obtained from nasal swabs and slit skin smears from patients grouped by village, dwelling or household in Indonesia, while Young et al. (45) combined (AT)15, (GTA)9 and (TTC)21 VNTR loci for identification of short and long range *M. leprae* transmission chains in areas within and surrounding the city of Hyderabad, India. Truman et al. (37) examined five VNTR loci [(GAA), (AT)17, (TA)18, (GTA)9 and (C)16(G)8] for the possibility of *M. leprae* strain differentiation, and they found that all loci are highly

discriminatory except one, (C)16(G)8 in which reproducible results were not obtained. They also observed that most of the genotypes remained stable upon passages in nude mice and armadillo. Zhang et al. extended the polymorphic loci to 32 from our list of 44 STR loci and selected 9 loci for determining strain identity in isolates collected from families with multiple cases (46). Matsuoka et al. expanded their study with the minisatellite locus (*rpoT*) to Latin America, where they found *M. leprae* isolates with four tandem repeats copies at that VNTR locus in Mexico and three copies in Peru and Paraguay (21). Our collaborators from China in their study in Qiubei County in Yunnan Province in South West China shared that multi locus VNTR analysis (MLVA) can be used to indentify different clusters of *M. leprae* isolates (38). Two of the clusters were found to be from the isolates of the patients who resided predominantly in the north and northwest parts of Qiubei County. Furthermore, conservation of *M. leprae* genotypes from patients of multicase families (MCFs) was also identified, indicating localized transmission from shared sources.

Monot *et al* identified three SNPs in the *M. leprae* genome after comparing genome sequences of 142 kb of DNA between a small panel of reference strains (25). Although these three SNPs can yield 64 genotype combinations, only four combinations that were termed as SNP type 1, 2, 3 and 4 were found, even when a larger set of global specimens were examined.

These four SNP- types were found to be geographically restricted. It was also suggested that SNP loci are more stable compared to that of VNTR loci. These finding were elaborated more in two separate studies by Monot et al. (24) and Young et al (44) using *M. leprae* isolates from Mali and Malawi respectively. In the Monot *et al* study since

SNP are too conserved and all the 42 *M. leprae* isolates in the patients from Mali, Africa belonged to SNP type 4; five *M. leprae* STR loci were studied (24). The results from these studies demonstrated heterogeneity in prevalent haplotypes indicating that genotype mapping with a small panel of one to five microsatellite VNTR loci was found to be insufficient to discern strain relatedness. The authors of these studies concluded that in these endemic areas, multiple rather than single dominant isolates are found and that additional genomic markers are necessary for strain typing.

1.2.2 RATIONALE AND OBJECTIVES

For these reasons, assays for amplification and differentiation of multiple genomic loci are needed. When these requirements have been met, it would be possible to undertake systematic strain typing studies that include suitable sampling strategies and conventional epidemiology methods for monitoring transmission and detecting clusters of cases. The hypothesis of our study was to prove that VNTRs are stable, reproducible and has the discriminatory ability; to differentiate *M. leprae* isolates from the clinical samples obtained from leprosy patients, so can be used for strain typing purpose.

The objectives of the study are as follows:

- Development of rapid and high throughput methods for *M. leprae* strain typing from clinical samples
- Extending these methods for *M. leprae* strain typing to obtain data for population based studies in the Cebu, Philippines region; to identify transmission patterns by analyzing and comparing the *M. leprae* genotypes with epidemiological criteria.

1.2.3 REFERENCES

1. 1977. Leprosy : cultivation of the etiologic agent, immunology, animal models : proceedings of the Workshop on Future Problems in the Microbiology of M. leprae, 12-15 October, 1976, vol. Pan American Health Organization, Pan American Sanitary Bureau, Regional Office of the World Health Organization, Washington.

2. **Binford, C. H.** 1968. The transmission of M. leprae to animals to find an experimental model. . Int J Lepr **36:**599-600.

3. **Binford, C. H.** 1965. Transmission of Mycobacterium leprae to animals. Nerve involvement in the ears of hamsters. Int J Lepr **33:**865-74.

4. **Clark-Curtiss, J. E., and M. A. Docherty.** 1989. A species-specific repetitive sequence in Mycobacterium leprae DNA. J Infect Dis **159:**7-15.

5. **Clark-Curtiss, J. E., W. R. Jacobs, M. A. Docherty, L. R. Ritchie, and R. Curtiss, 3rd.** 1985. Molecular analysis of DNA and construction of genomic libraries of Mycobacterium leprae. J Bacteriol **161:**1093-102.

6. **Clark-Curtiss, J. E., and G. P. Walsh.** 1989. Conservation of genomic sequences among isolates of Mycobacterium leprae. J Bacteriol **171:**4844-51.

7. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007-11.

8. **Cole, S. T., P. Supply, and N. Honore.** 2001. Repetitive sequences in Mycobacterium leprae and their impact on genome plasticity. Lepr Rev **72:**449-61.

9. **David, H. L., C. Asselineau, M. Daffe, M. A. Laneelle, F. Clement, and V. Levy-Frebault.** 1983. Taxonomy of mycobacterial strains isolated from the tissues of leprosy patients. Ann Microbiol (Paris) **134B:**367-77.

10. **David, H. L., S. Clavel, and F. Clement.** 1982. Adsorption of mycobacteriophages on Mycobacterium leprae: taxonomic significance. Ann Microbiol (Paris) **133:**93-7.

11. **de Wit, M. Y., and P. R. Klatser.** 1994. Mycobacterium leprae isolates from different sources have identical sequences of the spacer region between the 16S and 23S ribosomal RNA genes. Microbiology **140 (Pt 8):**1983-7.

12. **Fsihi, H., and S. T. Cole.** 1995. The Mycobacterium leprae genome: systematic sequence analysis identifies key catabolic enzymes, ATP-dependent transport systems and a novel polA locus associated with genomic variability. Mol Microbiol **16:**909-19.

13. **Groathouse, N. A., B. Rivoire, H. Kim, H. Lee, S. N. Cho, P. J. Brennan, and V. D. Vissa.** 2004. Multiple polymorphic loci for molecular typing of strains of Mycobacterium leprae. J Clin Microbiol **42:**1666-72.

14. **Grosskinsky, C. M., W. R. Jacobs, Jr., J. E. Clark-Curtiss, and B. R. Bloom.** 1989. Genetic relationships among Mycobacterium leprae, Mycobacterium tuberculosis, and candidate leprosy vaccine strains determined by DNA hybridization: identification of an M. leprae-specific repetitive sequence. Infect Immun **57:**1535-41.

15. **Hamilton, H. K., W. R. Levis, F. Martiniuk, A. Cabrera, and J. Wolf.** 2008. The role of the armadillo and sooty mangabey monkey in human leprosy. Int J Dermatol **47:**545-50.

16. **Hilson, G. R.** 1965. Observations on the inoculation of M. leprae in the foot pad of the white rat. Int J Lepr **33:**Suppl:662-6.

17. **Imaeda, T., W. F. Kirchheimer, and L. Barksdale.** 1982. DNA isolated from Mycobacterium leprae: genome size, base ratio, and homology with other related bacteria as determined by optical DNA-DNA reassociation. Lepr India **54:**801-8.

18. **Kirchheimer, W. F., and E. E. Storrs.** 1971. Attempts to establish the armadillo (Dasypus novemcinctus Linn.) as a model for the study of leprosy. I. Report of lepromatoid leprosy in an experimentally infected armadillo. Int J Lepr Other Mycobact Dis **39:**693-702.

19. **Matsuoka, M., S. Maeda, M. Kai, N. Nakata, G. T. Chae, T. P. Gillis, K. Kobayashi, S. Izumi, and Y. Kashiwabara.** 2000. Mycobacterium leprae typing by genomic diversity and global distribution of genotypes. Int J Lepr Other Mycobact Dis **68:**121-8.

20. **Matsuoka, M., L. Zhang, T. Budiawan, K. Saeki, and S. Izumi.** 2004. Genotyping of Mycobacterium leprae on the basis of the polymorphism of TTC repeats for analysis of leprosy transmission. J Clin Microbiol **42:**741-5.

21. **Matsuoka, M., L. Zhang, M. F. Morris, P. Legua, and C. Wiens.** 2005. Polymorphism in the rpoT gene in Mycobacterium leprae isolates obtained from Latin American countries and its possible correlation with the spread of leprosy. FEMS Microbiol Lett **243:**311-5.

22. **Medini, D., D. Serruto, J. Parkhill, D. A. Relman, C. Donati, R. Moxon, S. Falkow, and R. Rappuoli.** 2008. Microbiology in the post-genomic era. Nat Rev Microbiol **6:**419-30.

23. **Meyers, W. M., B. J. Gormus, and G. P. Walsh.** 1992. Nonhuman sources of leprosy. Int J Lepr Other Mycobact Dis **60:**477-80.

24. **Monot, M., N. Honore, C. Baliere, B. Ji, S. Sow, P. J. Brennan, and S. T. Cole.** 2008. Are variable-number tandem repeats appropriate for genotyping Mycobacterium leprae? J Clin Microbiol **46:**2291-7.

25. **Monot, M., N. Honore, T. Garnier, R. Araoz, J. Y. Coppee, C. Lacroix, S. Sow, J. S. Spencer, R. W. Truman, D. L. Williams, R. Gelber, M. Virmond, B. Flageul, S. N. Cho, B. Ji, A. Paniz-Mondolfi, J. Convit, S. Young, P. E. Fine, V. Rasolofo, P. J. Brennan, and S. T. Cole.** 2005. On the origin of leprosy. Science **308:**1040-2.

26. **Nakamura, M.** 2001. [Reasons why Mycobacterium leprae cells do not multiply under the cell-free condition]. Nihon Hansenbyo Gakkai Zasshi **70:**127-33.

27. **Pattyn, S. R.** 1973. The problem of cultivation of Mycobacterium leprae. A review with criteria for evaluating recent experimental work. Bull World Health Organ **49:**403-10.

28. **Rastogi, N., K. S. Goh, and M. Berchel.** 1999. Species-specific identification of Mycobacterium leprae by PCR-restriction fragment length polymorphism analysis of the hsp65 gene. J Clin Microbiol **37:**2016-9.

29. **Rees, R. J.** 1991. Evolution and contribution of animal models in leprosy. Indian J Lepr **63:**446-56.

30. **Sela, S., J. E. Clark-Curtiss, and H. Bercovier.** 1989. Characterization and taxonomic implications of the rRNA genes of Mycobacterium leprae. J Bacteriol **171:**70-3.

31. **Shepard, C. C.** 1960. Acid-fast bacilli in nasal excretions in leprosy, and results of inoculation of mice. Am J Hyg **71:**147-57.

32. **Shepard, C. C.** 1960. The experimental disease that follows the injection of human leprosy bacilli into foot-pads of mice. J Exp Med. **112:**445-454.

33. **Shepard, C. C., and D. H. McRae.** 1971. Hereditary Characteristic that Varies Among Isolates of Mycobacterium leprae. Infect Immun **3:**121-126.

34. **Shin, Y. C., H. Lee, G. P. Walsh, J. D. Kim, and S. N. Cho.** 2000. Variable numbers of TTC repeats in Mycobacterium leprae DNA from leprosy patients and use in strain differentiation. J Clin Microbiol **38:**4535-8.

35. **Storrs, E. E., G. P. Walsh, H. P. Burchfield, and C. H. Binford.** 1974. Leprosy in the armadillo: new model for biomedical research. Science **183:**851-2.

36. **Truman, R.** 2005. Leprosy in wild armadillos. Lepr Rev **76:**198-208.

37. **Truman, R., A. B. Fontes, A. B. De Miranda, P. Suffys, and T. Gillis.** 2004. Genotypic variation and stability of four variable-number tandem repeats and their suitability for discriminating strains of Mycobacterium leprae. J Clin Microbiol **42:**2558-65.

38. **Weng, X., Z. Wang, J. Liu, M. Kimura, W. C. t. Black, P. J. Brennan, H. Li, and V. D. Vissa.** 2007. Identification and distribution of Mycobacterium leprae genotypes in a region of high leprosy prevalence in China: a 3-year molecular epidemiological study. J Clin Microbiol **45:**1728-34.

39. **Williams, D. L., and T. P. Gillis.** 1989. A study of the relatedness of Mycobacterium leprae isolates using restriction fragment length polymorphism analysis. Acta Leprol **7 Suppl 1:**226-30.

40. **Williams, D. L., T. P. Gillis, R. J. Booth, D. Looker, and J. D. Watson.** 1990. The use of a specific DNA probe and polymerase chain reaction for the detection of Mycobacterium leprae. J Infect Dis **162:**193-200.

41. **Williams, D. L., T. P. Gillis, and F. Portaels.** 1990. Geographically distinct isolates of Mycobacterium leprae exhibit no genotypic diversity by restriction fragment-length polymorphism analysis. Mol Microbiol **4:**1653-9.

42. **Woods, S. A., and S. T. Cole.** 1990. A family of dispersed repeats in Mycobacterium leprae. Mol Microbiol **4:**1745-51.

43. **Woods, S. A., and S. T. Cole.** 1989. A rapid method for the detection of potentially viable Mycobacterium leprae in human biopsies: a novel application of PCR. FEMS Microbiol Lett **53:**305-9.

44. **Young, S. K., J. M. Ponnighaus, S. Jain, S. Lucas, S. Suneetha, D. N. Lockwood, D. B. Young, and P. E. Fine.** 2008. Use of Short Tandem Repeat Sequences to Study Mycobacterium leprae in Leprosy Patients in Malawi and India. PLoS Negl Trop Dis **2:**e214.

45. **Young, S. K., G. M. Taylor, S. Jain, L. M. Suneetha, S. Suneetha, D. N. Lockwood, and D. B. Young.** 2004. Microsatellite mapping of Mycobacterium leprae populations in infected humans. J Clin Microbiol **42:**4931-6.

46. **Zhang, L., T. Budiawan, and M. Matsuoka.** 2005. Diversity of potential short tandem repeats in Mycobacterium leprae and application for molecular typing. J Clin Microbiol **43:**5221-9.

Chapter 1.3

RAPID VARIABLE-NUMBER TANDEM REPEAT (VNTR) GENOTYPING FOR *MYCOBACTERIUM LEPRAE* CLINICAL SPECIMENS

This chapter has been published in *Journal of Clinical Microbiology*
Kimura et al, 2009

These studies were initiated by our former laboratory member Miyako Kimura. I contributed towards the standardization of methods, reproducibility of tests, reanalysis and preparation of manuscript.

Kimura M, Sakamuri RM, Groathouse NA, Rivoire BL, Gingrich D, Krueger-Koplin S, Cho SN, Brennan PJ, Vissa V. Rapid variable-number tandem-repeat genotyping for Mycobacterium leprae clinical specimens. J Clin Microbiol. 2009. Jun; 47(6):1757-1766. Epub 2009 Apr 22. PubMed PMID: 19386839

1.3.1 ABSTRACT

Mycobacterium leprae is the non-cultivable pathogen of leprosy. Since the availability of the genome sequence of an isolate of *M. leprae*, multiple locus variable number of tandem repeats (VNTR) analysis (MLVA) has been explored as a tool for strain typing and identification of chains of transmission of leprosy. In order to discover VNTRs and develop methods transferable to clinical samples MLVA was applied to global collection of *M. leprae* isolates derived from leprosy patients propagated in armadillo hosts. PCR amplification, agarose gel electrophoresis, and sequencing methods were applied to DNA extracts from these infected armadillo tissues (n=21). We identified polymorphism in 15 out of 25 short tandem repeat (STR) loci previously selected by *in silico* analyses of the *M. leprae* genome. We then developed multiplex-PCR for amplification of these 15 loci in four separate PCRs suitable for fluorescent fragment length analysis (FLA) and demonstrated STR profiles highly concordant with those from the sequencing methods. Subsequently, we extended this methodology to DNA extracts from human clinical specimens such as skin biopsies (n=30). With these techniques, mapping of multiple loci and differentiation of genotypes has been possible using total DNA extracts from limited amounts of clinical samples at reduced costs and time. These practical methods are therefore available and applicable to answer focused epidemiological questions, including the ability to monitor transmission of *M. leprae* in different leprosy endemic countries.

1.3.2 INTRODUCTION

The causative pathogen of leprosy is *Mycobacterium leprae.* Continued incidence, defying global campaigns to eliminate leprosy even after years of rigorous case finding and availability of multidrug therapy (MDT) regimens (27, 29, 30, 31) are attributed to sub-clinical human and environmental reservoir(s) of the pathogen (2, 8, 13). In recent years, molecular strain typing methodologies have complemented conventional infectious disease epidemiology. With the publication of the complete genome sequence of an isolate from Tamil Nadu (TN), India in 2001 (5), selection of potential polymorphic genomic markers for strain typing was feasible. The first genetic markers that showed polymorphism were short tandem repeats in the *M. leprae* genome. One was a six base pair intragenic sequence in the *rpoT* gene, and the second, a trinucleotide (TTC) repeat element upstream of a pseudogene (17, 23). These sequences exhibit variable number of tandem repeats (VNTR) when sequenced in different isolates. Based on these observations, we shortlisted 44 loci (including *rpoT* and TTC loci) by *in silico* analyses of the *M. leprae* genome and accomplished the screening of 11 Short Tandem Repeat (STR) loci, of which 9 were polymorphic when tested in a small panel of four human isolates derived from passage through armadillos (7). Five were minisatellites (6-50 bp repeat units) and four were microsatellites (1-5 bp repeat units). Since then, others have also showed that VNTR loci exist in *M. leprae* isolates (25, 33, 34). Three single nucleotide polymorphisms (SNPs) have been also discovered, by sequence comparisons of a limited number of strains (20).

The goal of our work has been to discover and apply DNA variation among *M. leprae* isolates to identify sources, and chains of transmission of leprosy in endemic regions. There are however, physiological and practical issues relevant to strain typing of

M. leprae in the clinical setting such as the long incubation period and low transmissibility of leprosy, and the requirement for clinical specimens such as slit skin smears and skin biopsies from leprosy patients due to the inability of *M. leprae* to grow in culture. During the course of the last 4 years, field studies in which STR mapping was implemented have been reported. Matsuoka et al. (16) applied the microsatellite (TTC)21 locus to strain type *M. leprae* obtained from nasal swabs and slit skin smears from patients grouped by village, dwelling or household in Indonesia, while Young et al. (33) combined (AT)15, (GTA)9 and (TTC)21 VNTR loci for identification of short and long range *M. leprae* transmission chains in areas within and surrounding the city of Hyderabad, India. Monot et al, mapped five *M. leprae* STR loci in patients from Mali, Africa (19). The results from these studies demonstrated heterogeneity in prevalent haplotypes indicating that genotype mapping with a small panel of one to five microsatellite VNTR loci was found to be insufficient to discern strain relatedness. However, within an intra-familial case three markers were congruent (33). The authors of these studies concluded that in these endemic areas, multiple rather than single dominant isolates are found and that additional genomic markers are necessary for strain typing.

For these reasons, assays for amplification and differentiation of multiple genomic loci are needed. When these requirements have been met, it becomes possible to undertake systematic strain typing studies that include suitable sampling strategies and conventional epidemiology methods for monitoring transmission and detecting clusters of cases. In light of these laboratory, field and clinical issues, we further explored MLVA techniques. In this paper, we report the development and testing of multiplex-PCR methods for MLVA for reference armadillo-derived *M. leprae* isolates (ADML) and

clinical materials and address allelic properties of individual loci, reproducibility and feasibility of the techniques. In an accompanying paper, we apply and extend these methods and the data for population based studies in Cebu, Philippines (Sakamuri et al, 2008, submitted for publication).

1.3.3 MATERIALS AND METHODS

1.3.3.1 Armadillo derived human *M. leprae* isolates

Clinical isolates of *M. leprae*, passaged in the armadillo were obtained from National Hansen's Disease Program (NHDP), Louisiana State University, Baton Rouge, LA, in the form of infected liver and spleen tissues, as reported in an earlier study (7). Frozen infected armadillo passaged clinical isolates obtained during the 1970-80s from the National Institute of Medical Research, Mill Hill, London (courtesy of the late Drs. Richard Rees and Joseph Colston), and from Florida Institute of Technology, Melbourne, FL (courtesy of Dr. Eleanor Storrs) were selected from the remaining inventory of tissues shipped to Colorado State University (CSU), following the discontinuation of the armadillo colonies at these institutes.

1.3.3.2 Clinical Samples

A set of 30 skin biopsies from leprosy patients were selected from previous approved studies. These specimens were from patients attending Leonard Wood Memorial (LWM), Skin Clinic in Cebu, Philippines collected in collaboration with

Yonsei University (YU). These biopsies were frozen immediately upon collection for storage and sectioning.

1.3.3.3 DNA purification

The Qiagen DNeasy Tissue kit (Qiagen, CA) was employed to obtain total DNA from small portions of infected armadillo tissues(~25 -100 mg). DNA was eluted in 100-200 µl elution buffer, and typically 1-2 µl of DNA was sufficient for one PCR. Genomic DNA from the Philippines sample collection was extracted from cryosections at Yonsei University using a previously described method (32). DNA extracts were provided to CSU. Reference *M. leprae* NHDP63 DNA was obtained from a large scale preparation of bacteria from infected liver or spleen obtained from NHDP as described previously (7).

1.3.3.4 VNTR mapping

The methods for STR amplification using PCR Supermix High Fidelity (Invitrogen, CA), and a touch-down thermocycler program have been described earlier (7). The copy number (allele) for all microsatellite loci was determined by direct sequencing of PCR products. Minisatellite copy number was inferred by 3% agarose gel electrophoresis using the EZ load^{TM} 20bp Molecular Ruler (Bio- Rad) for product length determination. Sequence was obtained for representative products at the Macromolecular Resource Facility at CSU and SeqWright, Houston, USA (Fisher Biotechnology).

1.3.3.5 *M. leprae* primers for multiplex-PCR and Fragment length analysis (FLA)

Fluorescent 5' labeled forward primers were custom synthesized at Applied Biosystems, CA, USA (Table 1.3.1) and the reverse desalted primers were custom

Multiplex PCR Combination Number	Locus[a]	Forward Primer	Tm (°C)	Reverse Primer	Tm (°C)	Amplicon size[b] (bp)
	(AC)8b	VIC-GCCCACTTACCTCAACCAAC	65.4	CCTATAACGGCACTCAGTCCA	61.4	390
	(GTA)9	NED-AGCCTTAGTCGCGCAGATG	62.6	TCCGCTGTCCGTCCGCTGA	68.8	307
	(GGT)5	6FAM-GCAGCGGTGTAACAGCATAGC	68.4	TGTCTGCCTTGCGAAACGGTC	65.5	242
	(AT)17	PET-TCTCCAACATGCTGCGACA	62.1	GTACAGCGGCCTGATCGAA	60.0	181
Combination 1	6-3a[c] (rpoT)	VIC-ATGCCGAACCGGACCTCGACGTTGA	60.0	TCGTCTTCGAGGTCGTCGAGA	59.5	91
	21-3	6FAM-GAATCTGACCTTTCGGAAATG	56.8	CGATGCAGCTTCCTACGG	60.7	312
	(AC)9	NED-AGCGCCCGTTGTCGATAGA	63.9	GACTGGATGTCGGCACCCC	65.5	236
	(AT)15	PET-CAATATGCGGGTTGGCGCTTCTG	66.2	CCGTCTGGCTCGATGGCTGGATTC	68.6	168
Combination 2	(AC)8a[c]	VIC-GTGTTACGCGGAACCAGGCA	65.5	CCATCTGTTGGTACTACTGA	53.5	124
	27-5	6FAM-ATTGAGCAGATGGCCGGTC	62.8	AGCAGTCGGCACGCCCTT	67.8	327
	6-7	VIC-GCCATCGTTGTCGGTTCATC	61.5	CGGAGGAGGTGGGTACGGT	66.1	268
	(TA)18	NED-CGTGCGTCGTGTGTAGGC	63.9	GACGTGGCAACATCGAAGTT	61.0	230
Combination 3	(GAA)21	PET-CTACAGGGGGCACTTAGCTC	62.1	GGACCTAAACCATCCCGTTTT	60.4	201
	18-8	PET-GCCCGTCTATCCGCATCAA	62.5	GCAAAGATCAGCACGCCAAT	61.8	348
	12-5	VIC-CTGGTCCACTTGCGGTACGAC	65.1	GGAGAAGGAGGCCGAATACA	61.4	289
	23-3	6FAM-CCGAAGCCCTGGACGAAG	63.1	GCCGTAAATCCGCTCCC	56.0	243
Combination 4	(TA)10[c]	PET-TAGATTCAAACGACCATGCA	60.0	TGATAATCACGTGTTTCCGC	60.0	185

a tandem repeat locus nomenclature per Groathouse *et al.*(7)
b amplicon sizes per the sequenced *M. leprae* TN strain (5)
c primer sequences per Zhang *et. al.*(34)
VIC, NED, 6FAM and PET refers to the 5' Fluorochrome.

TABLE 1.3.1 Multiplex-PCR primers and combinations for amplification of *M. leprae* short tandem repeat loci (STR)

synthesized at Integrated DNA technologies (IDT), Iowa, USA. These primers were reconstituted in 1 X TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0) at 100 μM concentration. Aliquots of these master stocks were stored at -20°C. Separate working stocks of the forward and reverse primer sets, were prepared according to the combinations listed in Table 1.3.1 and adjusted to 200 μl final volume with TE buffer, such that each primer was at a concentration of 2 μM.

1.3.3.6 Amplification and detection of STRs using multiplex-PCR

Multiplex-PCR was performed using the Multiplex PCR kit (Qiagen). Each reaction (20 μl final volume) was assembled in a PCR cabinet at room temperature and comprised 10 μl of 2X Qiagen multiplex-PCR Master mix, 2 μl Q solution, 2 μl each of the forward and reverse primer working stocks, and 1-2 μl of DNA template; the volume was adjusted with PCR grade water. The final concentration of each primer was thus 0.2 μM. Following an activation step at 95°C for 15 min, 40 cycles of PCR were run as follows: denaturation at 94°C for 30 sec, primer annealing at 60°C for 90 sec, and primer extension at 72°C for 90 sec. The PCR was terminated with a final extension at 72°C for 10 min. For DNA sequencing, the multiplex-PCR sample was simply diluted 10 fold; 1-2 μl was combined with 10 pmoles of the forward primer for the Big Dye cycle sequencing reaction at the Macromolecular Resource Facility at CSU.

1.3.3.7 Fragment length analysis (FLA)

After the multiplex-PCR, 1 μl of the PCR product was diluted 30-60 fold, and 1 μl of the diluted PCR product was combined with 12 μl deionized formamide (Applied Biosystems), and 0.3 μl LIZ-500 DNA standard (Applied Biosystems). The sample was

denatured at 94°C for 5 min, and subjected to capillary electrophoresis on the Applied Biosystems Genetic Analyzer 3130 at Macromolecular Resource Facility at CSU. The samples were injected into the capillary (50 cm length, POP-7 polymer) by applying injection voltage of 1kV for 22sec. The capillary electrophoresis was run at a voltage of 15kV at 60°C for 45min. Following the separation, the electropherograms were visualized and analyzed using GeneMapper Version 3.7 software (Applied Biosystems) to determine the major allele for each VNTR locus, in each multiplex-PCR combination.

1.3.3.8 VNTR data analyses

The Microsatellite Tool Kit (http://oscar.gen.tcd.ie/~sdepark/ms-toolkit/) was used to calculate allele frequency, and average allelic diversity.

1.3.4 RESULTS

1.3.4.1 Discovery of new VNTRs

In our earlier study (7), 44 loci suitable for VNTR strain typing were listed. VNTRs were discovered at 7 microsatellite loci containing one, two and three base pair repeat units on a panel of four *M. leprae* isolates. In this present study, we screened for polymorphisms within more microsatellite loci in the same set of isolates (Table 1.3.2, Taxa 15-18). Reliable sequence results for the monomeric G/C rich tracts such as for the (C)20, (G)22, and (C)18 (G)8 loci were not obtained even though we were able to obtain PCR products as reported by other groups (25, 34). Previously, we did not confirm VNTR polymorphism at the (AC)9, however, sequencing of the products revealed two

	Taxon	Isolate or Animal Code (Passage #)	Other passages (Passage #)	Year of Origin	Country of Origin of donor	Region	VNTR Pattern*															
							(AT)17	(TA)18	(AC)8a	(AT)15	(TA)10	(AC)8b	(AC)9	(TTC)21	(GTA)9	(GGT)5	6-7	12-5	18-8	21-3	23-3	27-5
A	1	TN		1980	India	IN	17	18	8	15	10	8	9	21	9	5	7	5	8	3	3	5
B	2	CD236(0)		1977	India	IN	16	N	8	15	10	8	9	21	9	5	7	5	8	3	3	5
	3	3045/334(2)	336(2)	1987	India	IN	N	11	8	13	8	7	10	15	10	5	5	4	7	1	2	5
	4	3044/329(2)	333(2), 325(2), 264(1), 267(1)	1988	India	IN	N	9	9	17	10	7	8	15	11	4	7	4	7	2	2	5
	5	I480(1)	I478(1)	1985	Philippines	OR	16	19	8	15	10	8	9	21	9	5	7	5	8	3	3	5
	6	A606(2)		N/A	Philippines	OR	11	17	11	12	13	6	7	13	9	5	5	4	8	3	2	5
	7	W88(1)		1976	Philippines	OR	12	18	11	16	11	8	8	28	15	5	7	4	8	3	2	5
	8	W112(2)		N/A	Philippines	OR	11	18	9	16	11	8	8	N	14	5	7	4	8	3	2	5
	9	I226(1)	A308(2)	N/A	Thailand	OR	11	16	10	12	13	6	7	13	9	5	5	4	8	3	2	5
	10	A344(3)		1980	Thailand	OR	11	21	10	14	10	8	7	14	9	4	8	4	8	2	2	5
	11	3039/321(2)	3017(1)	1987	Ethiopia	AFR	13	13	8	20	11	7	10	29	8	5	6	4	8	3	3	5
	12	I487(1)		N/A	Guyana	SAM	12	17	10	12	13	6	7	13	9	5	5	4	8	3	2	5
	13	2936/41(0)		1987	Malawi	AFR	12	16	8	13	N	8	7	13	10	5	8	3	7	2	2	5
	14	3035/299(2)		N/A	Malawi	AFR	12	16	8,9	12,13	8,9	7,8	7,8	13,16	8,10	4,5	6,8	3,4	6,7	2	2	5
C	15	NHDP98	-	1990	USA	NAM	11	12	8	14	11	7	8	10	9	4	7	5	8	2	2	4
	16	NHDP63	-	N/A	USA	NAM	13	16	10	16	13	7	8	10	10	4	7	5	8	2	2	4
	17	BR4923	-	1996	Brazil	SAM	15	18	7	20	8	7	8	12	12	4	6	4	3	2	2	5
	18	Thai-53	-	N/A	Thailand	OR	10	13	11	17	11	7	9	15	9	5	6	5	8	3	2	5
	19	NHDP10		1996	USA	NAM	13	16	10	17	12	7	8	11	10	4	7	5	8	2	2	4
	20	N1192		N/A	N/A	NAM	10	13	11	17	11	7	9	15	9	5	6	5	8	3	2	5
	21	N1200		N/A	N/A	NAM	13	16	10	16	14	7	8	10	10	4	7	5	8	2	2	4
	22	NHDP55		1996	USA	NAM	13	20	10	17	13	7	8	12	10	4	7	5	8	2	2	4

N: Not tested, PCR negative, or sequence negative.

*VNTR Pattern indicates the copy number found at the specified VNTR locus.

N/A: Not available.

The regions are coded as AFR: Africa, IN: India, NAM: North America, OR: Orient, SAM: South America.

TABLE 1.3.2 Origin and Multiple locus VNTR analyses (MLVA) of the armadillo derived *M. leprae* isolates used in this study. Year of origin refers to the year in which the original inoculum was obtained from a patient; **A**: Reference, sequenced strain (5); **B**: Isolates obtained from Mill Hill and FIT. The isolate or Animal Code refers to the armadillo in which leprosy patient-derived or subsequent armadillo-derived *M. leprae* inoculum (passage # 1 or 2 indicates passage from armadillo to armadillo) was inoculated; **C**: Armadillo derived human leprosy specimens obtained from NHDP (25).

alleles, with 8 or 9 copies (7). When we examined the dincucleotide (AC)8b, we found the 7 copy and not the 8 copy allele, as seen in the TN isolate in all of the 4 isolates. Of the (AT) rich loci, (AT)17, (TA)18 (AT)15 were already proven to be polymorphic (7, 34), so we did not screen any more (AT) dinucleotides STR loci at this stage. Four trinucleotide microsatellites, [(AGT)5a, (AGT)5b, (ACT)5, and (GGT)5] were screened; only (GGT)5 was polymorphic, with two alleles, 4 or 5. All four isolates had three copies of the pentanucleotide (CACCG) STR as in the TN strain. From these screens, (AC)9, (AC)8b, and (GGT)5 have therefore been included in the panel of microsatellite VNTR loci.

Likewise we screened the minisatellites 6-3b, 7-3, 10-4, and 15-3 and 23-3 loci for VNTRs within the four standard isolates. Only the 23-3 locus was polymorphic showing the two copy allele instead of the 3 copy allele seen in the TN strain (Table 1.3.2). Similar trends of polymorphisms in STR loci were detected in an independent study that included Asian isolates maintained by the foot pad inoculation system in nude mice (34).

1.3.4.2 VNTR mapping of an archived collection of *M. leprae* isolates from various parts of the world

We expanded our reference panel with a collection of frozen armadillo tissues sent from the National Institute of Medical Research, Mill Hill, London and Florida Institute of Technology, Melbourne, FL. The *M. leprae* inocula were originally obtained from leprosy patients from India, Ethiopia, Malawi, Guyana, Philippines and Thailand; such animal infections were performed during WHO studies to generate sufficient bacilli

for vaccine production (22). *M. leprae* isolates representing the original armadillo infection (human biopsy suspension to armadillo, i.e., passage number zero), and subsequent armadillo to armadillo passages (first, second or third) in these residual tissue banks were tested if available. We utilized these specimens to assess purity and stability of VNTR patterns (Table 1.3.2). Tissue records indicated that in some instances, biopsy homogenates from multiple patients and/or biopsies were combined for the first armadillo inoculation. A panel of 16 VNTR loci [10 microsatellites including (TA)10 and 6 minisatellites] was amplified and analyzed. All loci with the exception of 6-3a were polymorphic with new alleles also being identified within this expanded collection.

With regard to the purity and stability of VNTR profiles of *M. leprae* in these specimens, except in two instances (compare I480 with A606, and 2936/41 with 3035/299 in Table 1.3.2) we found a characteristic VNTR pattern for 16 loci for each tissue, despite the possibility that multiple biopsies were used in the first inoculation. In this VNTR mapping exercise, we were able to find two isolates (CD236 and I480) that closely matched the profile of the original sequenced TN isolate (5). The combined presence of the TN-like 23-3 locus and 21 copies at the TTC triplet, is a signature not found in any of the other isolates. In this collection of 21 ADML samples plus the *in silico* data from TN genome, the number of haplotypes were 16, with a genetic diversity of 0.5463, which is a measure of diversity across the 15 VNTR loci (Table 1.3.2) and was calculated in Microsatellite tool kit that uses definition per Nei's gene diversity (21).

1.3.4.3 Selection of VNTR loci and primer combinations for multiplex-PCR

We then developed a multiplex-PCR approach for strain typing of clinical isolates. We utilized primer sequences described by us or others, in various combinations to enable simultaneous amplification of three or four loci, such that the products were detectable in the 3% agarose gel, with minimum overlap (Fig. 1.3.1). The combinations listed in Table 1.3.1, proved to be satisfactory. At present there are four combinations for amplification of 15 loci. Primer sets for loci 6-3a and (TA)10 have been subsequently added to combinations 1 and 4 respectively allowing mapping at 17 STR loci.

1.3.4.4 Multiplex-PCR and fragment length analysis (FLA) for clinical samples

All the combinations were optimized with NHDP63 DNA obtained from tissue of armadillo infected with the isolate from a leprosy patient, and subsequently utilized as our reference DNA for VNTR mapping (7). In order to test the feasibility of multiplex-PCR and FLA for clinical samples we used DNA from 30 stored biopsy samples from leprosy patients (see Fig. 1.3.1, panels 1a, 2a, 3a and 4a). The VNTR alleles for NHDP63 and the 30 clinical isolates are shown in (Table 1.3.3).

The addition of Q solution, a co-solvent that can alleviate amplification problems associated with secondary structure and/or high GC content, was found to be useful for some loci, and was therefore routinely included in the PCR. In addition to conventional agarose gel electrophoresis, the products of the multiplex-PCRs were subjected to FLA (see Fig.1.3.1, panels 1b, 2b, 3b, 4b).

We also retested several of the armadillo derived *M. leprae* (Table 1.3.2) by this multiplex PCR- FLA method and obtained concordant alleles. Interestingly, the FLA profile from the tissue infected with 3035/299 from Malawi, which was previously

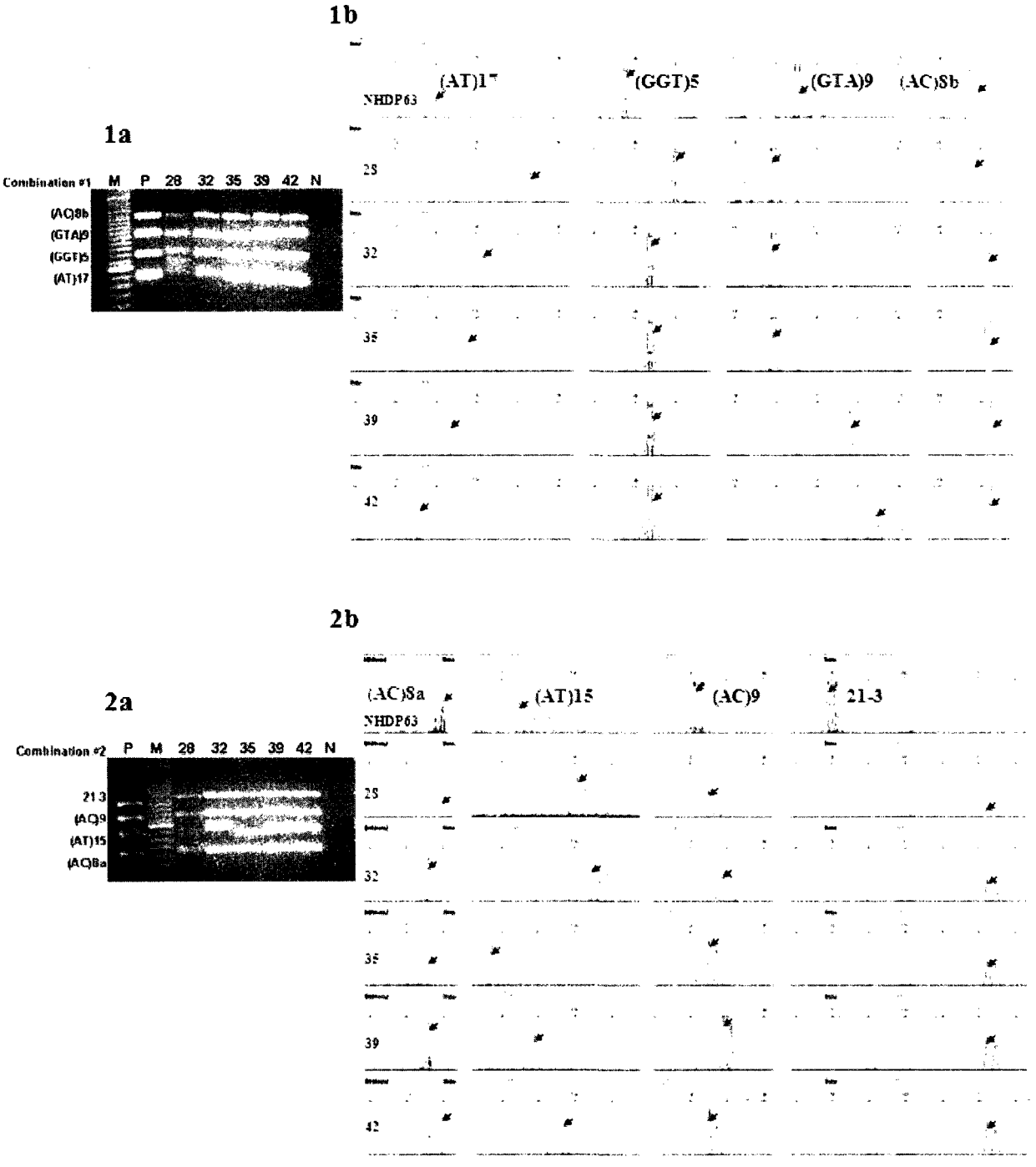
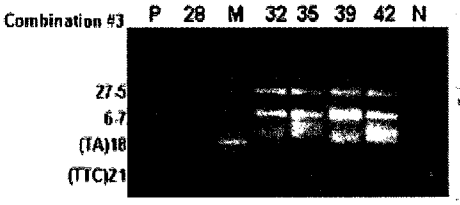


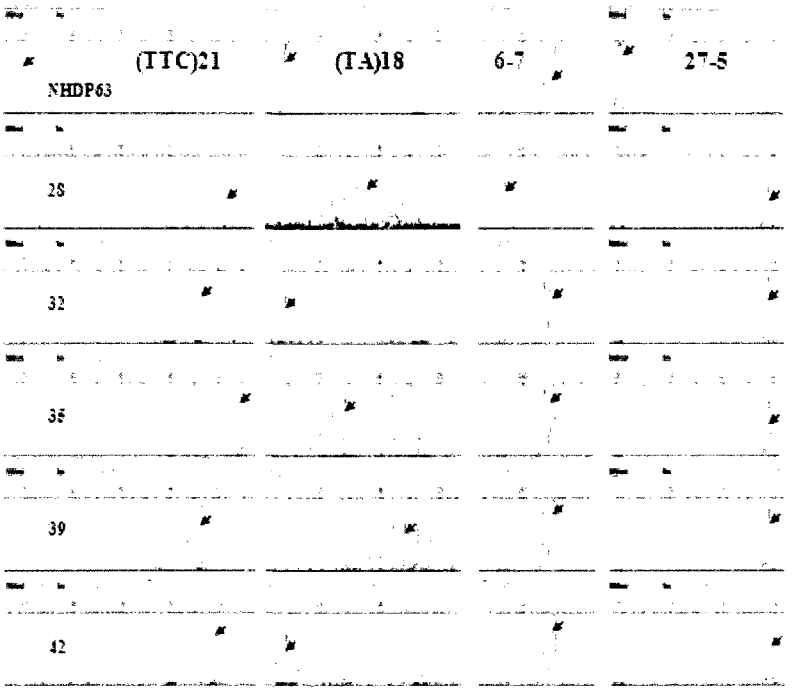
FIG. 1.3.1 Gel electrophoresis and fragment length analysis (FLA) of 15 VNTR loci from 5 Cebu clinical isolates amplified using the combination 1, 2, 3, and 4 primer sets.
Panels 1a, 2a, 3a and 4a: Gel electrophoresis of multiplex PCR products (combinations 1-4) Lanes: M: 20-bp DNA marker (Bio-rad), P: Multiplex PCR products from NHDP63 (positive control) 28, 32, 35, 39, 42: Multiplex PCR products from clinical isolates. N: Negative control; 1b, 2b, 3b and 4b are the electropherograms obtained from FLA, where the orange peaks represent the LIZ-500 size standards and red, blue, black and green peaks represents the VNTR products in each of the combinations; arrows indicate the peak selected as the allele for each locus.

(Fig. 1.3.1 continued..)

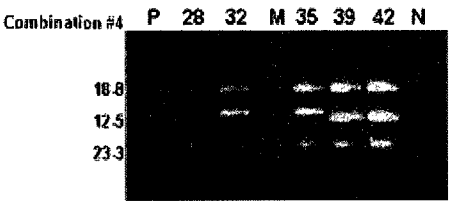
3a


Combination #3
P 28 M 32 35 39 42 N
27-5
6-7
(TA)18
(TTC)21

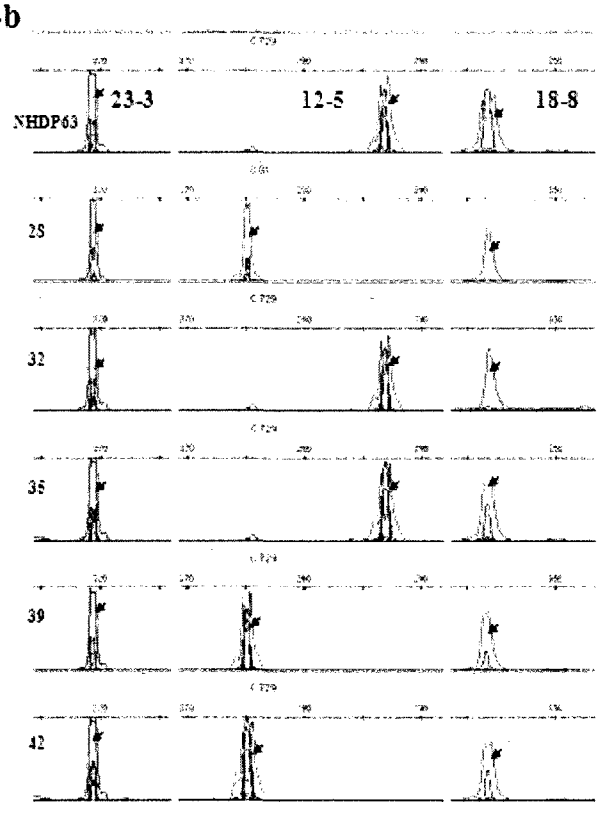
3b


(TTC)21
(TA)18
6-7
27-5
NHDP63
28
32
35
39
42

4a


Combination #4
P 28 32 M 35 39 42 N
18-8
12-5
23-3

4b


23-3
12-5
18-8
NHDP63
28
32
35
39
42

	Combination 1				Combination 2				Combination 3				Combination 4		
	(AC)8b	(GTA)9	(GGT)5	(AT)17	21-3	(AC)9	(AT)15	(AC)8a	27-5	6-7	(TA)18	(TTC)21	18-8	12-5	23-3
NHDP63	7	10	4	13	2	8	16	10	4	7	16	10	8	5	2
13	8	9	5	18	3	9	15	9	5	7	26	31	8	5	2
14	8	11	5	15	3	9	16	9	5	7	18	24	8	4	2
15	8[a]	11[a]	5[a]	12[a]	3	8	17	9	5	7[a]	18[a]	22[a]	8[a]	4[a]	2[a]
16	8	9	5	14	3	10	13	9	5	7	21	19	8	5	2
17	7	10	4	13	1[a]	9[a]	13[a]	9[a]	5[a]	8	15	14	7	4	2
18	8[a]	10[a]	5[a]	14[a]	3	8	16	9	5	7	20	36	8	4[a]	2[a]
19	7	8[a]	4	14	1[a]	9[a]	17[a]	9[a]	5[a]	8[a]	14[a]	16[a]	7[a]	4[a]	2[a]
20	8[a]	9[a]	5[a]	16[a]	3[a]	9[a]	12[a]	8[a]	5[a]	8[a]	18[a]	23[a]	8[a]	4[a]	2[a]
21	6	11	5	12	3[a]	8[a]	19[a]	9[a]	5[a]	7	25	29	8	4[a]	2[a]
22	6[a]	11[a]	5[a]	12[a]	2[a]	9[a]	13[a]	11[a]	5[a]	6[a]	13[a]	20[a]	8[a]	4[a]	2[a]
23	8	9	5	16	3	8	11	9	5	8	17	28	8	4	2
24	8	9	5	17	3	9	14	9	5	7	16	22[a]	8	5	3
25	7	10	4	14	1	8	15	9	5	8	15	20	7	4	2
26	8	8[a]	5	18	3	9	13	10	5	7	18	19[a]	8	5	3
27	8	8	5	16	3	10	15	9	5[a]	7[a]	18[a]	31[a]	8	5	2
28[c]	7	9	6	19	3	9	20	10	5	6	22	25	8	4	2
29	7	9	6	16	3[b]	9[b]	19[b]	10[b]	5[b]	6[b]	20[b]	24[b]	8	4	2
30	7	11	5	13	3[b]	8[b]	15[b]	8[b]	5[b]	7[b]	23[b]	24[b]	8	4	2
31	ND	11	ND	17	3	ND	15	9	4	ND	23	28	ND	5	2
32[c]	8	9	5	16	3	10	21	9	5	7	16	23	8	5	2
33	7[b]	9[b]	4[b]	13[b]	1[b]	9[b]	14[b]	9[b]	5[b]	8[b]	13[b]	16[b]	7[b]	4[b]	2[b]
34	8	9	5	15	3[b]	9[b]	15[b]	9[b]	5[b]	7[b]	21[b]	27[b]	8	5	2
35[c]	8	9	5	15	3	9	14	9	5	7	21	26	8	5	2
36	8	10	5	15	3[b]	9[b]	21[b]	9[b]	5[b]	6[b]	20[b]	26[b]	8	5	2
37	8	9	5	15	3[b]	9[b]	15[b]	10[b]	5[b]	7[b]	18[b]	24[b]	8	5	2
38[c]	8	9	ND	16	3	9	15	9	5	7	17	27	8	5	2
39[c]	8	12	5	14	3	10	17	9	5	7	26	23	8	4	2
40	ND	ND	ND	15	1[b]	9[b]	14[b]	9[b]	4[b]	4[b]	15[b]	25[b]	8	5[b]	2[b]
41	8	10	5	18	3[b]	9[b]	14[b]	9[b]	5[b]	8[b]	21[b]	31[b]	8	5	2
42[c]	8	13	5	12	3	9	19	10	5	7	16	24	8	4	2

[a] Also sequenced; [b] Two FLA runs using same PCR sample; [c] 3 PCRs and 3 FLA runs (2 PCRs with Qiagen Multiplex PCR Kit and 1 with ABI Amplitaq Gold PCR Master Kit); ND Not detected

TABLE 1.3.3 VNTR data from clinical samples from Cebu, Philippines using Multiplex-PCR- FLA.

thought to be different from 2936/41 (zero passage specimen) at multiple loci, was found to carry two alleles at these loci, one the same as that found in 2936/41 along with secondary alleles. The results were reproduced when a different scraping from the same frozen tissue was used for DNA extraction and typing.

1.3.4.5 Stutter phenomena and 3' base tailing and allele determination

Stutter products defined as products formed by slipped stand DNA synthesis at repeat regions were commonly observed for the microsatellite loci. We observed both minus and plus stutters ranging from -3 to +3 repeats with reference and clinical template DNAs (Table 1.3.4). The addition of an extra base at the 3'end of PCR products (non template 3' extension) was also observed. The extent and pattern of stutters and 3'tailing tended to be characteristic for each locus. Despite the stutter and 3' base tailing, we were able to select the peaks, as indicated by the arrows in the electropherograms of Fig. 1.3.1 (panels 1b, 2b, 3b, 4b), for the reportable alleles of the majority of the loci. A non-specific product peak at 242 bp (NED label, black peak) in combination 3 can be seen with human derived specimens.

1.3.4.6 Concordance between VNTR alleles determined by fragment length and DNA sequencing

In order to verify the FLA based allele calls, a subset of samples was submitted to multiplex-PCR using unlabeled primers and DNA sequencing of amplicons. For each combination, we randomly tested at least three DNA isolates. The alleles determined by sequencing of the amplicons were in concordance with those estimated by the FLA (Table 1.3.3).

Multiplex PCR Combination	Locus	Amplicon sizes*	3' tailing	Stutter products					
				-3	-2	-1	+1	+2	+3
#1	(AC)8b	384.15 ± 0.1589	+[b]	-	-	+	+	-	-
	(GTA)9	307.86 ± 0.4416	+[b]	-	-	+	+	-	-
	(GGT)5	238.53 ± 0.3047	+[b]	-	-	-	-	-	-
	(AT)17	174.83 ± 0.0961	-	+	+	+	+	+	+
#2	21-3	289.75 ± 0.3965	+[b]	-	-	+	-	-	-
	(AC)9	231.58 ± 0.2792	+[a]	-	-	+	-	-	-
	(AT)15	173.54 ± 0.2251	+[b]	+	+	+	+	+	+
	(AC)8a	125.17 ± 0.1555	+[b]	-	+	+	+	+	-
#3	27-5	300.61 ± 0.1253	+[b]	-	-	+	-	-	-
	6-7	263.95 ± 0.2714	+[a]	-	-	+	-	-	-
	(TA)18	224.51 ± 0.1618	-	+	+	+	+	+	+
	(GAA)21	169.80 ± 0.1354	-		+	+	+	-	-
#4	18-8	346.20 ± 0.2149	-	+[c]	+[c]	+[c]	-	-	-
	12-5	286.96 ± 0.2742	+[b]	-	-	-	-	-	-
	23-3	219.40 ± 0.2408	+[b]	-	-	-	-	-	-

FLA was performed on the ABI Genetic analyzer 3130.
* refers to the average amplicon sizes ± standard deviation for 10 independent multiplex PCR and FLA experiments.
'+' and '–' signs indicate the presence and absence respectively of 3' tailing and stutter products.
[a] 3'tailed and blunt PCR product are detected in approximately equal proportions.
[b] 3' tailing proceeds to near completion; the blunt PCR product is in limited abundance
[c] -1,-2 and -3 stutter products are weak. Instead, peaks at the 2 copy and 3 copy allele positions can be seen, which would be the -5 and -6 stutter products for NHDP63.

TABLE 1.3.4 Amplicon sizes, 3' tailing effect, and stutter products detected in FLA for NHDP63.

1.3.4.7 Reproducibility of multiplex-PCR and FLA

In order to test for run to run reproducibility of multiplex-PCR and FLA, multiplex-PCR was performed several times using NHDP63 DNA and the products from each PCR were submitted to FLA on separate days. Overall, the peak sizes and patterns remained stable between independent PCRs and FLA experiments. The average amplicon sizes and standard deviations for each locus determined from ten such experiments are shown in Table 1.3.4. The multiplex-PCR and FLA reproducibility was also verified for clinical isolates. Samples 28, 32, 35, 39 and 42 were subjected to multiplex-PCR and FLA at least 3 times, twice with Qiagen Multiplex enzyme and once with a high fidelity enzyme (AmpliTaq Gold, Applied Biosystems). Concordant FLA peaks were observed from all the three runs, which also demonstrated that the variation due to the DNA polymerase is minimal. Furthermore, when stored multiplex-PCR products were re-run for FLA, there were no significant changes in the profiles indicating that products can be stored (up to a year at -20°C) and analyzed at a subsequent convenient time and location.

1.3.5 DISCUSSION

The goal of our studies is to generate useful molecular epidemiological tools for monitoring the residual leprosy burden in endemic countries to complement local surveillance and control programs. Prolonged contact with a leprosy patient is a known risk factor, but only 10-30% of the new cases can be linked to an index case (3, 26); the number of new cases among children is not insignificant (a sign of recent transmission); isolates resistant to one or more antileprosy drugs are being detected in different endemic

countries and relapse rates amongst treated cases are ~13% (1, 6, 9, 14). In addition the concept of zoonotic transmission has not been excluded (18). Thus finding source(s) and chains of transmission by strain typing is a practical application. Special considerations for molecular epidemiology of leprosy include the long doubling rate of *M. leprae*, followed by an incubation period of 2-8 years before symptoms may be perceived or diagnosed, and the lack of laboratory cultivation methods for *M. leprae*. Furthermore, outbreaks are not common, though intra-familial clusters exist, and pockets of hyper-endemicity have been reported (11, 12).

In this context, we pursued our strategy to identify polymorphic markers for *M. leprae* for strain typing applications (7). Of the common methods for infectious pathogens, we and others have explored short tandem repeat analyses for leprosy transmission, because nucleotide substitutions, inversions, recombinations, and transpositions had been found to be rare (4, 28). The preferred method for measuring evolutionary distance between two samples takes into account the allele sizes (number of repeating units) at each VNTR locus. During the course of our experiments, three novel SNPs were discovered, which lead to a model for the origin and global spread of leprosy (20).

In light of field studies indicating that 1-3 microsatellites were insufficient for detection of strain diversity and transmission of *M. leprae*, we exploited the availability of a laboratory collection of 21 human clinical specimens passaged in armadillo for the purpose of identifying new genomic loci. Independently, a study performed with a different sample set composed of 27 mouse derived human isolates, primarily from Japan, but also including one each from Korea and Indonesia, and three from Thailand,

showed that the list of polymorphic loci are equivalent to that from our study. Locus 10-4 is comprised of four 10 base repeats, with a consensus repeat sequence 'ttATTAATAA'. Polymorphisms were not found either in the copy number or in the non-conserved nucleotides present at the first and/or second position in any of the repeats, even though 10-4 is situated in a non-coding region. The 7-3 locus is contained within a putative functional open reading frame in which VNTRs would cause frame shifting, perhaps the reason that the 7-3 locus was not polymorphic. The 23-3 locus was peculiar because the TN pattern with 3 copies was rare. We discovered a 5 base deletion in the third repeat in two isolates from Ethiopia. We have termed this allele as 23-3'. Zhang et al, showed that 32 of the 44 loci in our first list are polymorphic; of these, they selected 9 loci (all microsatellites) for determining strain identity in isolates collected from families with multiple cases (34).

To improve throughput, we have developed a multiplex-PCR approach suitable for rapid amplification of 15 VNTR loci from *M. leprae* containing specimens like slit skin smears and skin biopsies. The benefits are reduced PCR reagent consumption including DNA template and ease of set up. Of these 15 VNTR loci, 6 are minisatellites and 9 are microsatellites. Two more loci [6-3a and (TA)10] have since been included in the PCR. Microsatellites tend to exhibit higher allelic diversity even in localized collections, and when combined with minisatellites are useful in differentiating specimens from different countries/geographical regions (Table 1.3.5). For the 21 armadillo samples plus that of TN strain, a gene diversity of 0.5463 was observed for 16 loci. A similar trend was seen for the 30 clinical isolates from Philippines (gene diversity of 0.54 for 15 loci). For the same loci, the gene diversity is 0.6 when applied to the 27

(AC)8a		(AC)8b		(AC)9	
Allele	**Frequency**	**AC8b**	**Frequency**	**AC9**	**Frequency**
8	6.90	6	7.14	8	20.69
9	72.41	7	25.00	9	65.52
10	17.24	8	67.86	10	13.79
11	3.45				
(AT)15		**(AT)17**		**(TA)18**	
Allele	**Frequency**	**Allele**	**Frequency**	**Allele**	**Frequency**
11	3.45	12	13.79	13	3.45
12	3.45	13	13.79	14	6.90
13	13.79	14	17.24	15	6.90
14	3.45	15	20.69	16	6.90
15	24.14	16	20.69	17	13.79
16	20.69	17	3.45	18	17.24
17	6.90	18	10.34	19	3.45
18	3.45			20	3.45
19	6.90			21	13.79
20	6.90			22	10.34
22	6.90			25	3.45
				26	10.34
(GGT)5		**(GTA)9**		**(GAA)21**	
Allele	**Frequency**	**Allele**	**Frequency**	**Allele**	**Frequency**
4	15.38	8	10.71	14	3.45
5	80.77	9	46.43	16	6.90
6	3.85	10	17.86	19	6.90
		11	17.86	20	6.90
		12	3.57	22	13.79
		13	3.57	23	13.79
				24	13.79
				25	6.90
				26	6.90
				28	3.45
				29	3.45
				30	3.45
				31	6.90
				36	3.45
6-7		**12-5**		**18-8**	
Allele	**Frequency**	**Allele**	**Frequency**	**Allele**	**Frequency**
4	3.45	4	55.17	7	14.29
6	13.79	5	44.83	8	85.71
7	58.62				
8	24.14				
21-3		**23-3**		**27-5**	
Allele	**Frequency**	**Allele**	**Frequency**	**Allele**	**Frequency**
1	17.24	2	93.10	4	3.45
2	3.45	3	6.90	5	96.55
3	75.86				
4	3.45				

TABLE 1.3.5. Allele diversity detected for *M. leprae* STR loci in leprosy patients from Philippines.

isolates (mainly of Japanese origin) reported by Zhang et al (34). These diversity indices are comparable to those in other bacterial VNTR based typing systems. In *B. anthracis*, an 8 VNTR locus system was used for strain typing. An average diversity of 0.52 for 89 isolates was shown (10). In *S. enterica serotype Enteritidis*, Malorny et al, demonstrated a typing system based on nine VNTR loci (individual allelic diversity ranging from 0.07 to 0.65) which allowed further discrimination of isolates within single phage types (15). The average diversity would be 0.42 for this collection of 240 isolates. In the context of strain typing of *M. tuberculosis*, different genomic markers and combinations have been utilized. The locus diversity varied from 0.1 to 0.8 when 24 MIRU-VNTR loci were applied to a set of 90 isolates from a standardized worldwide collection. These loci have been recommended for high resolution phylogenetic studies (24). This dataset would represent an average genetic diversity of 0.58 for the 24 loci.

In addition, we developed the FLA method for mapping VNTR alleles, which is more rapid than the conventional sequencing method. To validate and standardize the technique, reference strain NHDP63 was used (Table 1.3.4) and included every time clinical samples were analyzed. Although the minisatellites banding profiles can be easily interpreted by visual inspection after electrophoresis on conventional agarose gels, it was convenient to include these in the multiplex-PCR combinations for FLA. Rapid mapping of 15 VNTRs, in a set of 30 clinical specimens obtained from one clinic has been demonstrated using the multiplex-PCR FLA technique. The amplicon sizes were found to be reproducible from run to run for the reference stain and the clinical isolates. However, the sizes (in bp) do not correspond to the calculated size of the DNA fragments (Table 1.3.4). This could be due to the nature of the gel matrix, differences during migration

between labeled PCR fragments, capillary length, ladder used for calibration, to the sequence composition and the models of Applied Biosystems Genetic Analyzers. So we recommend that a reference strain (such as NHDP63 in our experiments) be used in all of the runs for consistency in allele definitions for a given instrument within a laboratory. The allelic diversity for *M. leprae* in the Philippine sample collection is shown in Table 1.3.5.

For the minisatellite 18-8, minus stutter peaks were observed. These appear at the 2 and 3 copy positions. Even though some of the microsatellites are prone to the stutter phenomenon (*in vitro* and *in vivo*) and 3' tailing effect, electropherograms can be interpreted to assign alleles. The removal of the 3' tail artifacts by a high fidelity enzyme minimized the complexity of the peaks as was seen for the (AT)15 products (not shown). Stutter products are also often observed in conventional sequence traces, wherein the stutter product sequences merge making it difficult to read the sequence at the end of the repeat region. However in FLA, the main product and the stutter products are clearly discernible. This suggests that the FLA technique is more convenient for detecting such products and reporting the predominant alleles than by conventional sequencing. Strain typing allows us also to compare and confirm identity of laboratory reference strains. For example, the VNTR pattern of the Thai-53 isolate in our study is quite different from that described by Zhang et al, but is probably the same as that in Monot et al. The original strain was isolated in Japan from a Thai patient (25, 34).

VNTR stability to the extent of two passages in armadillo was noted in our studies except for two *M. leprae* suspensions. With the information available to us for the archived tissues, we surmise that 2936/41 and 3035/299 originated from one common

original inoculation in 1983. Tissue or animal 2936/41 represents the zero passage (patient to armadillo), while 3035/299 is that of the 2nd passage (armadillo to armadillo, date not known). Unfortunately, the intermediate 'first passage tissue' was not in our collection when we started this study. Therefore, we were unable to test or demonstrate when or how an additional VNTR profile emerged in the 2nd passage isolate. We could only propose that the alternate alleles in 3035/299 represent a secondary strain, rather than arising from VNTR instability sometime during incubation or passage because 11/16 loci carried two alleles. Endogenous infection of the host armadillos is another possibility. Similarly, regarding samples I480, I478 and A606, we inferred that they were derived from the same human inoculum. The zero passage animal was also common for all three. I480 and I478 were from 'first passage' infections, but in two different animals. These tissue specimens gave matching *M. leprae* VNTR patterns. The A606, which is a 2nd passage tissue, did not originate directly from either from I480 or I478, but from a different animal. The tissues corresponding to the zero passage and the precursor to A606 were not present in our sample bank for this study, so we cannot speculate on the origins of the different *M. leprae* genotype seen in A606 (alleles at 11/16 loci were altered) as compared to I480 and I478. For those armadillos infected with ~ 3 patient biopsy homogenates, it appears that one isolate was dominant in the first two passages. Truman et al, demonstrated that microsatellites were highly stable after multiple passages in armadillo and/or nude mice, albeit allele shifts by one or more repeat units were not infrequent (25). Matsuoka et al, demonstrated that the TTC (GAA) locus was stable after 11 passages in nude mice (16).

This study was undertaken mainly for developing methods for VNTR based strain typing. There was limited epidemiological or clinical information associated with the 30 coded clinical frozen samples (Table 1.3.3). However, it was possible to identify four isolates (17, 19, 25 and 33) which have the alleles 4 and 1 for the loci (GGT)5 and 21-3 respectively. These isolates cluster and separate when analyzed by the neighbor joining method (data not shown). For the same loci, the alleles are 6 and 3 for the pair of isolates (28 and 29) which also cluster, but into a different branch. There was no association with genotype and village or city. In a larger study which included these 30 samples and 170 more that were collected from patients in the same province since 2003, a well defined population structure, with distinct VNTR signatures was identified. Several pairs of isolates with matching genotypes associated with multi-case families or communities could be noted (Sakamuri et al, submitted for publication).

Our future work is focused on improving sensitivity, detection of mixed VNTR patterns, automated allele calling, elimination of non specific peaks (usually seen in samples with low bacterial numbers), with regard to the multiplex-PCR and FLA. We have initiated the generation of databases that include all available polymorphism data along with relevant clinical information for each sample so as to enable molecular geneticists and clinical epidemiologists objective means for tracing leprosy and implementation of preventive measures. These approaches were explored in the two accompanying paper (Sakamuri et al., submitted for publication).

1.3.6 ACKNOWLEDGEMENTS

This work was supported by NIH- NIAID grants AI-59644, AI-063457, AI-47197 and contract NOI AI-25469.

1.3.7 REFERENCES

1. Annual Report. 2003. International Federation of Anti-Leprosy Associations. The situation of activities supported by ILEP in 2002.

2. **Blake, L. A., B. C. West, C. H. Lary, and Todd JR 4th.** 1987. Environmental nonhuman sources of leprosy. Rev. Infect. Dis. **9**:562-577.

3. **Chen, S., L. Zhang, D. Liu, and B. Liu.** 2003. Should household contact examination in a low endemic situation of leprosy continue? Int. J. Lepr. Other Mycobact. Dis. **71**:95-100.

4. **Cole, S. T., P. Supply, and N. Honoré.** 2001. Repetitive sequences in *Mycobacterium leprae* and their impact on genome plasticity. Lepr. Rev. **72**:449-461.

5. **Cole, S.T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honoré, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature. **409**:1007-1011.

6. **dela Cruz, E., R. V. Cellona, M. V. Balagon, L. G. Villahermosa, T. T. Fajardo Jr, R. M. Abalos, E. V. Tan, and Walsh GP.** 1996. Primary dapsone resistance in Cebu, The Philippines; cause for concern. Int. J. Lepr. Other Mycobact. Dis. **64**:253-256.

7. **Groathouse, N. A., B. Rivoire, H. Kim, H. Lee, S. N. Cho, P. J. Brennan, and V. D. Vissa.** 2004. Multiple polymorphic loci for molecular typing of strains of *Mycobacterium leprae*. J. Clin. Microbiol. **42**:1666-1672.

8. **Izumi, S.** 1999. Subclinical infection by *Mycobacterium leprae*. Int. J. Lepr. Other Mycobact. Dis. **67**:S67-71.

9. **Jain, S., R. G. Reddy, S. N. Osmani, D. N. Lockwood, and S. Suneetha.** 2002. Childhood leprosy in an urban clinic, Hyderabad, India: clinical presentation and the role of household contacts. Lepr. Rev. **73**:248-253.

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

11. **Kerr-Pontes, L. R., A. C. Montenegro, M. L. Barreto, G. L. Werneck, and H. Feldmeier.** 2004. Inequality and leprosy in Northeast Brazil: an ecological study. Int. J. Epidemiol. **33**:262-269.

12. **Kumar, A., A. Girdhar, and B. K. Girdhar.** 2003. Epidemiology of leprosy in urban Agra. Lepr. Rev. **74**:31-34.

13. **Levy, L.** 1976. Studies of the mouse foot pad technique for cultivation of *Mycobacterium leprae*. 3. Doubling time during logarithmic multiplication. Lepr. Rev. **47**:103-106.

14. **Maeda, S., M. Matsuoka, N. Nakata, M. Kai, Y. Maeda, K. Hashimoto, H. Kimura, K. Kobayashi, and Y. Kashiwabara.** 2001. Multidrug resistant *Mycobacterium leprae* from patients with leprosy. Antimicrob. Agents Chemother. **45**:3635-3639.

15. **Malorny, B., E. Junker and R. Helmuth.** 2008. Multi-locus variable-number tandem repeat analysis for outbreak studies of *Salmonella enterica* serotype *Enteritidis*. BMC Microbiol. **8**:84.

16. **Matsuoka, M., L. Zhang, T. Budiawan, K. Saeki, and S. Izumi.** 2004. Genotyping of *Mycobacterium leprae* on the basis of the polymorphism of TTC repeats for analysis of leprosy transmission. J. Clin. Microbiol. **42**:741-745.

17. **Matsuoka, M., S. Maeda, M. Kai, N. Nakata, G-T. Chae, T. P. Gillis, K. Kobayashi, S. Izumi, and Y. Kashiwabara.** 2000. *Mycobacterium leprae* typing by genomic diversity and global distribution of genotypes. Int. J. Leprosy **68**: 121–128.

18. **Meyers, W. M., B. J. Gormus, and G. P. Walsh.** 1992. Nonhuman sources of leprosy. Int. J. Lepr. Other Mycobact. Dis. **60**:477-480.

19. **Monot, M., N. Honoré, N. Balière, B. Ji, S. Sow, P. J. Brennan, and S. T. Cole.** 2008. Are variable-number tandem repeats appropriate for genotyping *Mycobacterium leprae*? J. Clin. Microbiol. **46**:2291-2297.

20. **Monot, M., N. Honoré, T. Garnier, R. Araoz, J. Y. Coppée, C. Lacroix, S. Sow, J. S. Spencer, R. W. Truman, D. L. Williams, R. Gelber, M. Virmond, B. Flageul,**

S. N. Cho, B. Ji, A. Paniz-Mondolfi, J. Convit, S. Young, P. E. Fine, V. Rasolofo, P. J. Brennan, and S. T. Cole. 2005. On the origin of leprosy. Science. **308**:1040-1042.

21. **Nei, M. 1987.** Molecular Evolutionary Genetics, Columbia University Press, New York, NY, USA.

22. **Rees, R. J.** 1983. Progress in the preparation of an antileprosy vaccine from armadillo-derived *Mycobacterium leprae*. Int. J. Lepr. Other Mycobact. Dis. **51**:515-518.

23. **Shin, Y. C., H. Lee, H. Lee, G. P. Walsh, J. D. Kim, and S. N. Cho.** 2000. Variable numbers of TTC repeats in *Mycobacterium leprae* DNA from leprosy patients and use in strain differentiation. J. Clin. Microbiol. **38**:4535-4538.

24. **Supply, P, C. Allix, S. Lesjean, M. Cardoso-Oelemann, S. Rüsch-Gerdes, E. Willery, E. Savine, P. de Haas, H. van Deutekom, S. Roring, P. Bifani, N. Kurepina, B. Kreiswirth, C. Sola, N. Rastogi, V. Vatin, M. C. Gutierrez, M. Fauville, S. Niemann, R. Skuce, K. Kremer, C. Locht, and D. van Soolingen.** 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. J. Clin. Microbiol. **44**:4498-4510.
25. **Truman, R., A. B. Fontes, A. B. De Miranda, P. Suffys, and T. Gillis.** 2004. Genotypic variation and stability of four variable-number tandem repeats and their suitability for discriminating strains of *Mycobacterium leprae*. J. Clin. Microbiol. **42**:2558-2565.

26. **van Beers, S. M., M. Hatta, and P. R. Klatser.** 1999. Patient contact is the major determinant in incident leprosy: implications for future control. Int. J. Lepr. Other Mycobact. Dis. **67**:119-128.

27. Weekly Epidemiological Record. 2008. No. 33, 2008, 83, 293–300.

28. **Williams, D. L., T. P. Gillis, and F. Portaels.** 1990. Geographically distinct isolates of *Mycobacterium leprae* exhibit no genotypic diversity by restriction fragment-length polymorphism analysis. Mol Microbiol. **4**:1653-1659.

29. World Health Assembly, Elimination of leprosy: resolution of the 44th World Health Assembly. Geneva: World Health Organization; 1991 (Resolution No. WHA 44.9).

30. World Health Organization. 1982. Chemotherapy of leprosy for control programme. Technical Report Series. 768.

31. World Health Organization. 2003. The final push strategy to eliminate leprosy as a public health problem. Questions and answers. Second edition.

32. **Yoon, K. H., S. N. Cho, M. K. Lee, R. M. Abalos, R. V. Cellona, T. T. Fajardo Jr, L. S. Guido, E. C. Dela Cruz, G. P. Walsh, and J. D. Kim.** 1993. Evaluation of polymerase chain reaction amplification of *Mycobacterium leprae*-specific repetitive sequence in biopsy specimens from leprosy patients. J. Clin. Microbiol. **31**:895-899.

33. **Young, S. K., G. M. Taylor, S. Jain, L. M. Suneetha, S. Suneetha, D. N. Lockwood, and D. B. Young.** 2004. Microsatellite mapping of *Mycobacterium leprae* populations in infected humans. J. Clin. Microbiol. **42**:4931-4936.

34. **Zhang, L., T. Budiawan, and M. Matsuoka.** 2005. Diversity of potential short tandem repeats in *Mycobacterium leprae* and application for molecular typing. J. Clin. Microbiol. **43**:5221-5229.

Chapter 1.4

POPULATION BASED MOLECULAR EPIDEMIOLOGY OF LEPROSY IN CEBU, PHILIPPINES

This chapter has been published in *Journal of Clinical Microbiology*
Sakamuri et al, 2009

Clinical samples were obtained from LWM, Cebu, Philippines in collaboration with Dr. Robert Gelber Dr. Marivic Balagon and Dr. Sang Nae Cho. I contributed in the sample preparation, data collection, analysis and preparation of manuscript. PCR-RFLP method for SNP subtyping was designed by Dr. Vissa and Miyako Kimura. PCR-RFLP data was contributed by Dr. Wei Li (Post doctoral fellow) in our laboratory.

Sakamuri RM, Kimura M, Li W, Kim HC, Lee H, Kiran MD, Black WC 4th, Balagon M, Gelber R, Cho SN, Brennan PJ, Vissa V. Population-based molecular epidemiology of leprosy in Cebu, Philippines. J Clin Microbiol. 2009 Sep;47(9):2844-54. Epub 2009 Jul 1. PubMed PMID: 19571027;

1.4.1 ABSTRACT

To address the persisting problem of leprosy in Cebu, Philippines, we compiled a database of more than 200 patients who attend an established referral skin clinic. We described the patient characteristics in conventional demographic parameters, and also applied multiple locus variable number of tandem repeat (VNTR) analysis (MLVA) and single nucleotide polymorphism (SNP) typing for the *Mycobacterium leprae* in the biopsied skin lesions. These combined approaches revealed that transmission is ongoing with the affected including the young Cebuano population under 40 years of age in both crowded cities and rural areas in the island. Emergence of multicase families (MCF) is indicative of infection unconstrained by standard care measures. For the SNPs we designed a low cost polymerase chain reaction-restriction fragment length polymorphism typing method. MLVA in *M. leprae* was highly discriminatory in this population, yet could retain broad groups, as defined by the more stable SNPs implying temporal marker stability suitable for interpreting population structures and evolution. The majority of isolates belong to an Asian lineage (SNP type 1) and the rest to a putative post colonial lineage (SNP type 3). Specific alleles at two VNTR loci, (GGT)5 and 21-3 were highly associated with SNP type 3 in this population. MLVA identified *M. leprae* genotype associations in patients with known epidemiological links such as in MCFs and in some villages. These methods provide a molecular database and a rational framework for targeted approaches to search and confirm leprosy transmission in various scenarios.

1.4.2 INTRODUCTION:

During the last 4-5 years, genetic variation in *M. leprae* has been investigated for the purpose of strain typing. Although the *M. leprae* genome (4) has undergone reductive evolution and is highly mutated, limited genome variability has been found between global isolates and, except for loci prone to mutation such as variable-number of tandem repeats (VNTR) (7, 8, 12, 13, 16, 19, 22, 24, 25, 26), *M. leprae* are highly clonal species. Three single nucleotide polymorphisms (SNPs) were subsequently discovered by further genome sequencing efforts that allowed separation of global isolates into four subtypes (14).

Formal, systematic study of the *M. leprae* diversity by the application of the known polymorphic markers in defined endemic settings for studying extant population structures and leprosy transmission is limited. Previously, we presented the outcome of a focused study in Qiubei County in Yunnan Province in South West China (22). Using VNTR loci, we discovered subgroups within a major lineage. Differential geographical distribution of these subgroups was seen across the county. Furthermore, we noted conservation of *M. leprae* genotypes carried by patients of multicase families, indicative of localized transmission from shared sources.

We now extend such approaches to Cebu, Philippines, an island in the Central Visayas where leprosy is still in existence. From a case detection rate of 5.1 in 2001, to the current rate of 4.19 with an actual of about 300 registered cases for the entire island of Cebu, Multibacillary cases (MB) have consistently comprised 85%-90% of the total number of registered cases. We describe observations from conventional epidemiological and novel molecular studies. Currently, we can routinely map up to 15 VNTR loci using

multiplex-PCR and fragment length analysis methods (8). In this study, we demonstrate the development, feasibility and applicability of a low cost technology, PCR-RFLP (Polymerase chain reaction- Restriction Fragment Length Polymorphism) scheme for rapid SNP sub-typing of *M. leprae* DNA (14).

1.4.3 MATERIALS AND METHODS:

1.4.3.1 Armadillo derived *M. leprae* (ADML) strains of human origin

Four armadillo derived *M. leprae* (ADML) genomic DNAs (Thai-53, 3039/321, NHDP-63, and BR4923) were used as reference strains and taken from a panel of archived samples at CSU. These four isolates were previously characterized as belonging to the *M. leprae* SNP type 1, 2, 3 and 4 respectively (14). *M. leprae* DNA from the ADML isolates was isolated using the procedure described in our earlier paper (7).

1.4.3.2 Description of clinical study site, patients and biological samples

The patients in this study were drawn from those who consulted at the Cebu Skin Clinic (CSC), Leonard Wood Memorial Leprosy Research Centre, Philippines. For *M. leprae* strain typing purposes, we accessed skin biopsies and/or DNA extracts from three separate studies performed in collaboration with CSC staff. A total of 228 samples were thus compiled. These include DNA extracted from skin biopsies of patients from the 1980s (30 biopsy samples) (8); a collection of 100 frozen sectioned skin biopsies (from newly diagnosed untreated patients or those in the early stages of leprosy treatment) during the 2003-2004 period; and 98 ethanol-fixed skin biopsies from an ongoing study, started in 2006. These three sets of samples from the 1980s, 2003-2004 and 2006-2007 are henceforth referred to as 'R', 'M' and 'L' respectively.

Besides obtaining basic details of the medical history including the leprosy condition and treatment regimen, information about the patients' geographical area of residence (village, city and province), types /place(s) of employment and education, nature of dwelling, sources of drinking, cooking, and bathing water spanning a prior 20 year period, was captured with a standardized questionnaire at the time of recruitment since the 'L' study began. All studies involving biological sample collections were performed following approval from the governing human research ethical committee and informed consent procedures as necessary.

1.4.3.3 DNA extraction and multiplex-PCR, fragment length analysis (FLA) for 15 VNTRs

DNA was extracted from the biopsies using Qiagen DNeasy Tissue kit (Qiagen, Valencia, CA) as described earlier (7). A multiplex-PCR protocol described by Kimura et al comprising four reactions for the amplification of fifteen VNTR loci was achieved using the Multiplex PCR enzyme kit (Qiagen) and fluorescent 5' labeled forward primers and reverse unlabeled primers (8). These fifteen VNTRs include nine microsatellites [(AC)8a, (AC)8b, (AC)9, (AT)15, (AT)17, (TA)18, (GGT)5, (GTA)9, (TTC)21] and six minisatellites [6-7, 12-5, 18-8, 21-3, 23-3, and 27-5]. Multiplex-PCR products were separated by capillary electrophoresis [Applied Biosystems Genetic Analyzer 3130 at Macromolecular Resource Facility (MRF), CSU] and fragment length analysis to determine the major allele for each VNTR locus using GeneMapper Version 3.7 software. When DNA sequencing was performed, a multiplex-PCR sample was diluted 10 fold; 1-2 µl was combined with 10 pmoles of the unlabelled forward primer for the ABI Big Dye Cycle sequencing reaction at MRF, CSU (8).

1.4.3.4 Development of a PCR-Restriction fragment length polymorphism method for SNP subtyping of *M. leprae*

Known *M. leprae* SNP loci 1, 2 and 3 (nucleotide positions 14,676, 1,642,875, and 2,935,685 respectively on the sequenced TN strain) were amplified using published primer sequences (14). Single or multiplex PCRs were set up using 1-2 µl of the DNA extract. An additional forward primer was designed for locus 2 for a shorter product of 114 bp instead of 190 bp. The High Fidelity Super Mix (Invitrogen, Carlsbad, CA) was used for PCRs (final volume of 20 µl). The DNA products were resolved by agarose gel electrophoresis and detected by ethidium bromide staining.

The three SNP containing sequences were examined for restriction sites overlapping the SNP using NEBcutterV2.0 (http://tools.neb.com/NEBcutter2/index.php). Restriction digest assays were developed using *Sml*I , *CviKI*-1 and *BstU* I for SNP loci 1, 2 and 3 respectively (New England Biolabs, MA). The PCR products (5µl) were digested with 1 unit of the enzymes. The *Sml*I, *BstU*I and *CviK*I-1 reactions were performed at 55°C, 60°C and 37°C respectively for 1hr. The *Sml*I and *BstU*I uncut and cut DNAs were subjected to electrophoresis on 3% agarose gels. The uncut and *CviK*I-1 digested SNP locus 2 products were resolved on a 12% acrylamide gel. The gels were stained with ethidium bromide and visualized by UV trans-illumination.

1.4.3.5 Diversity and phylogeny

The Microsatellite Tool Kit (http://animalgenomics.ucd.ie/sdepark/ms-toolkit/) was used to calculate allele frequency, and average allelic diversity. PAUP4.0b software was used to apply the maximum parsimony algorithm with equal or unequal weighting for all the VNTR loci (17) for a maximum of 1000 trees. An allele difference of one

repeat unit (increase or decrease) per locus was calculated as one mutation step, difference of three as three steps, etc., as described previously (22). The 50% consensus trees were reported. The figure for the 50% majority rule tree was generated in PHYLIP program using the 'consensus' application (6).

1.4.4 RESULTS

1.4.4.1 Leprosy in Philippines and in the island province of Cebu

In 1998, about a decade after MDT was implemented nationwide, the Philippines reached its elimination goal with a prevalence of 0.9/ 10,000 with a total of 6,872 registered cases. Then after another decade of leprosy elimination campaigns, this has further gone down to a prevalence of 0.36/10,000 with a total of 3,156 registered cases. For Cebu, data from 2001-2006 period shows that the prevalence has sustained: 0.48 in 2001 and 0.46 in 2006.

Cebu Skin Clinic (CSC) diagnoses around 200 patients a year, more than half of the new cases in the island. The numbers of patients according to the Ridley-Jopling clinical classification is shown in Table 1.4.1. These patients come from all parts of the island province, voluntarily or as referrals. A large proportion of the CSC cases since 2003 are included in the current study.

Cebu is a narrow island, 225 km long and 25 km across at the widest point (Fig. 1.4.1A). The Metro Cebu comprises seven cities (Carcar City, Cebu City, Danao City, Lapu-Lapu City, Mandaue City, Naga City, and Talisay City) and 6 other municipalities on the east. Each city is administratively divided into barangays (akin to a county or village). Cebu City *per se* is located centrally within the Metro Cebu area, while Lapu

TYPE	*2000*	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*
IND	7	8	6	10	8	4	4	6
TT	6	4	0	5	2	4	1	0
BT	38	33	29	27	22	11	25	23
BB	15	10	8	4	11	13	2	0
BL	70	70	65	68	84	61	49	50
LL	69	62	72	64	86	80	83	94
TOTAL	**205**	**187**	**180**	**178**	**213**	**173**	**164**	**173**

IND, Indeterminate; TT Tuberculoid; BT, Borderline Tuberculoid; BB Borderline Borderline; BL Borderline Lepromatous; LL, Lepromatous Leprosy

TABLE 1.4.1 New leprosy cases diagnosed at Cebu Skin Clinic (CSC), LWM (2000-2007).

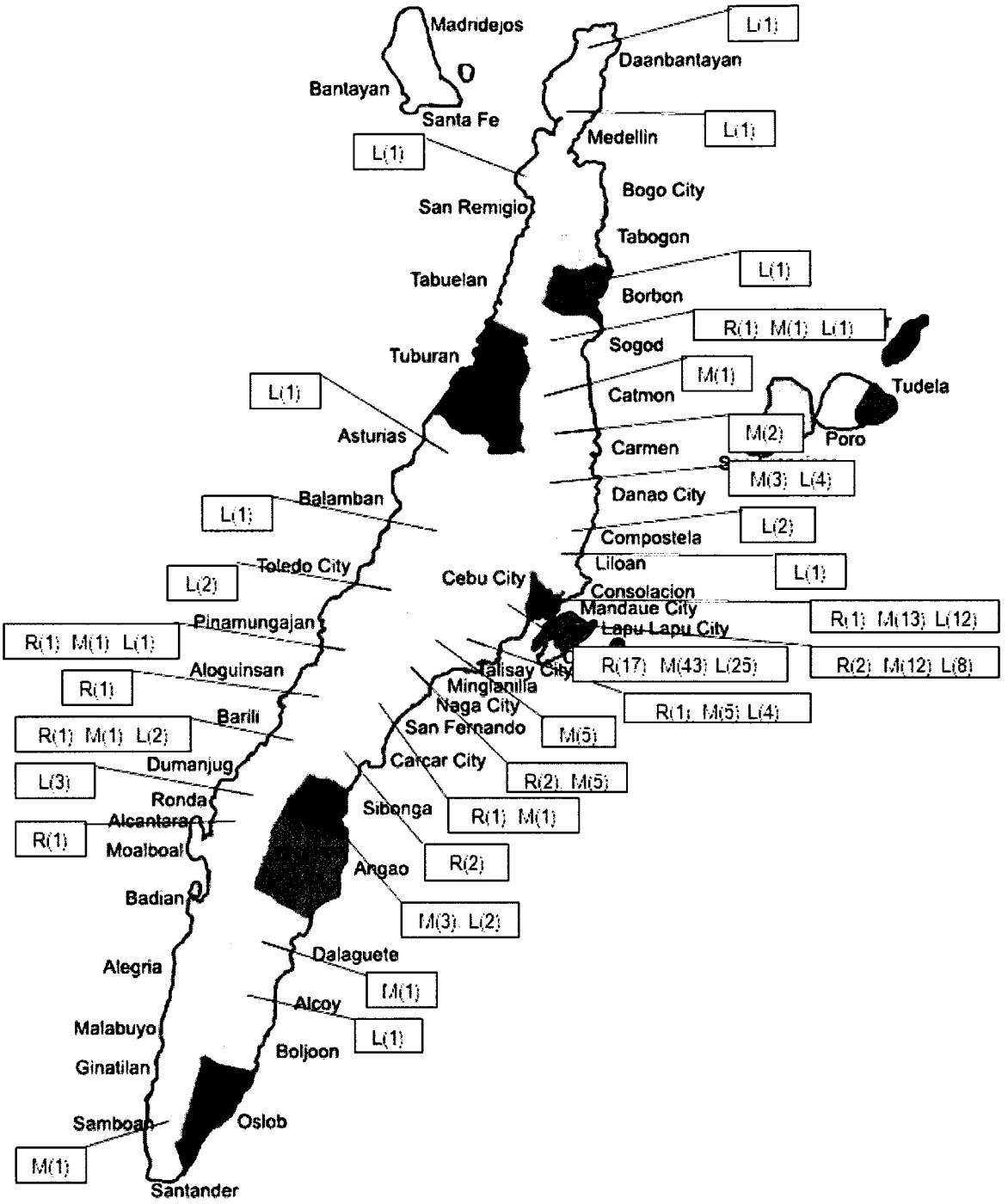


FIGURE 1.4.1A Map of Cebu Province, Philippines depicting patient's city of residence. One third of the patients in this study population are from the Cebu City. Letters R, M and L represent each of the sample collections followed in parenthesis by number of patients in this study.

Lapu City is in the Mactan island and is connected to the main island by a bridge. CSC is in Cebu City while the leprosarium is located in Mandaue City. It is slightly over 10 kms north of CSC. Of the 228 patients in this study, one third resides in Cebu City (Fig. 1.4.1B).

The majority of the patients were bacteriological index (BI) positive, as judged by microscopic examination of slit skin smears or biopsies and were prescribed the standard multibacillary WHO-MDT treatment (one year duration in the Philippines). Considering the 'L' and 'M' group (n=178), the age of the patients ranged from 5-71 years with a mean of 30 years. The proportion of patients under 30 and 40 years of age was 66%, and 80% respectively. The ratio of male to female patients was 3:1.

1.4.4.2 Multiple locus variable number of tandem repeat analysis (MLVA) based genomic variation of *M. leprae* in patients in Cebu, Philippines

Skin biopsies taken from leprosy patients at the time of diagnostic procedures or during treatment (as in the case of 38 of the 100 M samples) were the source of *M. leprae* DNA in these studies. MLVA was performed for a total of 228 specimens using multiplex-PCR and fragment length analysis (FLA) procedures described previously (8) to determine the number of repeats (hereafter known as alleles). The samples from patients (n=21) with low BI (BI < 1.5) did not yield data for all loci, and were excluded from the subsequent analysis.

The MLVA 15 locus data of the 207 Cebu samples, NHDP63 reference DNA and the sequenced *M. leprae* TN strain (4) were analyzed for allelic diversity and population structure. The allelic diversity across the 15 loci ranged from 0.1 to 0.94 and the heterozygosity for the entire data set is 0.5334 (21).

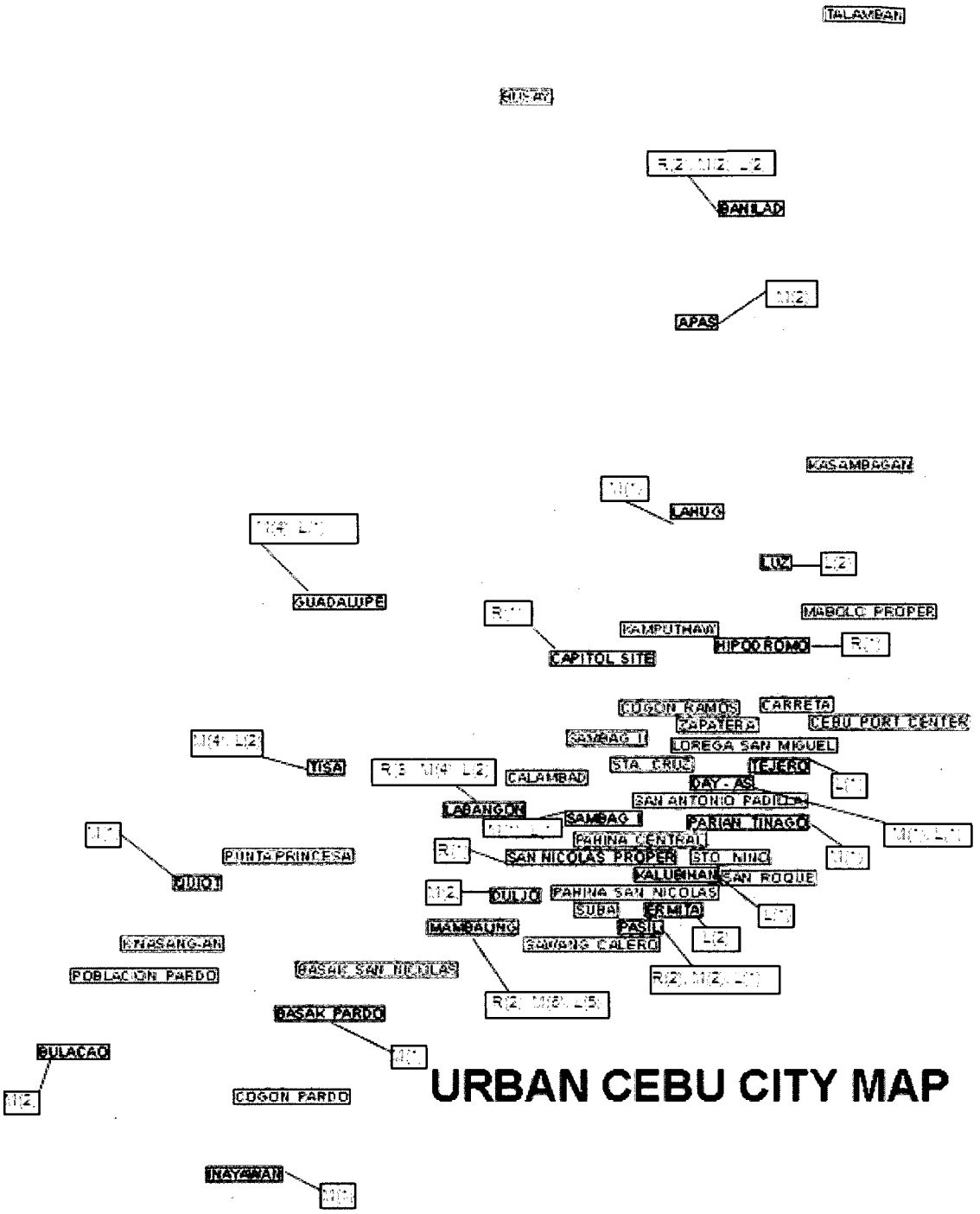


FIGURE 1.4.1B Distribution of patients from Cebu City according to barangay (village) of residence. Letters R, M and L represent each of the sample collections followed in parenthesis by number of patients in this study.

1.4.4.3 Comparison of *M. leprae* VNTR patterns in slit skin smears and skin biopsy

To compare the stability of VNTRs in different body compartments, VNTR data from ten high BI patients' biopsies and SSS samples were also analyzed (Table 1.4.2) and found to be mostly similar. We noted extra product peaks for the SSS in some of the more variable stutter prone microsatellites. Such patterns may represent variants of a dominant isolate. SSS is usually a mix of smears from the lesion site (biopsy) and as well as from other body sites.

1.4.4.4 Rapid PCR-RFLP SNP subtyping of *M. leprae* in patients in Cebu, Philippines

By using the NEBcutter2 program, we were able to identify commercially available enzymes suitable for detection and separation of the four SNP sub types according to the scheme shown in Fig. 1.4.2 and 1.4.3. The recognition sequences for, *Sml*I, *BstU*1 and *CviK*I-1 are CTYRAG, RGCY and CGCG respectively.

At first, we verified the feasibility of using one, two or three enzymes for SNP 1-4 sub-typing for reference *M. leprae* strains as shown in Fig. 1.4.3. The restriction enzyme products could be resolved on a 3% agarose gel for the *Sml*I and *BstU*I. A novel primer set was designed for the SNP 2 locus to eliminate non SNP *CviK*I-1 cutting sites. However, there are still two recognition sites one from the SNP 'T' allele that would yield fragments of 61, 11 and 42 bp; while amplicons lacking the SNP 'T' would produce 72 and 42 bp fragments. The 3% agarose gel system was not adequate to reliably differentiate the 72 and 61 bp *CviK*I-1 fragments; therefore a 12% acrylamide gel system was used.

	B01	SS01	B03	SS03	B07	SS07	B08	SS08	B09	SS09	B10	SS10	B11	SS11	B14	SS14	B19	SS19	B20	SS20
(AC)8b	7	7	6	6	8	8	8	8	8	8	6	6	8	8	7	7	8	8	8	8
(GTA)9	9	9	9	9	9	9	8	8	9	9	9	9	9	9	11	11	12	12	11	11
(GGT)5	4	4	5	5	5	5	ND	ND	5	5	5	5	5	5	5	ND	5	5	5	5
(AT)17	12	12	10	10	16	16	16	16	15	15	10	10	15	15/16	13	13	14	14	13	13
21-3	2	2	2	2	3	3	3	3	3	3	2	2	3	3	3	3	3	3	3	3
(AC)9	8	8	8	8	9	9	9	9	9	9	8	8	9	9	8	8	8	8	9	9
(AT)15	18	18	17	17	14	14	15	15	16	ND	20	ND	14	14	16	16	15	15	16	16
(AC)8a	8	8	9	9	8	8	9	9	9	9	10	10	9	9	9	9	9	9	9	9
27-5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6-7	7	7	7	7	7	7	9	9	7	7	7	7	7	7	8	8	7	7	7	7
(TA)18	13	13	16	16	18	18	17	17	19	19	17	16	18	18-20	18	18	18	18	17	17
(TTC)21	18	18	17	17	30	30/26	20	20	26	26/27	13	14	28	27/24	22	22	36	36	25	25
18-8	ND	ND	7	7	8	8	8	8	ND	ND	8	8	4	4	8	8	ND	ND	ND	ND
12-5	4	4	5	5	5	5	4	4	5	5	5	5	5	5	4	4	4	4	4	4
23-3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

ND not detected

TABLE 1.4.2 Comparison of *M. leprae* VNTR patterns between a skin biopsy (B) and pooled slit skin smears (S) in 10 patients

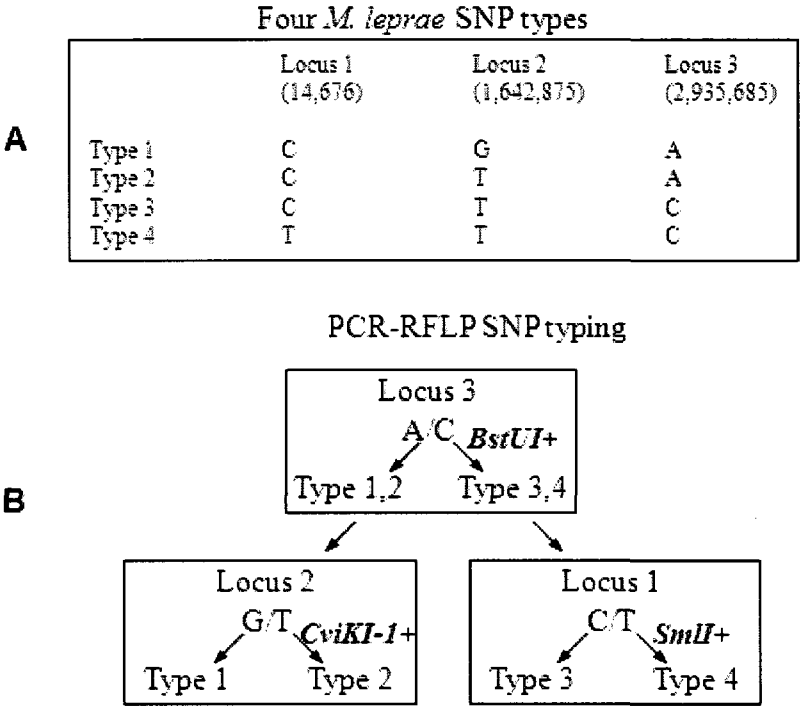


FIGURE 1.4.2 A schematic of the *M. leprae* SNP subtyping method for leprosy clinical samples based on PCR-RFLP. The four major SNP types are shown in panel A. The numbers below the SNP locus refer to the nucleotide position in the sequenced TN strain. Panel B shows the scheme.

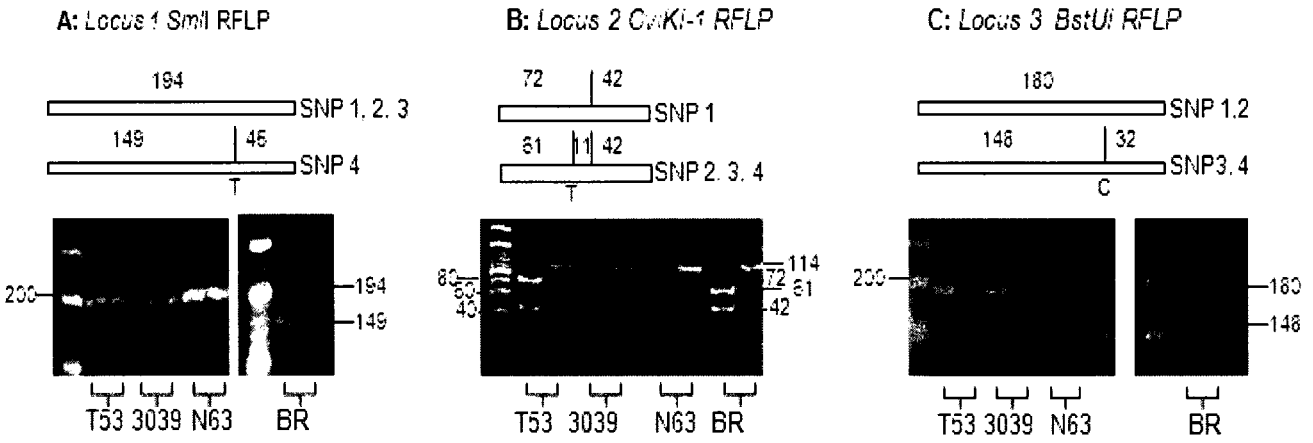


FIGURE 1.4.3 The PCR-RFLP pattern of four *M. leprae* reference strains. The PCR products of the SNP loci 1, 2 and 3 were subjected to enzyme digestion. The digested and undigested PCR products were resolved on agarose (panels A and C) or acrylamide gels (panel B) and run in pairs; the digested products for each of the strains are shown first. DNA sizing ladders are shown in the left lane in each gel. The numbers on the right side of each gel refer to the lengths (bp) of the PCR and digestion products.

After establishing the PCR-RFLP method for separating the SNP type 1-4 isolates, we applied it to all the 100 samples from the 'M' study. The first step of PCR amplification of the SNP locus 3 revealed that 99% of the DNAs yielded products and 18 samples could be digested with *BstU*I. One of the samples did not yield PCR products for all the three SNP loci. DNA sequencing of the PCR amplicons of the remaining samples confirmed that all *BstU*I sensitive amplicons carried the SNP allele 'C', while a random selection of the *BstU*I resistant samples carried the 'A' allele. In the next step, we examined susceptibility of the SNP locus 1 to *Sml*I. None of the amplicons were digested as was corroborated by the sequence of the products, and thus SNP- type 4 *M. leprae* strains are absent in this clinical sample set. Finally, to classify the strains carrying SNP 3 locus *BstU*I resistant phenotype into type 1 or type 2, the *CviK*I-1 enzyme digestion of the locus 2 PCR products was performed. One sample of type 2 was confirmed by sequencing at SNP loci 2 and 3. Another isolate that was restriction positive was sequenced confirmed to carry the 'T' allele; however PCR product for locus 3 was not obtained, so we were unable to classify as 2, 3 or 4. One of the samples gave mixed cut and un-cut *CviK*I-1 RFLP patterns. Furthermore, both the alleles G and T were observed in the DNA sequence. Interestingly the STR typing of this DNA (M55) also revealed mixed signals at multiple loci even when independent aliquots of biopsy sections of this sample were tested, indicating a mixed infection or contamination of the specimen. In summary, we were able to SNP subtype 97 of the 100 specimens.

1.4.4.5 VNTR alleles predictive of the SNP type 3 for Philippine isolates

Correlating the SNP subtypes with the VNTR profiles of 15 loci revealed that the alleles at the (GGT)5 and 21-3 loci were 4 and 1 or 2 respectively for the SNP type 3

isolates (Fig. 1.4.4A, B and C), while other combinations of alleles such as 4:3, 5:3, 5:1 and 5:2 were seen in isolates of the SNP- type 1 and 2.

To determine if this allelic signature is predictive of the SNP type 3 in this population, we reviewed the 207 VNTR patterns and identified 14 other SNP 3 candidates (10L and 4R). Of these we first tested 8L and 4R samples. As controls, we randomly tested other 'L' (n=6) and 'R' (n=1) samples that did not have the VNTR pattern of interest. All the 8L samples were of SNP type 3 (Fig. 1.4.5), while the 'controls' were not. The 'R' samples had degraded or very little was remaining at the time these new assays were being tested, and we could recover PCR amplicons for one sample (SNP type 3). Subsequently, we completed the PCR-RFLP SNP typing of the remaining 64L samples and demonstrated that the SNP 3 prediction rule was valid.

1.4.4.6 Population structure of *M. leprae* isolates from Cebu, Philippines

There are several computation models and software to infer clustering and/or phylogenetic relationships with the molecular dataset (1, 9, 10, 11, 15, 17). We applied parsimony (MP) principles, assumed a stepwise mutation model and executed the analysis using PAUP software program (17). Arbitrarily defining the sequenced TN strain as the outgroup taxon, 50% consensus trees were generated. Since the allelic diversity of each locus was found to be different in this study population, we weighted the loci according to the inverse of the individual allelic diversity index for MP. Neighbor joining (NJ) and UPGMA methods were also executed from the same data sets (data not shown) in PAUP.

Five major groups emerged in phylogenetic trees using MP (Fig. 1.4.6), NJ and UPGMA algorithms, which we have named A, B, C, D and T. The T group comprising

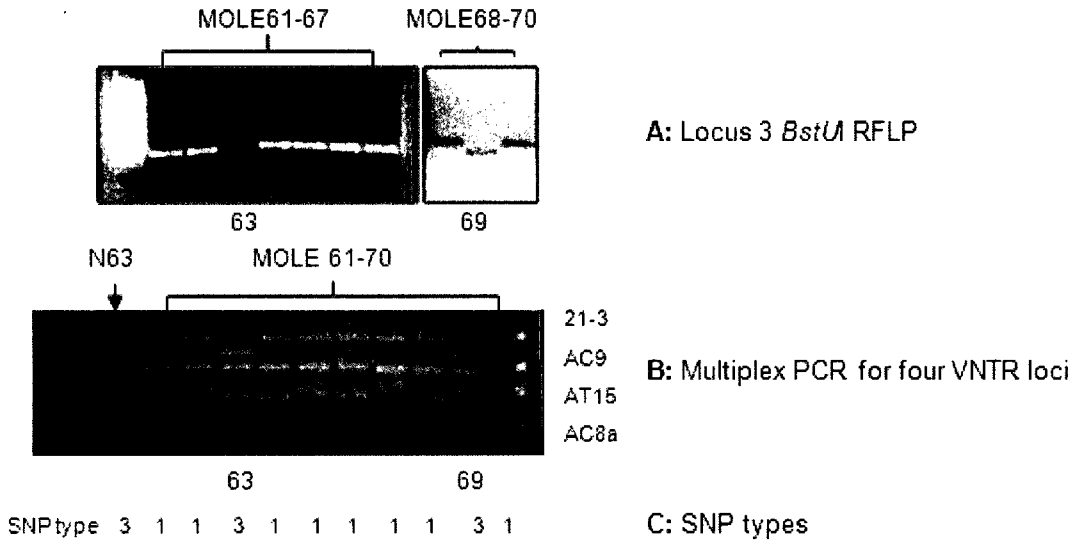
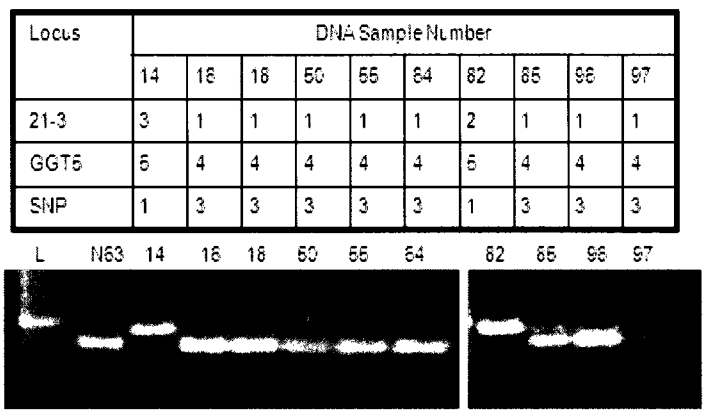
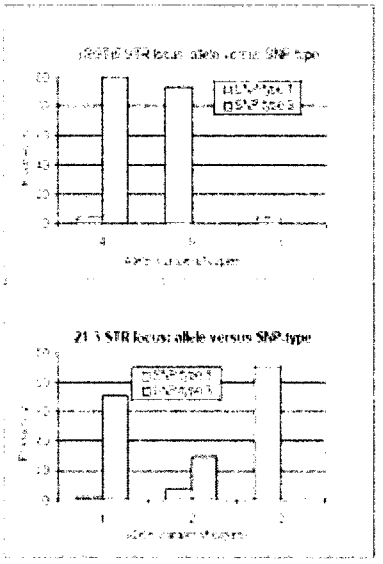


FIGURE 1.4.4 The relationship between 21-3 VNTR allele and the SNP locus *BstUI* cutting pattern for 10 Philippine *M. leprae* samples is shown in panels **A** and **B** respectively. Panel **A** is the *BstUI*-RFLP gel, and Panel **B** is an agarose gel showing products of multiplex PCR for four VNTR loci. The 21-3 product is the largest. VNTRs at all loci can be seen within this sample set. The SNP types are shown in Panel **C**. Samples 63 and 69 are indicated below panels **A** and **B**. N63 is NHDP63 (SNP type 3).

Locus	DNA Sample Number									
	14	16	18	50	55	64	82	85	96	97
21-3	3	1	1	1	1	1	2	1	1	1
GGT5	5	4	4	4	4	4	5	4	4	4
SNP	1	3	3	3	3	3	1	3	3	3



Locus 3 *BstUI* RFLP



The association of STR and SNP alleles in Cebu

FIGURE 1.4.5 The 21-3 and (GGT)5 alleles are indicative of SNP type 1 and 3 in Philippine *M. leprae*. The table indicates the VNTR alleles for 21-3 and (GGT)5 for 10 *M. leprae* specimens. The gel shows the corresponding SNP locus 3 *BstUI* cutting patterns for these 10 specimens. N63 is NHDP63 (SNP type 3). The graphs show the allele frequency of these loci for SNP type 1 and 3 for 100 *M. leprae* samples.

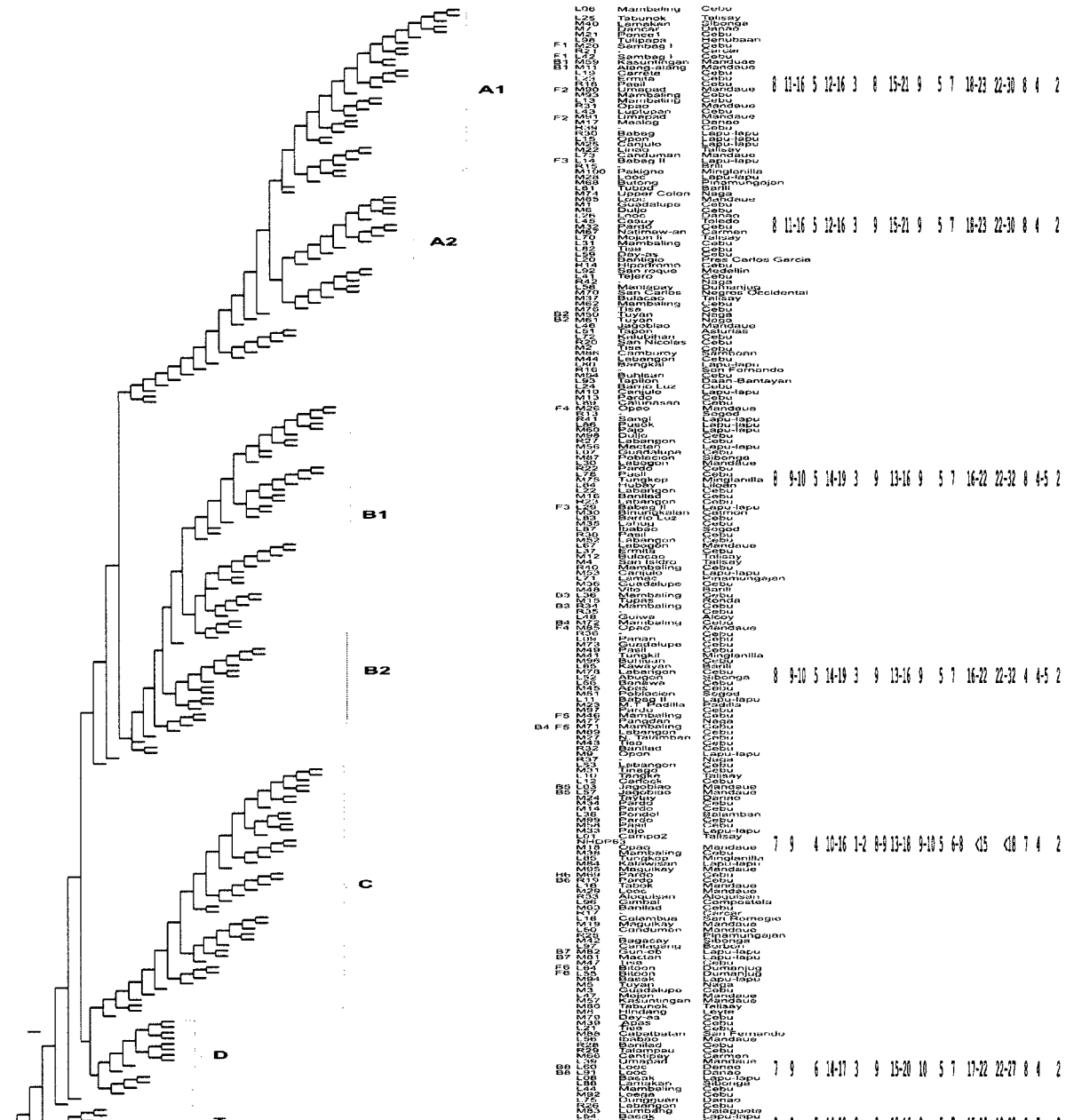


FIGURE 1.4.6 *M. leprae* population structure based on VNTR markers and the dominant genotypes identified in Cebu, Philippines. On the left, a fifty percent consensus phylogenetic tree generated using MP algorithm and the major branches are marked as A, B, C, D and T. The source of each *M. leprae* isolate is indicated in the three columns to the right: study (R, M and L) sample number, barangay and city respectively. The conserved allelic patterns within each branch are indicated at the extreme right. The locus order is (AC)8b, (GTA)9, (GGT)5, (AT)17, (21-3), (AC)9, (AT)15, (AC)8a, 27-5, 6-7, (TA)18, (TTC)21, 18-8, 12-5 and 23-3. Clustering of genotypes within multicase families and/or barangays is seen; represented by letter codes F1-F6 (Table 1.4.3), and B1-B8 (Table 1.4.4).

ten isolates is seen proximal to the sequenced TN strain. This grouping may be attributed to the 3 copies of 23-3 VNTR locus rare in the entire data set, and other shared alleles with the TN strain.

The majority of the remaining isolates fall within the groups A, B, C and D. Groups A and B are separated from C and D due to (AC)8b locus. Groups A and B have 8 copies of the VNTR, whereas in the groups C and D, the allele is 7. Group A is then branched into A1 and A2; the (AC)9 allele is 8 or 9 respectively. The 18-8 locus, with alleles of 8 or 4 separates B1 from B2.

The C group comprising 45 clinical isolates is characterized by low allele numbers for (AC)8b, (TA)18, (GGT)5, (TTC)21, 21-3 and 18-8 VNTR loci. The dominant alleles for (AC)8b, (GGT)5, and 18-8 VNTR loci are 7, 4 and 7. The range of alleles for (TA)18, (TTC)21 and 21-3 are 12-17, 14-20 and 1-2 respectively, while in A, B and D they are 21-26, 21-32 and 3 copy respectively (Fig. 1.4.6 and 1.4.7).

The D group is characterized by a distinctive 6 copy allele of the (GGT)5 locus and all isolates have 10 repeats at the (AC)8a locus.

The majority of the Cebu *M. leprae* isolates are of SNP type 1. Of the 207 samples for which VNTR patterns were mapped, 100M, 78L and 5R samples have been tested for SNP types. From these 10L, 18M and 1R sample were confirmed as SNP type 3. In the rest of the 'R' samples for which DNA is no longer available for testing, three more carry the 4: 1 alleles for (GGT)5:21-3. Therefore, we estimate that 16% (32/207) of the study population are of SNP type 3. Interestingly in trees that were derived solely based on VNTR alleles, the SNP type 3 isolates converged in Group C described above, separating from the remaining at a deep or intermediate branch level. The allele profiles

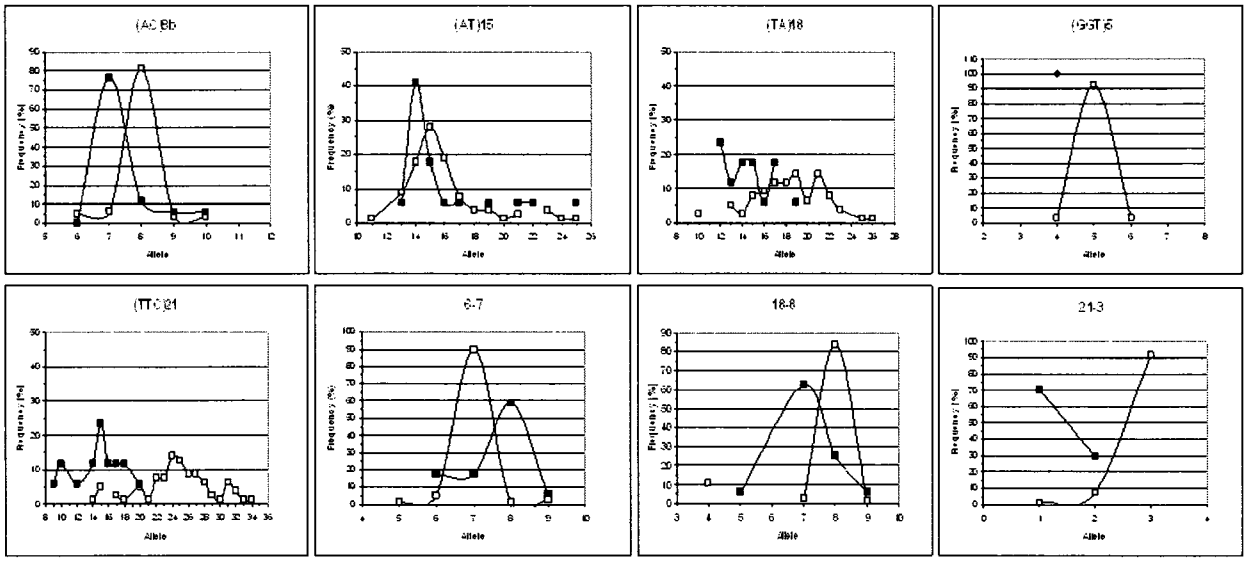


FIGURE 1.4.7 Comparison of VNTR patterns in SNP type 3 *versus* SNP type 1 isolates in Cebu, Philippines. The allelic frequencies (Y axis) were plotted against the VNTR alleles (X axis) for the locus indicated on the top of each panel. Solid squares (■) and hollow squares (□) represent SNP type 3 and SNP type 1 isolates respectively.

for the SNP type 3 versus SNP type 1 isolates are shown in Fig. 1.4.7. A few non-SNP type 3 (of SNP type 1, n= 13) isolates also group with the SNP type 3 isolates in Group C, probably due to their low allele numbers typically seen in the SNP type 3 isolates (Fig. 1.4.7) but not in the majority of the SNP type 1 isolates.

1.4.4.7 Identification of plausible transmission clusters in Cebu by genotyping

Exposure to the *M. leprae* source can occur within and outside the household, such as in the community, place of employment, education, etc. The barangay of residence at the time of sampling was known for nearly 200 patients. However, daily, weekly or monthly commute for work or other reasons is not uncommon for residents in the metro Cebu region.

The patient population studied included several MCFs. MCFs presenting in the clinic within a period of four years of each other, serve as a model for recent infections from one to the other, or from another common source. The 15 locus VNTR profiles of *M. leprae* from these patients are shown in Table 1.4.3. In general, the mismatched alleles in the MCFs occurred in the more polymorphic and stutter prone VNTR loci. In the two pairs F3 and F4 there were differences in many loci. Accordingly, four of the six MCF pairs, appeared closely in the consensus tree despite allelic differences at one or more loci (Fig. 1.4.6). Based on these trends, we tentatively define transmission clusters in Cebu as patients/isolates whose *M. leprae* VNTR alleles are shared in 12/15 loci (includes all minisatellites) and that co-localize in the weighted consensus tree.

Besides examining the genotypes of *M. leprae* in MCFs, we sorted the VNTR profiles according to the barangays. There is not much evidence of geographic segregation of strains according to VNTR or SNP profiles. However, several matched

	B01	SS01	B03	SS03	B07	SS07	B08	SS08	B09	SS09	B10	SS10	B11	SS11	B14	SS14	B19	SS19	B20	SS20
(AC)8b	7	7	6	6	8	8	8	8	8	8	6	6	8	8	7	7	8	8	8	8
(GTA)9	9	9	9	9	9	9	8	8	9	9	9	9	9	9	11	11	12	12	11	11
(GGT)5	4	4	5	5	5	5	ND	ND	5	5	5	5	5	5	5	ND	5	5	5	5
(AT)17	12	12	10	10	16	16	16	16	15	15	10	10	15	15/16	13	13	14	14	13	13
21-3	2	2	2	2	3	3	3	3	3	3	2	2	3	3	3	3	3	3	3	3
(AC)9	8	8	8	8	9	9	9	9	9	9	8	8	9	9	8	8	8	8	9	9
(AT)15	18	18	17	17	14	14	15	15	16	ND	20	ND	14	14	16	16	15	15	16	16
(AC)8a	8	8	9	9	8	8	9	9	9	9	10	10	9	9	9	9	9	9	9	9
27-5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6-7	7	7	7	7	7	7	9	9	7	7	7	7	7	7	8	8	7	7	7	7
(TA)18	13	13	16	16	18	18	17	17	19	19	17	16	18	18-20	18	18	18	18	17	17
(TTC)21	18	18	17	17	30	30/26	20	20	26	26/27	13	14	28	27/24	22	22	36	36	25	25
18-8	ND	ND	7	7	8	8	8	8	ND	ND	8	8	4	4	8	8	ND	ND	ND	ND
12-5	4	4	5	5	5	5	4	4	5	5	5	5	5	5	4	4	4	4	4	4
23-3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

ND not detected

TABLE 1.4.2 Comparison of *M. leprae* VNTR patterns between a skin biopsy (B) and pooled slit skin smears (S) in 10 patients

pairs emerged (Table 1.4.4) that are suitable for formal secondary epidemiological investigations of these cases and active search for other new cases in these barangays; activities not within the scope of the current study.

1.4.5 DISCUSSION

Global standardized guidelines for diagnosis and treatment of leprosy with a goal of elimination have been available and in practice for more than two decades (23). Yet, new patients are being detected at unchanged rates in some endemic regions leaving fundamental questions about source and routes of transmission and efficacy of leprosy control measures unanswered. Classical epidemiology and leprosy control programs have frequently identified household contacts of patients as being at higher risk of infection than the general population (2, 3, 20); explanations or tools are not routinely in place for tracing the origins and relationships of these and remaining new cases (5, 18)

In susceptible individuals leprosy manifests as single or multiple skin lesions often remaining unnoticed or not reported and undiagnosed for extended durations. Laboratory testing including slit skin smear testing and biopsy examinations are not universally required for diagnosis. *M. leprae* is non-cultivable in the laboratory even after detection. These and social factors impede tracing leprosy infections in communities. Nevertheless, molecular technologies of strain typing of the infectious agent directly from skin lesions and skin smear specimens when available are emerging and being explored to provide novel information and overcome these challenges in traditional leprosy control.

Residence Barangay	Barangay number-Sample ID	STR locus														
		AC8b	GTA9	GGT5	AT17	21-3	AC9	AT15	AC8a	27-5	6-7	TA18	TTC21	18-8	12-5	23-3
Alang-alang	B1-M11	8	13	5	13	3	8	17	9	5	7	19	29	8	4	2
	B1-M59	8	13	5	13	3	8	17	9	5	7	21	31	8	4	2
Tuyan	B2-M50	8	11	5	13	3	9	16	9	5	7	21	22	8	4	2
	B2-M61	8	11	5	13	3	9	16	9	5	7	22	22	8	4	2
Mambaling	B3-R34	8	9	5	15	3	9	15	9	5	7	21	27	8	5	2
	B3-L36	8	9	5	14	3	9	15	9	5	7	22	27	8	5	2
Mambaling	B4-M71	8	9	5	18	3	9	15	9	5	7	19	25	8	5	2
	B4-M72	8	9	5	16	3	9	16	9	5	7	22	25	8	5	2
Jagobiao	B5-L03	6	9	5	10	2	8	17	9	5	7	16	17	7	5	2
	B5-L57	6	9	5	10	2	8	18	9	5	7	17	17	7	5	2
Pardo	B6-R19	7	8	4	14	1	9	17	9	5	8	14	16	7	4	2
	B6-M69	7	8	4	14	1	9	19	9	5	8	14	16	7	4	2
Mactan	B7-M81	7	9	4	14	1	9	14	9	5	8	17	18	7	4	2
	B7-M82	7	9	4	14	1	9	14	9	5	8	19	16	7	4	2
Looc	B8-L60	7	9	6	17	3	9	17	10	5	7	19	22	8	4	2
	B8-L91	7	9	6	16	3	9	17	10	5	7	18	20	8	4	2

TABLE 1.4.4 Identification of pairs of *M. leprae* isolates with matching MLVA patterns in various barangays in Cebu

We present findings and interpretations on molecular strain typing of *M. leprae* based on fifteen VNTRs genomic markers by studying nearly 170 patients residing in Cebu, Philippines from a recent four year period. Another thirty patients from 20 years ago were also included. Sub-typing on the basis of three SNP markers was applied to approximately half the study population. In this regard, the Cebu study represents a comprehensive and descriptive molecular analysis encompassing a large cohort of patients.

Skin biopsy specimens from a single lesion per patient were collected and used as the source of the *M. leprae* DNA for genotyping. For the Cebu patient group, the panel of 15 VNTR markers exhibited slow, medium and rapid rates of evolution. Therefore, all loci were informative revealing a heterogeneous diverse *M. leprae* population. In contrast, we observed that in Qiubei, China, the *M. leprae* isolates appeared to have descendants of a dominant genotype with only few of the 15 VNTR loci being polymorphic. Thus not all *M. leprae* VNTR loci are highly variable; markers suitable for genotyping can be exploited, but these need to be determined empirically for each population or endemic region.

Finding conserved VNTR profiles in MCF pairs is consistent with the potential for MLVA in tracing sources of infection or identifying a transmission network. While, MLVA that was able to demonstrate the genotypes were shared for patients with known epidemiological links, the estimated phylogenetic trees contain clusters with no known links, together with those from disparate residential histories. Multiple factors including the long incubation period of leprosy, missing data due to partial patient coverage (capture) in this study (Table 1.4.1), the lack of sensitivity of MLVA in paucibacillary

biopsies, and patient mobility are probable causes. The search for new genomic markers should continue and be evaluated as they are discovered.

Monot et al described a classification of global isolates into four types based on the discovery of three SNPs (14). SNP type 1 *M. leprae* was shown to be prevalent in Asia (Nepal, India, Korea, Thailand, and the Philippines), Pacific islands of New Caledonia and East Africa, while type 3 was found in the European, North African and American continents, and also in New Calendonia. In our study, we also found that the major circulating isolate in Cebu is SNP type 1; 16% belong to SNP type 3. Cebu was colonized by the Spaniards in the 16th century and therefore the finding of SNP type 3 isolates could be attributed to this history. However, SNP type 3 isolates, are also prevalent in Japan, Korea and China (14, 22). Therefore the exact origin of the Cebu strains is not definite. Interestingly, VNTR markers could separate the SNP type 3 isolates from the majority of the SNP type 1. Distinctive allelic patterns were found in SNP type 3 isolates; typically contractions rather than expansions at several microsatellite and minisatellite loci when compared to SNP type 1 isolates. In addition, a two locus signature was detected that is highly predictive of the SNP type 3 within the Cebu population. These data indicate that there is sufficient stability in the VNTR markers such as to retain phylogenetic relationships during transmission. The detection of the polymorphism for the 21-3 minisatellite simply requires a 3% agarose gel system; the detection of 1 or 2 copy alleles being a good indicator of the SNP type 3 in the studied population. The two SNP subtypes were not found to be localized to any part of the island suggesting there has been considerable circulation and dispersion of leprosy in this community.

Intra-patient variability of VNTRs has been raised as a concern in two studies (13, 24). These studies examined VNTR profiles in more than one clinical sample per patient taken at the same and/or different time points. In the study of 42 patients in Mali, Africa, 36 patients showed no allelic variation in all five loci, four of which were microsatellites (13). In 6 patients, one or two copy variations were noted at one or all loci. It was suggested that a larger panel of markers exhibiting a broader range of allelic diversities would be necessary to assess if there were any genotypic relationships within this patient population. SNP typing was unable to further discriminate the isolates as they were all of the type 4.

There are several indications from our findings in Cebu that VNTRs can be used as markers both for the study of strain evolution and leprosy transmission: the distinctive clustering of SNP type 3 isolates by VNTRs; the contrasting allelic properties between the Cebu and Qiubei leprosy patients; and the closely matching profiles in MCFs despite the overall diversity in the population and the analysis of one time point single biological samples.

Regarding the possibility of detection of mixed infections, we found one biopsy sample (M55) from which two distinct alleles were detected at several VNTR loci (11/15) in the electropherograms by the FLA technique, which would have been missed with conventional DNA sequencing. Separate aliquots of frozen sections gave the similar results.

The absolute definition for strain identity and the apriori choice of markers is not yet clear because these will be dependent on the strain diversity in the target population and have to be experimentally described. The mechanisms by which *M. leprae* isolates

are established during infection (and re-infection) and evolve are not known and could be influenced by the endemicity (incidence rate), the population size, host susceptibility, leprosy control measures, and other unknowns such as environmental reservoirs. Furthermore, the methodology and analysis can vary on whether the application is for discerning localized transmission or for delving into broader evolutionary relationships or both as we have discerned in this study using 15 markers in Cebu. The baseline molecular epidemiology database and approaches from this work will aid in designing targeted studies to monitor selected genotypes/clusters, perform active case finding, trace and thus uncover transmission arising in settings beyond the household such as school, employment and socialization or from environmental sources. The collection and banking of samples such as skin biopsies and slit skin smears, and a tiered approach to strain typing such as described could be adopted in other reference clinics.

1.4.6 ACKNOWLEDGEMENTS

These studies were funded by NIH-NIAID grants AI-59644, AI-063457 and contract NO1-AI-25469. We thank staff at CSC, LWM for clinical work, patients who volunteered for the research studies, and Macromolecular Resources, CSU for DNA sequencing and typing services. We thank Michelle Price and Niccole Carner for the graphics.

1.4.7 REFERENCES

1. **Bohannan, B. J., and J. Hughes.** 2003. New approaches to analyzing microbial biodiversity data. Curr. Opin. Microbiol. **6**:282-287.

2. **Cardona-Castro, N., J. Beltrán-Alzate, and R. Manrique-Hernández.** 2008. Survey to identify *Mycobacterium leprae*-infected household contacts of patients from prevalent regions of leprosy in Colombia. Mem Inst Oswaldo Cruz. **103**:332-336.

3. **Cellona, R. V., G. P. Walsh, T. T. Fajardo Jr, R. M. Abalos, E. C. dela Cruz, L. Guido-Villahermosa, M. V. Felicio-Balagon, G. J. Steenbergen, and J. T. Douglas.** 1993. Cross-sectional assessment of ELISA reactivity in leprosy patients, contacts, and normal population using the semisynthetic antigen natural disaccharide octyl bovine serum albumin (ND-O-BSA) in Cebu, The Philippines. Int. J. Lepr. Other Mycobact. Dis. **61**:192-198.

4. **Cole, S.T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honoré, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels, C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell.** 2001. Massive gene decay in the leprosy bacillus. Nature. **409**:1007-1011.

5. **Deps, P. D., B. V. Guedes, J. Bucker Filho, M. K. Andreatta, R. S. Marcari, and L. C. Rodrigues.** 2006. Delay in the diagnosis of leprosy in the Metropolitan Region of Vitória, Brazil. Lepr Rev. **77**:41-47.

6. **Felsenstein, J.** 1989. PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics* **5:** 164-166.

7. **Groathouse, N. A., B. Rivoire, H. Kim, H. Lee, S. N. Cho, P. J. Brennan, and V. D. Vissa.** 2004. Multiple polymorphic loci for molecular typing of strains of *Mycobacterium leprae*. J. Clin. Microbiol. **42**:1666-1672.

8. **Kimura, M., R. M. Sakamuri, N. A. Groathouse, B. L. Rivoire, D. Gingrich, S. Krueger-Koplin, S. N. Cho, P. J. Brennan, and V. Vissa.** 2009. Rapid Variable-Number Tandem Repeat (VNTR) genotyping for *Mycobacterium leprae* clinical specimens. J. Clin. Microbiol. **47**:1757-1766.

9. **Kolaczkowski, B., and J. W. Thornton.** 2004. Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. Nature **431**:980-984.

10. **Kumar, S., K. Tamura, and M. Nei. 2004.** MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform **5**: 150–163.

11. **Martin, A. P.** 2002. Phylogenetic approaches for describing and comparing the diversity of microbial communities. Appl. Environ. Microbiol. **68**:3673-3682.

12. **Matsuoka M., S. Maeda, M. Kai, N. Nakata, G-T. Chae, T. P. Gillis, K. Kobayashi, S. Izumi, and Y. Kashiwabara.** 2000. *Mycobacterium leprae* typing by genomic diversity and global distribution of genotypes. Int. J. Leprosy **68**: 121–128.

13. **Monot, M., N. Honoré, N. Balière, B. Ji, S. Sow, P. J. Brennan, and S. T. Cole.** 2008. Are variable-number tandem repeats appropriate for genotyping *Mycobacterium leprae*? J. Clin. Microbiol. **46**:2291-2297.

14. **Monot, M., N. Honoré, N. Garnier, R. Araoz, J. Y. Coppée, C. Lacroix, S. Sow, J. S. Spencer, R. W. Truman, D. L. Williams, R. Gelber, M. Virmond, B. Flageul, S. N. Cho, B. Ji, A. Paniz-Mondolfi, J. Convit, S. Young, P. E. Fine, V. Rasolofo, P. J. Brennan, and S. T. Cole.** 2005. On the origin of leprosy. Science. **308**:1040-1042.

15. **Saitou, N., and M. Nei.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

16. **Shin Y. C., H. Lee, H. Lee, G. P. Walsh, J. D. Kim, and S. N. Cho.** 2000. Variable numbers of TTC repeats in *Mycobacterium leprae* DNA from leprosy patients and use in strain differentiation. J. Clin. Microbiol. **38**:4535-4538.

17. **Swofford, D. L.** 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

18. **Tadesse Argaw, A., E. J. Shannon, A. Assefa, F. S. Mikru, B. K. Mariam, and J. B. Malone.** 2006. A geospatial risk assessment model for leprosy in Ethiopia based on environmental thermal-hydrological regime analysis. Geospat. Health. **1**:105-113.

19. **Truman, R., A. B. Fontes, A. B. De Miranda, P. Suffys, and T. Gillis.** 2004. Genotypic variation and stability of four variable-number tandem repeats and their suitability for discriminating strains of *Mycobacterium leprae*. J Clin Microbiol. **42**:2558-2565.

20. **van Beers, S. M., M. Hatta, and P. R. Klatser.** 1999. Patient contact is the major determinant in incident leprosy: implications for future control. Int. J. Lepr. Other Mycobact. Dis. **67**:119-128.

21. **Weir, B. S.** 1990. Genetic data analysis: methods for discrete population genetic data analysis. Sinauer Associates, Inc., Sunderland, Mass.

22. **Weng X., Z. Wang, J. Liu, M. Kimura, W. C. Black 4th, P. J. Brennan, H. Li, and V. D. Vissa.** 2007. Identification and distribution of *Mycobacterium leprae* genotypes in a region of high leprosy prevalence in China: a 3-year molecular epidemiological study. J. Clin. Microbiol. **45**:1728-1734.

23. **World Health Organization.** 2008. Global Leprosy situation, beginning of 2008. Wkly Epidemiol. Rec. 2008. **83**(33):293–300.

24. **Young S. K., J. M. Ponnighaus, S. Jain, S. Lucas, S. Suneetha, D. N. Lockwood, D. B. Young, and P. E. Fine.** 2008. Use of Short Tandem Repeat Sequences to Study *Mycobacterium leprae* in Leprosy Patients in Malawi and India. PLoS Negl. Trop. Dis. **2**:e214.

25. **Young S. K., G. M. Taylor, S. Jain, L. M. Suneetha, S. Suneetha, D. N. Lockwood, and D. B. Young.** 2004. Microsatellite mapping of *Mycobacterium leprae* populations in infected humans. J. Clin. Microbiol. **42**:4931-4936.

26. **Zhang L., T. Budiawan, and M. Matsuoka.** 2005. Diversity of potential short tandem repeats in *Mycobacterium leprae* and application for molecular typing. J. Clin. Microbiol. **43**:5221-5229.

Chapter 1.5

A CONTINUATION: STUDY AND CHARACTERIZATION OF *M. LEPRAE* SHORT TANDEM REPEAT GENOTYPES AND TRANSMISSION OF LEPROSY IN CEBU, PHILIPPINES

The work in this chapter has been submitted to *Leprosy Review* journal Sakamuri et al, 2009 (in press)

Clinical samples were obtained from LWM, Cebu, Philippines in collaboration with Dr. Marivic Balagon who is the current director for LWM. These studies were initiated by our former lab member Jordan Harrison who extracted DNA from clinical samples. I contributed towards the completion of MLVA analysis, data analysis and preparation of manuscript.

Sakamuri RM, Harrison J, Gelber R, Saunderson P, Brennan PJ, Balagon M, and Vissa V. 2009. A continuation: Study and characterization of *M. leprae* short tandem repeat genotypes and transmission of leprosy in Cebu, Philippines. Lep. Rev. (in press)

1.5.1 ABSTRACT

Objective To study the stability and allelic diversity of tandem repeat loci in *M. leprae* in leprosy patients of Cebu, Philippines, and the suitability of multilocus variable number of tandem repeat (VNTR) analysis (MLVA) typing for detecting transmission.

Methods Total DNA from seventy newly diagnosed leprosy patients consulting at the Leonard Wood memorial, Cebu Skin Clinic Total DNA was extracted from slit skin smear (SSS) scrapings of each patient and used for amplification of thirteen *M. leprae* VNTR loci by single locus or multiplex PCR. Number of repeats for each VNTR locus was obtained by DNA sequencing or fragment length analysis methods. Medical, social and geographic details were included in the molecular epidemiology database.

Results and conclusions Multiplex PCR and fragment length analysis methods were found to be more efficient and accurate compared to single short tandem repeat (STR) amplification and DNA sequencing. Intra-patient MLVA patterns from four different samples were conserved in the minisatellites, while differences in one or more of the polymorphic and stutter prone microsatellites was observed, in four of five patients. The thirteen loci could differentiate *M. leprae* strains in Cebu. However, MLVA patterns were stable enough during incubation and transmission between individuals within multi-case families. Thus *M. leprae* MLVA has potential for strain typing and transmission studies in Cebu.

1.5.2 INTRODUCTION

A micro-epidemiology project under the auspices of the IDEAL consortium has been active in several countries, including the Philippines, which has a national case detection rate (CDR) of 2.6 per 100,000. The Philippines consists of around 7000 islands and Cebu, one of the larger islands, has historically had a higher incidence of leprosy than most others. The leprosy CDR for Cebu is now 5.4 per 100,000 although there is considerable variation between districts; in Lapu Lapu (an urban area adjacent to Cebu City) the CDR was 16.6 per 100,000 in 1997 and 10.3 per 100,000 in 2007. In 1930, the Leonard Wood Memorial research facility was established in Cebu and the epidemiology of leprosy was intensively studied in the district of Cordova near to Cebu City (3).

Because it is an island with a well-functioning health infrastructure, it was felt that Cebu would be an ideal setting in which to study the micro-epidemiology of leprosy. The movements of patients and their inter-relationships should be easier to map.

1.5.3 MATERIALS AND METHODS

1.5.3.1 Clinical Samples

Seventy slit skin smears (SSS) and biopsy samples were collected in 70% ethanol from the patients attending Leonard Wood Memorial (LWM), Skin Clinic in Cebu. These samples were shipped to Colorado State University (CSU) in two separate batches.

Besides obtaining basic details of the medical history including the leprosy condition and treatment regimen, information about the patients' geographical area of residence (village, city and province), types /place(s) of employment and education,

nature of dwelling, sources of drinking, cooking, and bathing water spanning a prior 20 year period, was captured with a standardized questionnaire at the time of recruitment. The average bacteriological index was 3.5. Nineteen patients reported knowledge of at least one known leprosy patient in their families. Two Multi-case families (MCFs) were identified in this IDEAL study patient population. In some patients, several specimens were obtained to identify any intra-patient variations which may be present.

The SSS samples were processed to obtain the DNA. Briefly, the SSS containing tubes were centrifuged to pellet the particulate material. The pellet was rehydrated in 0.5 ml PBS buffer (Invitrogen) at room temperature for 1 hour, followed by a centrifugation step to recover the SSS material. Buffer (0.1M Tris, 0.05% Tween 20) was prepared, filter sterilized and distributed into 3-5 ml aliquots. Proteinase K powder was added to the buffer (10 mg/ml) just before its use and filter sterilized. This lysis buffer (200 µl) was added into the rehydrated SSS sample and allowed to incubate overnight at 60 °C. Then the lysate was centrifuged at 13,000 rpm for 10 minutes, the supernatant was divided into two aliquots. One aliquot was stored at -20 °C and the other tube was stored at 4 °C for further analysis. DNA from the biopsy samples were obtained by using DNeasy Tissue kit (Qiagen).

1.5.3.2 VNTR mapping

Single short tandem repeat (STR) amplification was done according to the IDEAL protocol (1) and the Multiplex PCR (M-PCR) amplification was done according to a protocol developed at CSU (4). M-PCR was performed to amplify four or five VNTR loci simultaneously from the clinical isolates, by utilizing the primer sequences described by us or others, in four combinations (2, 4, 6). Primer sequences used for M-PCR

combinations are from Kimura et al. (4). Ten microsatellites [(AC)8a, (AC)8b, (AC)9, (AT)15, (AT)17, (TA)10, (TA)18, (GGT)5, (GTA)9, (TTC)21] and seven minisatellites [6-7, 12-5, 18-8, 21-3, 23-3, 27-5, and rpoT] were included in this study. Minisatellite copy number was inferred by 2.5-3% agarose gel electrophoresis using the EZ load^{TM} 20 bp Molecular Ruler (Bio-Rad) for product length determination. The copy number (allele) for microsatellite loci was determined by direct sequencing of PCR products. Sequence was obtained for representative products at the Macromolecular Resource Facility (MRF) (now known as Proteomics and Metabolomics Facility) at CSU.

M-PCR products were subjected to fragment length analysis (FLA) for determining the allele number for each VNTR locus (4). After the M-PCR, the diluted PCR product was combined with deionized formamide (Applied Biosystems), and LIZ-500 DNA standard (Applied Biosystems). The sample was denatured and subjected to capillary electrophoresis on the Applied Biosystems Genetic Analyzer 3130 at MRF/PMF at CSU. Following the separation, the electropherograms were visualized and analyzed using GeneMapper Version 3.7 software (Applied Biosystems) to determine the major allele for each VNTR locus, in each Multiplex-PCR combination (4).

1.5.4 RESULTS AND DISCUSSION

DNA from sixty SSS and some of their paired biopsy samples were extracted. Single PCR for 13 VNTR loci were done for some of the samples. Allele numbers for the minisatellite loci like 12-5, 21-3, 23-3 and 27-5 were obtained from 3% agarose gel. However, the sequence data obtained from the PCR products of the microsatellite loci

were hard to analyze to determine the exact allele number. Therefore, M-PCR and FLA methods were used for obtaining the VNTR profiles for all samples (Table 1.5.1).

In order to test the reproducibility of the VNTR data, nine SSS DNA samples were also sent to another laboratory. Thirteen VNTR loci data for the nine SSS samples were obtained by using single PCR method and conventional sequencing at a partner laboratory, whereas at CSU M- PCR and FLA methods were used. When the VNTR data from both the laboratories were compared 14 mismatches were observed. When the sequence files from were re-read for 12 of these mismatches, it was found that 7 of the sequence data was hard to analyze or 5 of them were misinterpreted due to mixed signals, stutter or unidentifiable end points of repeats in the sequence. The FLA results were therefore accepted. In general, M-PCR followed by FLA was found to be easier to interpret for allele numbers for the VNTR loci and the data obtained from this approach may be more accurate and also useful to detect mixed alleles.

In order to study the intra-patient *M. leprae* VNTR profiles, SSS from different anatomical sites were collected. A biopsy specimen was also obtained from one of the lesions. These SSS and biopsy samples were subjected to M-PCR and FLA. It was found that the patterns were mostly similar in the minisatellites, while variations were observed in the more variable and stutter prone microsatellites like (AT)17 and (TTC)21 (Table 1.5.2).

The VNTR profiles of the isolates from the SSS samples within the two MCFs were compared. IDEAL50 and IDEAL53 of MCF-1 were found to have identical VNTR profiles, indicating that the infections were from the same isolate. In another MCF (IDEAL68, IDEAL69 and IDEAL70) the isolates in the patients IDEAL69 and IDEAL

Table 1.5.1 *M. leprae* VNTR profiles in leprosy patients in the Cebu study population

Sample ID	(AC)8b	(GTA)9	(GGT)5	(AT)17	21-3	(AC)9	(AT)15	(AC)8a	27-5	6-7	(TA)18	(TTC)21	18-8	12-5	23-3	(TA)10	rpoT	BI	Group (5)
IDEAL6	8	9	5	14	3	9	15	9	5	7	20?	24	8	5	2	12	3	2.0+	B1
IDEAL7	9	9	5	15	2	8	17	9	5	6	14	36	8	4	2	11	3	3.8+	B1
IDEAL8	7	9	5	17	3	9	16	10	5	7	18	21	8	5	2	11	3	2.33+	B1
IDEAL9	6	9	5	10	2	8	17	9	5	7	15	16	7	5	2	7	3	3.66+	C
IDEAL10	7	9	6	14	3	8	20	10	5	7	24	23	8	4	2	11	3	3.0+	D
IDEAL11	8	9	5	13	3	9	15	9	5	8	21	19	8	4	2	11	3	3.17+	B1
IDEAL12	7	10+12	4	14	3+1	9	14	9	5	7	17	20	7/8	4	2	11	3	2.8+	C
IDEAL13	7	10	4	14	1	9	14	9	5	7	17	20	7	4	2	11	3	4.5+	C
IDEAL14	8	9	5	15	3	9	14	9	5	7	19	25	8	5	2	11	3	5.0+	B1
IDEAL15	8	10	5	15	3	9		8	5	7	21	26	8	4	2	11		4.7+	B1
IDEAL16	8	9	5	14	3	8	17	9	5	7	21	25	8	4	2		3	1.2+	B1
IDEAL17	8	8	5	13	3	8	14	10	5	6	22	27	8	4	2	11	3	3.33+	
IDEAL18	8	11	5	13	3	10	18	8	5	7	14	31	8	4	2	10		3.5+	A2
IDEAL19	8	11	4	15	2	8	19	8	5	6	13	9	8	4	2	8	3	2.2+	A1
IDEAL20	7	12	5	13	3	8	22	9	5	7	21	31	8	4	2	11	3	4.0+	A1
IDEAL21	8	9	5	11	3	9	13	9	5	8	20	18	8	4	2	11		4.66+	B1
IDEAL22	8	11	5	15	3	10	16	9	5	7	15		8	4	2	10	3	4.5+	A2
IDEAL23	8	9	5	20	3	9	15	8	4	8	25	19	8	4	2	10	3	5.33+	B1
IDEAL24	8	9	5	13	3	9	11	8	5	7	14	16	8	4	2	9	3	5.0+	B1
IDEAL25	8	9	4/5	15	3	9	15	9	5	7	16/19	24	8	5	2	13	3	5.5+	B1
IDEAL26	8	11	5	12	3	9	16	9	5	7	19	21	8	4	2	9	3	3.8+	A2
IDEAL27	8	11	5	15	3	8	14	9	5	6	14	22	8	4	2	13	3	5.0+	A1
IDEAL28	8	12	5	13	3	10		9	5	7	23	28	8	4	2	11		1.7+	A2
IDEAL29	8	13	5	13	3	9		9	5	7	15	31	8	4	2	10		4.8+	A2
IDEAL30	8		5	13	3	8		9	5	7	15	19			2	10	3	1.0+	
IDEAL31	8	9	5	15	3	9	13	10	5	7	23	19		4	2	12	3	2.5+	B1
IDEAL32		12	4	13	3	8	17	9		7	15	12		4	2		3	0.33+	A1
IDEAL33	8		6	18	2/3	8		9	4	7	15		8	4	2		3	0.17+	
IDEAL34	7	9	6	18	3	9	23	10	5	7	11	22	8	4	2	16		2.5+	D
IDEAL35	8	10	5	14	3	9	15	9	5	7	18	25	8	5	2	12	3	4.33+	B1
IDEAL36	8	12	5	13	3	8/10	17	8	5	5/7	16	23	8	4	2	10	3	5.0+	A1
IDEAL37	8	11	5	14	3	10	16	8	5	7	14	30/33	8	4	2	10	3	4.17+	A2
IDEAL38	8	9	5	21	3	9	15	8		7	15	27	8	6	2	11	3	4.7+	B1
IDEAL39	8	13	5	13	3	9	15	9	5	7	16	31	8	4	2	10	3	4.5+	A2
IDEAL40	8	11	5	14	3	9	15	8	5	7	17	26	8	4	2	10	3	4.2+	A2
IDEAL41	7	9	6	17	3	9	22	10/9	5	6	17	23	8	4	2	13	3	4.2+	D
IDEAL42					2	8	15	9	6			32	8	5	2		3	0.33+	
IDEAL43	8	11	5	14/10	3	8		10	5	7	17	26	8	4	2	10		0.17+	A1
IDEAL44	8	12	5	14	3	9	16	11	5	7	17	23	8	4	2	14		4.66+	A2
IDEAL45	8	13	4	15	3	9	16	9	5	7	21	27/25	8	4	2	11	3	4.5+	A2
IDEAL46	6	12	4	12	2	8	16	9	5	6	14	23	8	4	2	7	3	4.0+	C
IDEAL47	7	9	6	17	3	9	15	9	5	7		26	8	4/5	2		3	1.0+	D
IDEAL48	8	9	4/5	15/16	3	9	15	9	5	7	17	25	8	5	2	11	3	5.0+	B1

IDEAL49	8	9	5	12	2	10	14	9	5	7	15		8/7	4	2	10	3	4.2+	B1
IDEAL50	8	10	5	14	3	8	15	8	5	5	18	24	8	4	2	10		5.0+	B1
IDEAL51	7	8	4	17	1	9	14	9	5	8	14	18	7	4	2	10	3	5.0+	C
IDEAL52	6	9	5	10	2	8	18	9	5	6	13	14	8	5	2	7	3	5.0+	C
IDEAL53	8	10	5	14	3	8	15	8	5	5	16	24/25	8	4	2	10	3	3.5+	B1
IDEAL54	8	12	5	14	3	8	18	9	5	7	23	18	7	4	2	11	3	3.17+	A1
IDEAL55	8	11	5	17	3	9	14	10	5	8	18	17	8	4	2	11	3	3.2+	A2
IDEAL56	7		6	12	2/3	9	14	11	5	7		18	8	4	2		3	1.2+	D
IDEAL57	8	12	5	13	3	9	15	9	5	7	19	27	6	4	2	10	3	4.0+	A2
IDEAL58	8	9	5	15	3	9	11	8	5	8	16	17		4	2	9	3	4.2+	B1
IDEAL59	8	9	5	18	3	9	15	8	5	7	21	24	4	5	2	12	3	5.0+	B2
IDEAL60																		0.17+	
IDEAL61	8	9	5	15	3	9		9	5	7	17	18	8	5	3	11	3	3.7+	T
IDEAL62	8	9	5	15	3	9	15	9	5	7	18	25	8	5	2	12	3	4.0+	B1
IDEAL63	8	10	5	24	3	9	16	9	5	7	16	22	8	5	2	14	3	2.7+	B1
IDEAL64	8	10	5	17	3/2	10	13	9	5	7	18	25	8	4	2	12	3	4.2+	B1
IDEAL65	7	9/10	6		3	9	16	10	5	7	15	23	8	4	2	12	3	1.5+	D
IDEAL66	8		5	14	3	9	13	9	5	8	15	19	8	4	2	11	3	2.7+	
IDEAL67	8	11	5	15	3	9	16	8	5	6		27	8	4	2	11	3	5.0+	A2
IDEAL68	7	9	6	15	3	9	14	9	5	7	17	26	8	4	2	13	3	5.0+	D
IDEAL69	9	9	5	15	3	9	14	9	5	7	17	23	8	5	2	13	3	4.8+	B1
IDEAL70	9	9	5	15	3	9	14	9	5	7	17	25	8	5	2	13	3	4.2+	B1

Weal PCR product NO PCR product

+ indicates two alleles for the VNTR locus; / indicates either of the alleles, small font number indicates a minor allele of the locus

Table 1.5.1 *continuation*

Sample ID	Sample type	STR locus													Sample collection sites
		(AC)8a	(AC)8b	(AC)9	(TA)10	(AT)17	(GGT)5	(GTA)9	(GAA)21	6-7	12-5	21-3	23-3	27-5	
IDEAL1A	SSS	9	8	9	10	13	5	12	28	7	4	3	2		Ear lobe right
IDEAL1B	SSS	9	8	9	10	13	5	12	29	7	4	3	2	5	Elbow right
IDEAL1C*	SSS	9	8	9	10	13	5	12	29	7	4	3	2	5	Elbow left
IDEALB01	Biopsy	9	8	9	10	13	5	12	29	7	4	3	2	5	Elbow left
IDEAL2A	SSS	10	8	8	11	13	5	11	30+34	7	4	3	2	5	Ear lobe right
IDEAL2B*	SSS	10	8	8	11	13	5	11	31	7	4	3	2	5	Back right A
IDEAL2C*	SSS	10	8	8	11	13	5	11	30	7	4	3	2	5	Back right C
IDEALB02	Biopsy	10	8	8	11	13	5	11	31	7	4	3	2	5	Back right
IDEAL3A	SSS	9	NA	8	NA	10	NA	9	12	NA	4	3	NA	5	Ear lobe left
IDEAL3B	SSS	10	6	8	7	10	5	9	15	7	5	2	2	5	Arm left posterior
IDEAL3C*	SSS	10	6	8	NA	10	5	10	14	7	5	2	2	5	Back left
IDEALB03	Biopsy	10	6	8	7	10	5	10	14	7	5	2	2	5	Back left
IDEAL4A	SSS	10	7	9	NA	18	6	9	24	7	4	3	2	5	Ear lobe right
IDEAL4B*	SSS	10	7	9	NA	19	6	9	23	7	4	3	2	5	Back right
IDEAL4C	SSS	10	7	9	12	19	6	9	23	7	4	3	2	5	Back left
IDEALB04	Biopsy	10	7	9	12	19	6	9	23	7	4	3	2	5	Back right
IDEAL5A	SSS	10	6	8	7	9	5	9	14	7	5	2	2	5	Ear lobe left
IDEAL5B	SSS	10	6	8	7	9	5	9	14	7	5	2	2	5	Arm left posterior A
IDEAL5C*	SSS	10	6	8	7	9	5	9	14	7	5	2	2	5	Arm left posterior B
IDEALB05	Biopsy	10	6	8	7	9	5	9	14	7	5	2	2	5	Arm left posterior B

Weak product

* SSS is from the biopsy site, NA: No PCR product

Table 1.5.2 Comparison of VNTRs from different body sites; Biopsy (B) and Slit Skin Smears (A, B, C)

70 were found to be closely matching, but different from that of IDEAL68 (Table 1.5.3) at 4/13 VNTR loci. When the epidemiological information for the second MCF was considered it was found that this family has a history of 5 other leprosy patients (2 uncles, 2 sisters and a brother). Therefore, multiple sources of infection may be possible in this family context.

MLVA differentiated *M. leprae* in this study of 70 leprosy patients consulting at the Cebu Skin Clinic; adding to an existing strain type database of more than 200 Cebu patients. Although a formal phylogenetic or cluster analysis of the strain types in the 70 patients is not shown here, the alleles, their frequencies and diversity, and the VNTR patterns are consistent with earlier findings. The strain types can be assigned to five major groups (5). The two multi-case families in this study (as in others identified in the earlier study (5)) are of interest in that two individuals in each family had identical strains of *M. leprae*, indicating that the VNTR profile can remain stable for at least as long as the incubation period of these cases. This is also suggested by the finding that the VNTR profile is stable within one individual. This suggests that if the right specimens can be collected, evidence of direct transmission between individuals could be obtained.

1.5.5 ACKNOWLEDGEMENTS

These studies were funded by HEISER program for research in Leprosy and Tuberculosis of The New York Community Trust as apart of the IDEAL (Initiative for Diagnostic and Epidemiological Assays for Leprosy) consortium, NIH-NIAID grants AI-063457 and Contract NO1-AI-25469. We thank the clinical and administrative staff at Cebu Skin Clinic, LWM for their support and contributions to make this project possible.

Family Number	Patient ID	STR locus													Barangay	City
		(AC)8a	(AC)8b	(AC)9	(TA)10	(AT)17	(GGT)5	(GTA)9	(GAA)21	6-7	12-5	21-3	23-3	27-5		
MCF-1	IDEAL50	8	8	8	10	14	5	10	24	5	4	3	2	5	Kinalumsan	Lapu-lapu
	IDEAL53	8	8	8	10	14	5	10	24/25	5	4	3	2	5	Kinalumsan	Lapu-lapu
MCF-2	IDEAL68	9	7	9	13	15	6	9	26	7	4	3	2	5	Upper Gahab	Sibonga
	IDEAL69	9	9	9	13	15	5	9	23	7	5	3	2	5	Upper Gahab	Sibonga
	IDEAL70	9	9	9	13	15	5	9	25	7	5	3	2	5	Upper Gahab	Sibonga
	Alleles Differences when compared with *M. leprae* found in siblings in the multi-case family															

Table 1.5.3 *M. leprae* MLVA profiles of two multi-case families

1.5.6 REFERENCES

1. **Gillis T., V. Vissa, M. Matsuoaka, S. Young, J. H. Richardus, R. Truman, B. Hall, P. J. Brennan, and IDEAL Consortium Partners**. 2009. Characterization of Short tandem repeats for genotyping *Mycobacterium leprae.* Lep Rev. (in press)

2. **Groathouse, N. A., B. Rivoire, H. Kim, H. Lee, S. N. Cho, P. J. Brennan, and V. D. Vissa.** 2004. Multiple polymorphic loci for molecular typing of strains of *Mycobacterium leprae.* J Clin Microbiol **42:**1666-72.

3. **Guinto R. S., J. A. Doull, H. Bancroft, J. N. Rodriguez**. 1951. A field study of leprosy in Cordova, Philippines; resurvey in 1941 after eight years. Int J Lepr. **19**:117-135.

4. **Kimura M., R. M. Sakamuri, N. A. Groathouse, B. L. Rivoire, D. Gingrich, S. Krueger-Koplin, S. N. Cho, P. J. Brennan, and V. Vissa**. 2009. Rapid variable-number tandem-repeat genotyping for *Mycobacterium leprae* clinical specimens. J Clin Microbiol. **47**:1757-1766.

5. **Sakamuri R. M., M. Kimura, W. Li, H. C. Kim, H. Lee, M. D. Kiran, W. C. Black 4th, M. Balagon, R. Gelber, S. N. Cho, P. J. Brennan, and V. Vissa**. 2009. Population based molecular epidemiology of leprosy in Cebu, Philippines. J Clin Microbiol. **47**:2844-2854.

6. **Zhang L., T. Budiawan, and M. Matsuoka**. 2005. Diversity of potential short tandem repeats in *Mycobacterium leprae* and application for molecular typing. J Clin Microbiol **43**:5221-5229.

Chapter 1.6

Evaluation of primary drug resistance mutations in *M. leprae* isolates from newly diagnosed patients from Cebu, Philippines

1.6.1 ABSTRACT

Drug resistance is a cause for concern in the treatment and control of any infectious disease. Drug resistance in leprosy was known to be caused during the dapsone monotherapy control program and patients' irregular treatment. WHO MDT which includes anti-leprosy drugs such as dapsone, rifampicin and clofazimine was introduced in 1982 for the treatment and control of leprosy. However drug resistance was reported in even in the areas where the MDT was implemented successfully and there is no estimate of drug resistance of leprosy. So to evaluate the primary drug resistance in newly diagnosed leprosy patients who are being treated with the standardized MDT regimen, we developed a multiplex PCR for amplifying the drug resistance determining regions (DRDR) in genes *rpoB, folP1*, *gyrA* and *gyrB*, which are targets of the drugs rifampicin, dapsone and oflaxacin, respectively. No mutations were detected in *rpoB* and *folP1* DRDRs for 77 samples and in the *gyrA* DRDR for 15 samples tested thus far.

1.6.2 INTRODUCTION

1.6.2.1 Anti-leprosy drugs

Even though *Mycobacterium leprae* has been identified from leprosy patients as early as 1873, preliminary screening for potential anti-leprosy drugs was not possible in the laboratory, since the bacilli are un-culturable in artificial media (*in vitro* conditions). So the chemotherapy for leprosy was started mainly based on the compounds' efficacy against *M. tuberculosis* and other bacteria which were then tested in leprosy patients. Promin (Glucosulfone sodium) was used for leprosy by Faget (9), which was shown to possess anti-tubercular activity in guinea pigs (10). However this had to be administered

intravenously (IV), so the search for an oral sulphone lead to the discovery of dapsone (DDS: 4,4'- diamino-diphenyl sulphone) (6). Initially dapsone was used at high doses (600-700mg/week) which was found to be toxic and unnecessary, and then the dapsone dosage was reduced. Clofazimine had been developed as an anti-tubercular drug but became popular in leprosy chemotherapy. Barry et al. (1) synthesized this clofazimine which is a riminophenazine dye (Lamprene-Geigy; B663). When tested against *M. leprae*, it was found that it was not only as effective as dapsone but it also had anti-inflammatory action which was an added benefit for suppressing leprosy reactions. However, when exposed to sunlight skin becomes red at first, later brown while the lesions get darkened and turn black or grey. However these colors slowly disappear after stopping the treatment. Red tint may appear in the urine, sputum and sweat. In case of *M. leprae* clofazimine was found to be bacteriostatic whereas dapsone is bacteriocidal. Another drug which was adopted from tuberculosis chemotherapy is rifampicin. Rifampicin is a semisynthetic drug, a derivative of 4-methyl-piperazinaminyl group and also obtained from rifamycin B which is derived from a bacterium (*Amycolatopsis rifamycinica)* (39). It was shown to be bactericidal in both *in vitro* and clinical conditions against *M. tuberculosis* (32, 48). When tested in the mouse foot-pad system it was found to be highly active against *M. leprae* (14, 35). Initially dapsone was commonly used as a monotherapy for leprosy patients from 1940s. Because of the increase in dapsone resistance in leprosy patients, in 1982, WHO recommended multidrug therapy (MDT) which includes dapsone, rifampicin and clofazimine for the treatment and control of multibacillary leprosy (52). The current regimen for patients with paucibacillary (PB) leprosy is, DDS plus rifampicin for 6 months and three drugs DDS, rifampicin and

clofazimine for multibacillary (MB) leprosy (Table 1.1.1). Over the years MDT was found to be very effective for the treatment of leprosy however in some cases resistance to drugs was observed. The MDT regimen has been changed from time to time in the duration of treatment. Meanwhile, minocycline (tetracycline derivative), clarithromycin (macrolide), several fluoroquinolone derivatives like oflaxacin are the drugs that were added subsequently into the list of anti-leprosy drugs.

1.6.2.2 Mode of action for anti-leprosy drugs and their mechanism of drug resistance

Similar to that of sulfonamides, dapsone is mainly involved in the inhibition of folic acid synthesis, particularly in the conversion of 7,8-dihydro-6-hydroxymethylpterin pyrophosphate and para-aminobenzoate (PABA) to 7,8-dihydropteroate with the help of a bacterial enzyme dihydropteroate synthase (DHPS) which is encoded as *folp1* gene in *M. leprae*. This dihydropteroate is converted into tetrahydropteroate which acts as a cofactor in the biosynthesis of purines, pyramidines and aminoacids. Dapsone is a structural analog to PABA and competes for the active site of the DHPS enzyme. The mechanism of dapsone resistance is found to be associated with the development of mutations in the *folp1* (DHPS) gene of *M. leprae* (12, 19, 21, 51). Mutations in the codons 53 and 55 of the *folp1* gene were observed in clinical isolates with dapsone resistance (12, 18, 24, 28, 51) (Fig. 1.6.1).

Rifampicin targets the beta subunit of the RNA polymerase which is encoded by *rpoB* gene, thus preventing the transcription within the bacteria. It was found that the rifampicin-RNA polymerase is a highly stable complex, where the aromatic amino acids

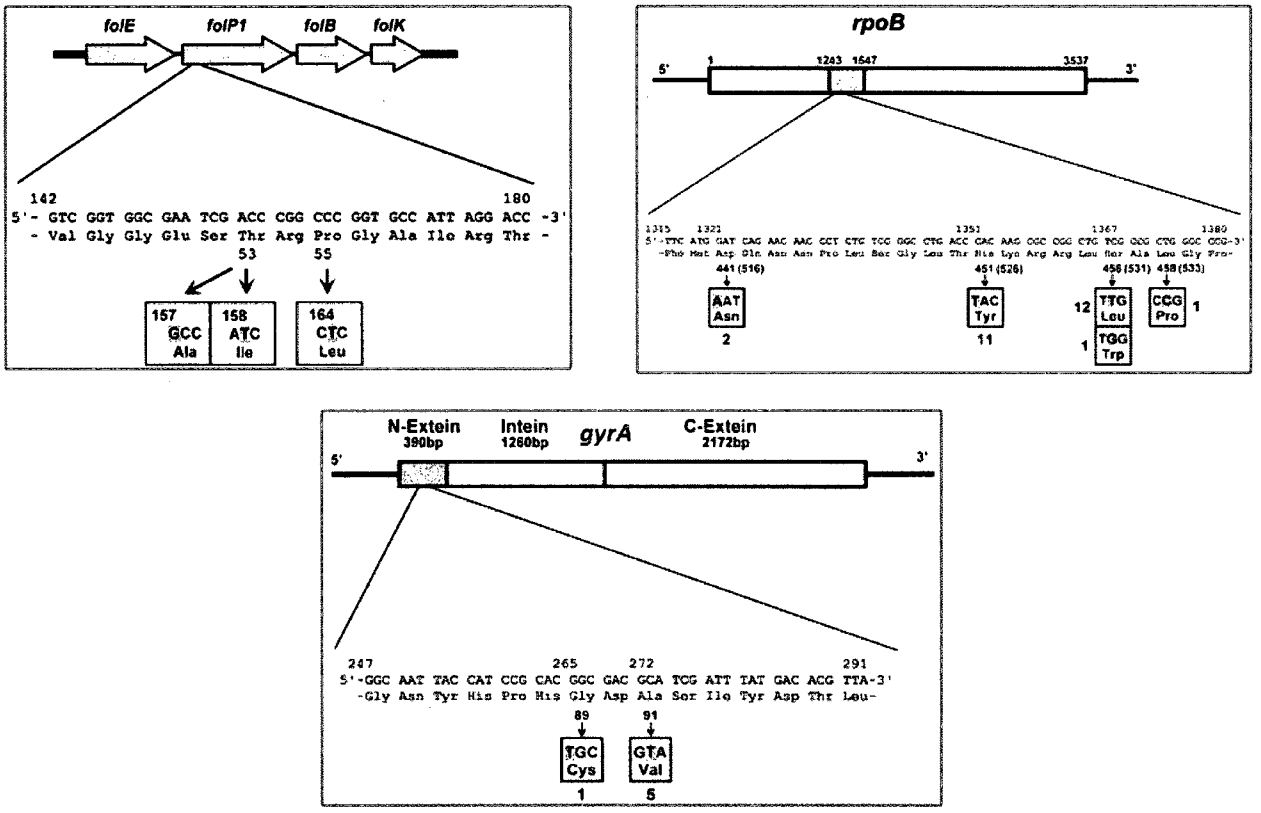


Figure 1.6.1 The DNA sequence and common mutations in the drug resistance determining regions (DRDR) of *M. leprae folP1*, *rpoB*, and *gyrA* genes conferring resistance to dapsone, rifampin, and fluoroquinolones.

of the RNA polymerase are tightly bound to the naphthoquinone of rifampcin (2). The drug resistance for rifampicin in the bacteria is developed when the aromatic amino acids of the RNA polymerase in the active binding area (rifampicin-RNA polymerase complex) are replaced with non aromatic amino acids. Similar to that of *Escherichia coli*, mutations in the *rpoB* gene of *M. leprae* are clustered in a small region (Cluster I, region II) in the middle of the gene (15). So, the mutations at the codons 407, 410, 420, 425 and insertions between 408 and 409 in the *rpoB* gene have been indentified to be associated with rifampicin resistance (Fig. 1.6.1).

The compound B663 (clofazimine) was found be involved somewhere in respiration (1). Interestingly it was found that peroxide formation inside the bacteria may be helpful for the action of this B663 compound, since it was found that catalase negative (isoniazid resistant) strains of tuberculosis were more susceptible to B663 than that of parent or catalase positive strains. Rensburg et al. showed that clofazimine mainly effects membrane phospholipid metabolism, but phospholipase activity is unaffected by clofazimine and its analog B669 (47). In another study, it was found that the drug binds preferentially to mycobacterial DNA at base sequences containing guanine resulting in inhibition of mycobacterial replication and growth (29, 30). In the recent studies in *Staphylococcus aureus* (31), it was found that clofazimine has no specific inhibition of RNA polymerase, but non-specifically inhibits DNA, RNA and protein synthesis. It was mainly found to be a membrane disrupting agent. However the exact mechanism of action of this drug is not yet understood, so the molecular surveillance of drug resistance is currently out of reach.

Regarding the new drugs for leprosy treatment, fluoroquinolones mainly target the DNA metabolism by the inhibition of DNA topoisomerase II (DNA gyrase) and DNA topoisomerase IV. In Gram negative bacteria, fluoroquinolones were predominantly found to inhibit DNA gyrase, whereas in Gram-positive organisms such as *S. aureus*, topoisomerase IV was recently found to be the principal target. The function of DNA gyrase is to supercoil linear DNA double helix to obtain highly condensed three dimensional structure of the DNA. Topoisomerase IV was found to have two functions: the first one being the decatenation step where in the new replicated DNA strands of the daughter chains after chromosome replication are separated and the second being to relax the positive supercoils like that of DNA gyrase in Gram negative bacteria. The DNA gyrase and the topoisomerase IV are found to have two subunits namely *gyrA* and *gyrB* that form a heteroduplex ($gyrA_2gyrB_2$) (5). It was found that mutations in the codons 89 and 91 of the *gyrA* and in the codon 205 of the *gyrB* are associated with quinolone resistance in *M. leprae* (25) (Fig. 1.6.1).

Until recently drug resistance in the *M. leprae* isolates were mainly detected in the mouse foot-pad (MFP) assay.

1.6.2.3 Drug susceptibility tests

Mouse foot-pad technique: Leprosy research gained momentum only after the successful growth and transmission of *M. leprae* in mouse foot-pads by Shepard in 1960 (33, 40, 41, 45). Anti-leprosy drugs were then tested for inhibition of growth of *M. leprae* in mouse foot-pads (34). Initially the drugs were tested from the day of inoculation to test the ability of a drug to suppress the growth of the *M. leprae* inoculum (33). Later these tests

are extended to established infections in the mouse foot-pad (44). In 1967, Shepard developed a kinetic technique for detecting the bactericidal activity of a drug. In this approach mice were inoculated with 5000-10000 bacilli. The drug is given to mice from 60 days to 150days after the inoculation of bacilli. Number of bacilli in the mouse foot pads was measured at regular intervals from 60 days to 150 days and also after 150 days. If the growth of *M. leprae* is less in drug treated mice compared to that of control untreated mice and this effect (limited growth in the drug treated mice) is restricted to the time of drug administration i.e. 150 days, then the drug is found to have bacteriostatic effect. When the growth of *M. leprae* continued to be less in the drug treated mice compared to that of control mice even after 150days, then the drug is found to have bactericidal effect (42). The other method developed by Colston et al, for leprosy drug testing in mouse foot-pads is 'proportional bactericidal test' (7). In this method mouse foot-pads are inoculated with 10^1, 10^2, 10^3 and 10^4 *M. leprae* bacilli, mice are treated with the drugs for the initial 60 days, then the foot-pads are harvested after one year and the number of bacilli is counted (11). Finally the percentage of the bacterial killing is analyzed by a Spearman and Karber method (43). However proportional bactericidal test method is not as popular as Shepard's Kinetic method.

In the classical and continuous method for evaluating the anti-leprosy drugs in mouse foot-pads for the clinical samples homogenates of the biopsy specimens from leprosy patients are used as inoculum. Inocula containing 5000-10000 bacilli are injected into hind mouse foot-pads. Then all the drugs to be tested are given to the mice through diet from day zero until the harvest (6-12 months). Different concentration of each drug is tested in the mouse foot-pad assays. In general, dapsone is tested in three concentrations

(0.0001%, 0.001% and 0.01%/ per gm in diet), clofazimine in two concentrations (0.001% and 0.01% per gm in diet), rifampcin in two concentrations (0.003% and 0.03% per gm in diet) for drug susceptibility testing. The number of bacilli in the control and drug treated mice are measured and compared after harvesting the mouse foot-pads. A susceptible *M. leprae* isolate is one with a significant fold increase (10^5) of bacterial numbers in the control mice compared to that of drug treated mice whereas in a resistant *M. leprae* isolate, the bacterial numbers are significantly higher in the drug treated mice compared to that of control mice for a drug (22, 41).

Molecular methods: As described in section 1.6.2.2, the mutations in the target genes like *folp1*, *rpoB*, *gyrA* and *gyrB* are tested for detection and monitoring of drug resistance for dapsone, rifampicin and oflaxicin respectively in the leprosy clinical isolates. Amplification of target region by PCR and sequencing of the PCR amplicon is most common and direct method of detecting mutations (15, 19, 24, 28, 49-51). Several alternative methods like single-strand confirmation polymorphism (SSCP), hybridization and microarray techniques were also developed for rapid detection of mutations in the target regions (drug resistance determining region, DRDR) from the clinical specimens (16, 17, 26, 36, 37).

It was found that in most of the cases, drug resistance detected through the mouse foot-pad method correlates with the molecular method which shows a mutation in the DRDR for that drug (3). In dapsone, the degree of resistance is found to be variable. It was reported that all of the *M. leprae* isolates with high degree (0.01% of dapsone concentration) of dapsone resistance in the mouse foot-pad method had mutation in the DRDR whereas few isolates with intermediate degree of resistance had mutations, while

none of the low degree of resistance had any mutations in the DRDR of *folp1* gene (4, 27, 38). Results of molecular method and *in vivo* drug testing (mouse foot-pads) for rifampicin and oflaxicin resistance were found to be mostly concordant with very few exceptions in case of rifampicin (27, 46).

1.6.2.4 Rationale and Objective:

Dapsone monotherapy during 1940-80s caused drug resistance within *M. leprae.* Even though MDT treatment is highly successful in controlling the disease, Linder et al. (23) questioned about the MDT duration, since all the patients irrespective of BI receives same treatment i.e., patients with BI 6.0 and 2.0 get the same MB MDT for 1 year, even though it was found that BI decreases with a rate of only 0.5–1 log-unit/year after starting treatment. Bacteriological index was found to be 1.0 in some patients even after release from treatment with an initial BI of 3.6 (23). In these scenarios there is a possibility of development of drug resistance within *M. leprae* and then relapse of the disease. Patients' irregular treatment can also be one of the contributing factors for drug resistance. These drug resistance *M. leprae* isolates from the relapse patients can infect healthy individuals.

Excerpts from a WHO document regarding the drug resistance surveillance program (53):

- It is assumed that a combination of three drugs, if taken regularly, will prevent the emergence of drug resistance. In addition, there is limited information on patient adherence with the unsupervised components of MDT.
- Lack of priority and the absence of information on current magnitude of drug resistance in leprosy is not evidence of an absence in drug resistance.

- Although at this point in time the problem of drug resistance is not acute, it is important that data is collected systematically and trends monitored carefully so that effective measures to combat this problem can be developed for the future. The recent development of a more practical and quick DNA sequencing methods to detect drug resistance has greatly helped in establishing a sentinel surveillance net-work.
- As a first step, WHO held an informal meeting in November 2006 in Agra, India to review the problem of drug resistance and recent advances in technology for testing drug resistance based on DNA sequencing. As a follow-up to this meeting, another meeting was held in Hanoi, Viet Nam in October 2008 to develop guidelines for sentinel surveillance. Based on the discussions and recommendations of these two meetings, a Guideline for Global Surveillance of Drug Resistance in Leprosy *(SEA-GLP-2009.2)* was published (53).

Hence, this study has been conducted with the following objective:

- To monitor if drug resistance is a factor in the leprosy incidence in Philippines, we intent to study primary drug resistance from newly diagnosed leprosy patients from Cebu, Philippines by molecular approach.

1.6.3 MATERIAL AND METHODS:

The patients in this study were enrolled from those who consulted at the Cebu Skin Clinic (CSC), Leonard Wood Memorial Leprosy Research Centre, Philippines. Skin biopsies from 98 newly diagnosed leprosy patients were collected in 70% ethanol and were used this study. Detailed medical history including the leprosy condition and

treatment regimen, information about the patients' geographical area of residence (village, city and province), types /place(s) of employment and education, nature of dwelling, sources of drinking, cooking, and bathing water spanning a prior 20 year period, was captured with a standardized questionnaire at the time of recruitment since the study began. All the procedures involving biological sample collections were performed following the approval from the governing human research ethical committee and informed consent procedures as necessary.

The Qiagen DNeasy Tissue kit (Qiagen, CA) was used to obtain total DNA from the skin biopsy. DNA was eluted in 100-200 µl elution buffer, and typically 1-2 µl of DNA was sufficient for one PCR. For the amplification of DRDRs within the *folp1*, *rpoB*, *gyrA* and *gyrB* genes, multiplex-PCR was performed using the Multiplex PCR kit (Qiagen). The sequence information for the primers is listed in Table 1.6.1. Two separate working stocks of both the forward and reverse primer mixes, were prepared and adjusted to 200 µl final volume with TE buffer, such that each primer was at a concentration of 2 µM.

Each PCR reaction (20 µl final volume) was assembled in a PCR cabinet at room temperature and comprised 10 µl of 2X Qiagen multiplex-PCR Master mix, 2 µl Q solution, 2 µl each of the forward and reverse primer working stocks, and 2 µl of DNA template; the final volume was adjusted with PCR grade water (2 µl). The final concentration of each primer was thus 0.2 µM. Following an activation step at 95°C for 15 min, 40 cycles of PCR were run as follows: denaturation at 94°C for 30 sec, primer annealing at 60°C for 90 sec, and primer extension at 72°C for 90 sec. The PCR was terminated with a final extension at 72°C for 10 min. Armadillo derived *M. leprae* DNA

	Forward Primer sequence	Reverse Primer sequence	Size (in bp)
rpoB	CAGGACGTCGAGGCGATCAC	TCG TCA GCG GTC AAG TA	386
folP1	TTCGTTCTCAGATGGCGGAC	GCCCACCAGACACATCGTTG	281
gyrA	CCG TAG CCA CGC TAA GTC A	CCG GCG AAC CGA AAT TGC C	186
gyrB	ACTGATCCTCGAAGTTCTGAACTG	CAATGCCGTAATAATTTGCTTGAA	158

Table 1.6.1 Multiplex-PCR primers for amplification of *M. leprae* drug resistance determining regions (DRDR) within *rpoB*, *folp1*, *gyrA* and *gyrB* genes for detecting mutations of drug resistance for dapsone, rifampicin and ofloxicin.

(NHDP63) was used as a positive control (13, 20). The DNA products were resolved by agarose gel electrophoresis and detected by ethidium bromide staining. For DNA sequencing, the multiplex-PCR sample was either simply diluted 10 fold or ethanol precipitated and reconstituted; then 1-2 µl of this product was combined with 10 pmoles of the forward primer for the Big Dye cycle sequencing reaction at the Macromolecular Resource Facility (now called as Proteomics and Metabolomics Facility) at CSU.

1.6.4 RESULTS AND DISCUSSION

All the patients in this study were treated with WHO MDT, one year for MB and six months for PB leprosy. After examining the clinical details, it was found that the age of the patients ranged from 12 to 68 years, ten patients were found to have leprosy reactions before the treatment. The patients' average bacteriological index (BI) is 3.1.

We have developed a multiplex-PCR approach for amplifying all the four DRDRs within the *folp1*, *rpoB*, *gyrA* and *gyrB* genes in a single PCR for detecting drug resistance mutations in the clinical isolates (Fig. 1.6.2). Multiplex-PCR was performed for ninety eight samples; however amplification could not be obtained for twenty one samples, which are from the patients with low BI (< 1.5). The multiplex-PCR amplicons from the remaining seventy seven samples were subjected to DNA sequencing with *folp1* and *rpoB* forward primers separately. Sequencing results from direct multiplex-PCR products (after diluting 10 fold) were sometimes unreadable due to mixed signals, and interference with dye bobs. These problems could be eliminated or reduced by ethanol precipitation of multiplex-PCR products, reconstitution of the DNA in TE and followed by sequencing.

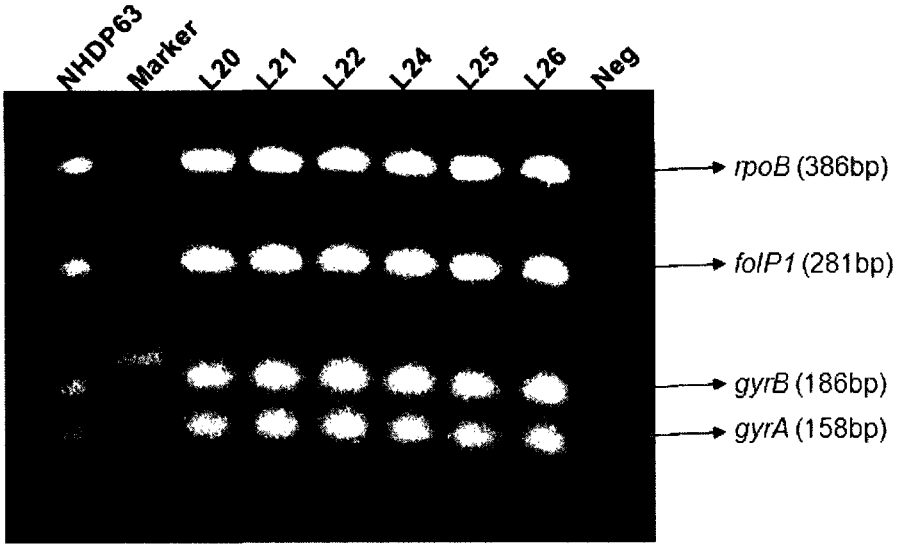


Figure 1.6.2: Gel electrophoresis for detection of the amplification of *M. leprae* DRDRs within *rpoB*, *folp1*, *gyrA* and *gyrB* genes from the reference strain (NHDP63) and 6 Cebu clinical isolates (L20, L21, L22, L24, L25 and L26). Marker: 20-bp DNA marker (Bio-rad); Neg: negative control (no template).

There were no mutations in the DRDRs of *folp1* and *rpoB* and specifically at codons 53 and 55 of *folp1* and at codons 407, 410, 420, 425 of *rpoB* in the seventy seven *M. leprae* isolates, indicating a low possibility of resistance to dapsone and rifampicin. During a sampling of newly diagnosed leprosy patients in Cebu mutations were not found in the fifteen *M. leprae* isolates tested, at codons 89 and 91 of the *gyrA* gene for quinolone resistance.

The primary dapsone resistance in Cebu, Philippines was previously reported to be 3.6% in the period 1975-1978, 8.1% and in the period 1979-1982 by mouse foot-pad assays (8). In a recent study by Matsuoka et al. (27), mutations were found in the *folp1* DRDR in 2 out of 77 new cases tested i.e. 2.6% of primary dapsone resistance in Cebu, Philippines, but primary resistance was not seen in case of rifampicin and ofloxacin. However, in our findings we do not see any primary resistance to dapsone, rifampicin and oflaxacin based on the sampling of 77 patients from Cebu, Philippines during 2006-2007. It is not clear from our results if there is any change in the level of primary dapsone resistance due to the limitation in the sample size or improved compliance. Long- term follow up of the new cases and large sample sizes are required to monitor systematic surveillance.

1.6.5 References:

1. **Barry, V. C., J. G. Belton, M. L. Conalty, J. M. Denneny, D. W. Edward, J. F. O'Sullivan, D. Twomey, and F. Winder.** 1957. A new series of phenazines (rimino-compounds) with high antituberculosis activity. Nature **179:**1013-5.

2. **Blanchard, J. S.** 1996. Molecular mechanisms of drug resistance in Mycobacterium tuberculosis. Annu Rev Biochem **65:**215-39.

3. **Cambau, E., P. Bonnafous, E. Perani, W. Sougakoff, B. Ji, and V. Jarlier.** 2002. Molecular detection of rifampin and ofloxacin resistance for patients who experience relapse of multibacillary leprosy. Clin Infect Dis **34:**39-45.

4. **Cambau, E., L. Carthagena, A. Chauffour, B. Ji, and V. Jarlier.** 2006. Dihydropteroate synthase mutations in the folP1 gene predict dapsone resistance in relapsed cases of leprosy. Clin Infect Dis **42:**238-41.

5. **Champoux, J. J.** 2001. DNA topoisomerases: structure, function, and mechanism. Annu Rev Biochem **70:**369-413.

6. **Cochrane, R. G.** 1948. A comparison of sulphone and hydnocarpus therapy of leprosy. Int J Lepr **16:**139-44.

7. **Colston, M. J., Hilson, G. R. F., and Banerjee, D. K.** 1978. The 'proportional bactericidal test': a method for assessing bactericidal activity of drugs against *Mycobacterium leprae* in mice. Lepr. Rev. **49:**7- 15.

8. **dela Cruz, E., R. V. Cellona, M. V. Balagon, L. G. Villahermosa, T. T. Fajardo, Jr., R. M. Abalos, E. V. Tan, and G. P. Walsh.** 1996. Primary dapsone resistance in Cebu, The Philippines; cause for concern. Int J Lepr Other Mycobact Dis **64:**253-6.

9. **Faget G.H., P. R. C., Johansen F.A., Dinan J.F., Prejean B.M., Eccles C.G.** 1943. The promin treatment of leprosy, A progress report. . United States Government Printing Office.

10. **Feldman, W. H., A. G. Karlson, and H. C. Hinshaw.** 1948. Promin in experimental tuberculosis; antituberculosis effects of sodium P,P'-diaminodiphenylsulfone-N,N'-didextroxe sulfonate (Promin) administered subcutaneously (a preliminary report). Mayo Clin Proc **23:**118-25.

11. **Gelber, R. H.** 1994. Chemotherapy of lepromatous leprosy: recent developments and prospects for the future. Eur J Clin Microbiol Infect Dis **13:**942-52.

12. **Gillis, T. P., and D. L. Williams.** 2000. Dapsone resistance in *Mycobacterium leprae*. Lepr Rev **71 Suppl:**S91-5.

13. **Groathouse, N. A., B. Rivoire, H. Kim, H. Lee, S. N. Cho, P. J. Brennan, and V. D. Vissa.** 2004. Multiple polymorphic loci for molecular typing of strains of *Mycobacterium leprae*. J Clin Microbiol **42:**1666-72.

14. **Hilson, G. R., D. K. Banerjee, and I. B. Holmes.** 1971. The activity of various antituberculous drugs in suppressing experimental *Mycobacterium leprae* infection in mice. Int J Lepr Other Mycobact Dis **39:**349-53.

15. **Honore, N., and S. T. Cole.** 1993. Molecular basis of rifampin resistance in *Mycobacterium leprae*. Antimicrob Agents Chemother **37:**414-8.

16. **Honore, N., E. Perrani, A. Telenti, J. Grosset, and S. T. Cole.** 1993. A simple and rapid technique for the detection of rifampin resistance in *Mycobacterium leprae*. Int J Lepr Other Mycobact Dis **61:**600-4.

17. **Honore, N., P. W. Roche, J. H. Grosset, and S. T. Cole.** 2001. A method for rapid detection of rifampicin-resistant isolates of *Mycobacterium leprae*. Lepr Rev **72:**441-8.

18. **Kai, M.** 2004. [Diaminodiphenylsulfone resistance of *Mycobacterium leprae* due to mutations in the dihydropteroate synthase gene]. Nihon Hansenbyo Gakkai Zasshi **73:**221-6.

19. **Kai, M., M. Matsuoka, N. Nakata, S. Maeda, M. Gidoh, Y. Maeda, K. Hashimoto, K. Kobayashi, and Y. Kashiwabara.** 1999. Diaminodiphenylsulfone resistance of *Mycobacterium leprae* due to mutations in the dihydropteroate synthase gene. FEMS Microbiol Lett **177:**231-5.

20. **Kimura, M., R. M. Sakamuri, N. A. Groathouse, B. L. Rivoire, D. Gingrich, S. Krueger-Koplin, S. N. Cho, P. J. Brennan, and V. Vissa.** 2009. Rapid variable-number tandem-repeat genotyping for *Mycobacterium leprae* clinical specimens. J Clin Microbiol **47:**1757-66.

21. **Kulkarni, V. M., and J. K. Seydel.** 1983. Inhibitory activity and mode of action of diaminodiphenylsulfone in cell-free folate-synthesizing systems prepared from Mycobacterium lufu and *Mycobacterium leprae*. A comparison. Chemotherapy **29:**58-67.

22. **Levy, L., and B. Ji.** 2006. The mouse foot-pad technique for cultivation of *Mycobacterium leprae*. Lepr Rev **77:**5-24.

23. **Linder, K., M. Zia, W. V. Kern, R. K. Pfau, and D. Wagner.** 2008. Relapses vs. reactions in multibacillary leprosy: proposal of new relapse criteria. Trop Med Int Health **13:**295-309.

24. **Maeda, S., M. Matsuoka, N. Nakata, M. Kai, Y. Maeda, K. Hashimoto, H. Kimura, K. Kobayashi, and Y. Kashiwabara.** 2001. Multidrug resistant

Mycobacterium leprae from patients with leprosy. Antimicrob Agents Chemother **45:**3635-9.

25. **Matrat, S., E. Cambau, V. Jarlier, and A. Aubry.** 2008. Are all the DNA gyrase mutations found in *Mycobacterium leprae* clinical strains involved in resistance to fluoroquinolones? Antimicrob Agents Chemother **52:**745-7.

26. **Matsuoka, M., K. S. Aye, K. Kyaw, E. V. Tan, M. V. Balagon, P. Saunderson, R. Gelber, M. Makino, C. Nakajima, and Y. Suzuki.** 2008. A novel method for simple detection of mutations conferring drug resistance in *Mycobacterium leprae*, based on a DNA microarray, and its applicability in developing countries. J Med Microbiol **57:**1213-9.

27. **Matsuoka, M., T. Budiawan, K. S. Aye, K. Kyaw, E. V. Tan, E. D. Cruz, R. Gelber, P. Saunderson, V. Balagon, and V. Pannikar.** 2007. The frequency of drug resistance mutations in *Mycobacterium leprae* isolates in untreated and relapsed leprosy patients from Myanmar, Indonesia and the Philippines. Lepr Rev **78:**343-52.

28. **Matsuoka, M., Y. Kashiwabara, and M. Namisato.** 2000. A *Mycobacterium leprae* isolate resistant to dapsone, rifampin, ofloxacin and sparfloxacin. Int J Lepr Other Mycobact Dis **68:**452-5.

29. **Morrison, N. E., and G. M. Marley.** 1976. Clofazimine binding studies with deoxyribonucleic acid. Int J Lepr Other Mycobact Dis **44:**475-81.

30. **Morrison, N. E., and G. M. Marley.** 1976. The mode of action of clofazimine DNA binding studies. Int J Lepr Other Mycobact Dis **44:**133-4.

31. **Oliva, B., A. J. O'Neill, K. Miller, W. Stubbings, and I. Chopra.** 2004. Anti-staphylococcal activity and mode of action of clofazimine. J Antimicrob Chemother **53:**435-40.

32. **Pallanza, R., V. Arioli, S. Furesz, and G. Bolzoni.** 1967. Rifampicin: a new rifamycin. II. Laboratory studies on the antituberculous activity and preliminary clinical observations. Arzneimittelforschung **17:**529-34.

33. **Rees, R. J.** 1964. Limited Multiplication of Acid-Fast Bacilli in the Foot-Pads of Mice Inoculated with *Mycobacterium leprae*. Br J Exp Pathol **45:**207-18.

34. **Rees, R. J.** 1969. New prospects for the study of leprosy in the laboratory. Bull World Health Organ **40:**785-800.

35. **Rees, R. J., J. M. Pearson, and M. F. Waters.** 1970. Experimental and clinical studies on rifampicin in treatment of leprosy. Br Med J **1:**89-92.

36. **Sapkota, B. R., C. Ranjit, and M. Macdonald.** 2006. Reverse line probe assay for the rapid detection of rifampicin resistance in *Mycobacterium leprae*. Nepal Med Coll J **8:**122-7.

37. **Sapkota, B. R., C. Ranjit, K. D. Neupane, and M. Macdonald.** 2008. Development and evaluation of a novel multiple-primer PCR amplification refractory mutation system for the rapid detection of mutations conferring rifampicin resistance in codon 425 of the rpoB gene of *Mycobacterium leprae*. J Med Microbiol **57:**179-84.

38. **Sekar, B., N. Elangeswaran, E. Jayarama, M. Rajendran, S. S. Kumar, R. Vijayaraghavan, D. Anandan, and K. Arunagiri.** 2002. Drug susceptibility of *Mycobacterium leprae*: a retrospective analysis of mouse foot-pad inoculation results from 1983 to 1997. Lepr Rev **73:**239-44.

39. **Sensi, P., Margalith, P., Timbal, M. T.** 1959. Rifomycin, a new antibiotic—preliminary report. . *Farmaco Ed Sci* **14:**146-147.

40. **Shepard, C. C.** 1960. Acid-fast bacilli in nasal excretions in leprosy, and results of inoculation of mice. Am J Hyg **71:**147-57.

41. **Shepard, C. C.** 1960. The experimental disease that follows the injection of human leprosy bacilli into foot-pads of mice. J Exp Med. **112:**445-454.

42. **Shepard, C. C.** 1967. A kinetic method for the study of activity of drugs against *Mycobacterium leprae* in mice. Int. J. Lepr. **35:**429-435.

43. **Shepard, C. C.** 1982. Statistical analysis of results obtained by two methods for testing drug activity against *Mycobacterium leprae*. Int J Lepr Other Mycobact Dis. **50:**96-101.

44. **Shepard, C. C., and Chang, Y. T.** 1967. Effect of DDS on established infections with *Mycobacterium leprae* in mice. Int. J. Lepr. **35:**52-57.

45. **Shepard, C. C., and D. H. McRae.** 1965. *Mycobacterium leprae* in Mice: Minimal Infectious Dose, Relationship between Staining Quality and Infectivity, and Effect of Cortisone. J Bacteriol **89:**365-72.

46. **Shetty, V. P., A. V. Wakade, S. Ghate, V. V. Pai, R. Ganapati, and N. H. Antia.** 2003. Viability and drug susceptibility testing of M. leprae using mouse footpad in 37 relapse cases of leprosy. Int J Lepr Other Mycobact Dis **71:**210-7.

47. **Van Rensburg, C. E., G. K. Joone, J. F. O'Sullivan, and R. Anderson.** 1992. Antimicrobial activities of clofazimine and B669 are mediated by lysophospholipids. Antimicrob Agents Chemother **36:**2729-35.

48. **Verbist, L., and A. Gyselen.** 1968. Antituberculous activity of rifampin in vitro and in vivo and the concentrations attained in human blood. Am Rev Respir Dis **98:**923-32.

49. **Williams, D. L., and T. P. Gillis.** 1999. Detection of drug-resistant *Mycobacterium leprae* using molecular methods. Indian J Lepr **71:**137-53.

50. **Williams, D. L., and T. P. Gillis.** 2004. Molecular detection of drug resistance in *Mycobacterium leprae*. Lepr Rev **75:**118-30.

51. **Williams, D. L., L. Spring, E. Harris, P. Roche, and T. P. Gillis.** 2000. Dihydropteroate synthase of *Mycobacterium leprae* and dapsone resistance. Antimicrob Agents Chemother **44:**1530-7.

52. **World Health Organization.** 1982. Chemotherapy of leprosy for control programme. . Geneva: WHO. Technical Report Series. **768.**

53. **World Health Organization**. 2009. Report of the Workshop on Sentinel Surveillance for Drug Resistance in Leprosy, 20-22 October, 2008, Hanoi, Vietnam.

Chapter 1.7

Summary and Conclusions for Molecular epidemiology of leprosy in Cebu, Philippines

Even though leprosy is the oldest known human disease and its cause by an obligate intra cellular bacterium *Mycobacterium leprae*, has been known since 100 years (6), the exact mechanism of transmission is not yet clearly understood whether it is human to human, environmental and/or by other means (1, 9-11, 13, 18). Tools to address these studies were also not available until genome sequencing of *M. leprae* was completed (2). Prior to this, all the isolates of *M. leprae* obtained from different sources were found to non-polymorphic in nature when assessed by a number of genetic techniques (3, 19, 20). Molecular markers such as short tandem repeats (STRs) and single nucleotide polymorphisms (SNPs) have since then been identified by genomic and comparative genomic analysis of *M. leprae* strains (4, 14, 21).

In this regard, we have undertaken a molecular approach to study leprosy in endemic situations. Philippines is one of the study sites. Initially we developed methods and applied them in the field setting. In this regard we have developed a multiplex-PCR approach for rapid amplification of fifteen VNTR loci in four PCRs (combinations) compressing of 9 microsatellites and 6 minisatellites in order to reduce the consumption of DNA template. We also developed a high throughput fragment length analysis (FLA) method to determine the VNTR allele for each locus in the multiplex PCR combination. We validated and standardized the FLA using reference DNA (NHDP63), a human isolate of *M. leprae* propagated in armadillos. We then demonstrated the utility of these methods for rapid mapping of 15 VNTRs using the 30 stored clinical specimens initially. We also developed a PCR-RFLP method for rapid sub-typing of SNPs (type 1, 2, 3 and 4). The SNP type 3 or 4 can be distinguished from type 1 or 2, by testing susceptibility of locus 3 (2935685, base pair position in *M. leprae* TN strain) to BstUI. SNP type 4 is

sensitive to SmlI at locus 1 (14676), while SNP types 1, 2 and 3 are not. Similarly, SNP types 2, 3 and 4 are sensitive to CviKI-1 at locus 2 (1642875), while SNP type 1 is not.

Skin biopsies and slit skin smear (SSS) samples were collected at Cebu skin clinic (CSC), Leonard Wood memorial (LWM) in Cebu, Philippines and sent to CSU for further analysis for the ongoing study (L samples). Stored skin biopsies were also obtained. DNA was extracted from the clinical samples and VNTR data for 15 VNTR loci was obtained from 207 samples by multiplex PCR and FLA methods. The allelic diversity for the VNTR loci ranged from 0.1 to 0.94, indicating that some markers highly stable (low discriminatory) while others are highly variable (high discriminatory) which is believed to be good scenario for strain typing. VNTR data from ten high BI patients' biopsies and SSS samples were found to be similar indicating the stability of VNTRs in different body compartments. Five groups emerged in the phylogenetic tree which were named as A, B, C, D and T, each with a distinct VNTR signature pattern. A total of 37 SNP type 3 samples (17% of the study population) were detected. In the phylogentic tree which was based on VNTRs, all the SNP type 3 isolates clustered in Group C and the VNTR allele profiles for SNP type 3 were quite distinct from that of SNP type 1 isolates. Conserved VNTR profiles were observed within the multicase families (MCF) which validate the ability of the strain typing method in tracing sources of infection or identifying a transmission network.

The overall population structure of *M. leprae* isolates from Cebu remained similar with stable clusters in the phylogentic tree, even with the addition of 70 samples. Another three MCFs were studies in this group. The VNTR profiles of *M. leprae* within the individual MCF were found to be stable enough during incubation indicating

transmission between the individuals. Multiplex PCR and FLA methods were found to be better compared to single PCR and conventional DNA sequencing methods.

Since *M. leprae* could not be cultured in artificial media the hence the only method to test drug resistance is mouse footpad assay which is a time consuming (12 months) and laborious method. So for these reasons molecular approach is the most common method to detect drug resistance in *M. leprae* isolates. To study primary drug resistance to leprosy drugs from newly diagnosed leprosy patients from Cebu (L samples) by molecular approach; a multiplex PCR was developed for amplifying four drug resistance determining regions (DRDR) within the *folp1*, *rpoB*, *gyrA* and *gyrB* genes in a single PCR for detecting drug resistance mutations. In the sample set of 77 samples we examined no DRDR mutations in *rpoB* and *folP1* (targets for rifampicin and dapsone). Mutations were also not observed in the *gyrA* (target for oflaxacin) DRDR from 15 samples tested.

Future directions

Even though major clusters with distinct VNTR signatures were observed in the phylogenetic tree, relationships of *M. leprae* isolates within the clusters could not be defined, since the samples used for this study were from different time periods with few/no known epidemiological linkages. In the samples from the current study the patient coverage in the island of Cebu is only partial. The movement of patients from place to place and long incubation time of leprosy are the other hurdles in tracking transmission patterns.

So to address these issues a long term molecular epidemiology database with maximum coverage of patients in a geographical region is needed in monitoring genotypes/clusters and thus to uncover transmission patterns combined with epidemiological investigations. So with this long term approach we can identify the evolutionary scenario of *M. leprae* isolates in that region (i.e. microevolution). If the *M. leprae* strains circulating in Cebu, Philippines are same then the major clusters obtained through phylogenetic analysis will also remains the same.

With the help of epidemiological information, we can also identify the known leprosy patient contacts of a new patient and obtain samples from them. We can obtain the VNTR data from both the leprosy patient contact and new patient and then compare the MLVA profiles. By this way a transmission network can be generated which is similar to that of clustered cases of tuberculosis (7). Incorporation of VNTR data from these epidemiological linked samples into the phlyogentic analysis helps in defining a cluster, if it is true or not. Global positioning system (GPS) information can also be incorporated into the epidemiological information database to track the leprosy endemic pockets within a geographical region.

VNTR data from paucibaillary cases ($1/10^{th}$ of the study population) was missing; possibility of propagating the bacilli in the mouse foot pads is an alternative approach.

Addition of new VNTR markers in the phlyogenetic analysis was recommended for enhancing accuracy of VNTR-based phylogenetic trees based on simulation study by Hall (5). It was also proposed that for addition of a single marker increases the accuracy of the phylogentics analysis by 2%. A hierarchical phylogenetic analysis could be

adopted similar to that of *Bacillus anthracis* (8), by initially resolving with SNPs, then by low mutation rate VNTR loci and finally with high mutation rate VNTR loci.

Homoplasy is a condition in which two isolates share common character state (for eg. same VNTR profiles), these isolates might not necessarily originate from a common ancestor. But the common character state in the two isolates might be due to multiple independent events. It was hypothesized that homoplasy occurs mainly with the markers with high mutational rates like (AT)15, (AT)17, (TA)18 and (TTC)21. Weighting of markers was found to minimize homoplasy (15). Since we utilized this weighting method in our phylogenetic analysis we might have minimized the effect of homoplasy.

Even though several hybridization techniques were developed (12, 16, 17), other than the common sequencing method for rapid detection of mutations in the target regions (DRDR) from the clinical specimens, the sensitivity and specificity of these methods are low; so if a mutation is detected in a *M. leprae* isolate it needs to be reconfirmed by sequencing method and also new mutations cannot be identified by these methods. However, a rapid screening for detecting common drug resistance mutations can be done by these hybridization methods.

REFERENCES

1. **Chakrabarty, A. N., and S. G. Dastidar.** 2001. Is soil an alternative source of leprosy infection? Acta Leprol **12:**79-84.

2. **Cole, S. T., K. Eiglmeier, J. Parkhill, K. D. James, N. R. Thomson, P. R. Wheeler, N. Honore, T. Garnier, C. Churcher, D. Harris, K. Mungall, D. Basham, D. Brown, T. Chillingworth, R. Connor, R. M. Davies, K. Devlin, S. Duthoy, T. Feltwell, A. Fraser, N. Hamlin, S. Holroyd, T. Hornsby, K. Jagels,**

C. Lacroix, J. Maclean, S. Moule, L. Murphy, K. Oliver, M. A. Quail, M. A. Rajandream, K. M. Rutherford, S. Rutter, K. Seeger, S. Simon, M. Simmonds, J. Skelton, R. Squares, S. Squares, K. Stevens, K. Taylor, S. Whitehead, J. R. Woodward, and B. G. Barrell. 2001. Massive gene decay in the leprosy bacillus. Nature **409:**1007-11.

3. **Cole, S. T., P. Supply, and N. Honore.** 2001. Repetitive sequences in Mycobacterium leprae and their impact on genome plasticity. Lepr Rev **72:**449-61.

4. **Groathouse, N. A., B. Rivoire, H. Kim, H. Lee, S. N. Cho, P. J. Brennan, and V. D. Vissa.** 2004. Multiple polymorphic loci for molecular typing of strains of Mycobacterium leprae. J Clin Microbiol **42:**1666-72.

5. **Hall, B. G.** 2009. Molecular Epidemiology of Mycobacterium leprae: a solid beginning. Lep Rev **(in press)**.

6. **Hansen, G. H. A.** 1874. Undersogelser angaende spedalskhedens aasager. Norsk Magazin for Laegervidenskaben **4 (Suppl.):**1-88.

7. **Houben, R. M., and J. R. Glynn.** 2009. A systematic review and meta-analysis of molecular epidemiological studies of tuberculosis: development of a new tool to aid interpretation. Trop Med Int Health **14:**892-909.

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

9. **Lahiri, R., and J. L. Krahenbuhl.** 2008. The role of free-living pathogenic amoeba in the transmission of leprosy: a proof of principle. Lepr Rev **79:**401-9.

10. **Lavania, M., K. Katoch, V. M. Katoch, A. K. Gupta, D. S. Chauhan, R. Sharma, R. Gandhi, V. Chauhan, G. Bansal, P. Sachan, S. Sachan, V. S. Yadav, and R. Jadhav.** 2008. Detection of viable Mycobacterium leprae in soil samples: insights into possible sources of transmission of leprosy. Infect Genet Evol **8:**627-31.

11. **Lavania, M., K. Katoch, P. Sachan, A. Dubey, S. Kapoor, M. Kashyap, D. S. Chauhan, H. B. Singh, V. D. Sharma, R. S. Jadhav, and V. M. Katoch.** 2006. Detection of Mycobacterium leprae DNA from soil samples by PCR targeting RLEP sequences. J Commun Dis **38:**269-73.

12. **Matsuoka, M., K. S. Aye, K. Kyaw, E. V. Tan, M. V. Balagon, P. Saunderson, R. Gelber, M. Makino, C. Nakajima, and Y. Suzuki.** 2008. A novel method for simple detection of mutations conferring drug resistance in Mycobacterium leprae, based on a DNA microarray, and its applicability in developing countries. J Med Microbiol **57:**1213-9.

13. **Matsuoka, M., S. Izumi, T. Budiawan, N. Nakata, and K. Saeki.** 1999. Mycobacterium leprae DNA in daily using water as a possible source of leprosy infection. Indian J Lepr **71:**61-7.

14. **Monot, M., N. Honore, T. Garnier, R. Araoz, J. Y. Coppee, C. Lacroix, S. Sow, J. S. Spencer, R. W. Truman, D. L. Williams, R. Gelber, M. Virmond, B. Flageul, S. N. Cho, B. Ji, A. Paniz-Mondolfi, J. Convit, S. Young, P. E. Fine, V. Rasolofo, P. J. Brennan, and S. T. Cole.** 2005. On the origin of leprosy. Science **308:**1040-2.

15. **Philippe, H., G. Lecointre, H. L. VanLe, and H. LeGuyader.** 1996. A critical study of homoplasy in molecular data with the use of a morphologically based cladogram, and its consequences for character weighting. Molecular Biology and Evolution **13:**1174-1186.

16. **Sapkota, B. R., C. Ranjit, and M. Macdonald.** 2006. Reverse line probe assay for the rapid detection of rifampicin resistance in Mycobacterium leprae. Nepal Med Coll J **8:**122-7.

17. **Sapkota, B. R., C. Ranjit, K. D. Neupane, and M. Macdonald.** 2008. Development and evaluation of a novel multiple-primer PCR amplification refractory mutation system for the rapid detection of mutations conferring rifampicin resistance in codon 425 of the rpoB gene of Mycobacterium leprae. J Med Microbiol **57:**179-84.

18. **Truman, R.** 2008. Armadillos as a source of infection for leprosy. South Med J **101:**581-2.

19. **Williams, D. L., and T. P. Gillis.** 1989. A study of the relatedness of Mycobacterium leprae isolates using restriction fragment length polymorphism analysis. Acta Leprol **7 Suppl 1:**226-30.

20. **Williams, D. L., T. P. Gillis, and F. Portaels.** 1990. Geographically distinct isolates of Mycobacterium leprae exhibit no genotypic diversity by restriction fragment-length polymorphism analysis. Mol Microbiol **4:**1653-9.

21. **Zhang, L., T. Budiawan, and M. Matsuoka.** 2005. Diversity of potential short tandem repeats in Mycobacterium leprae and application for molecular typing. J Clin Microbiol **43:**5221-9.

Chapter 2.1

Essentiality and characterization of an arabinosyltransferase involved in the cell wall synthesis of mycobacteria.

2.1.1 ABSTRACT

The cell envelope of mycobacteria are comprised of a core macromolecule known as mycolyl arabinogalactan (mAG) which is covalently attached to the peptidoglycan layer forming the mAGP complex and other macromolecules such as lipomannan and lipoarabinomannan. The mAGP complex is essential for survival of the bacteria and is a permeability barrier. The synthesis and assembly of this mAGP complex is still not fully understood. In this regard, through comparative genome approaches we identified *Rv3805c* as a putative glycosyltransferase located in the cell wall synthesis gene cluster. Rv3805c predicted to have ten transmembrane domains, contains a conserved glycosyltransferase and proline rich motif also present in Emb proteins which are involved in cell wall biosynthesis. In this study we identified that, *Rv3805c* encodes an arabinosyltransferase through an *in vitro* arabionosyltransferase assay. Furthermore, a knock out mutant of an ortholog gene *NCgl2780* was generated in *C. glutamicum*. The structural analysis of AG of *NCgl2780* mutant in *C. glutamicum* revealed the loss of β(1→2) Ara*f* residues at the non-reducing ends. We also indentified that the ortholog *MSMEG_6400* is essential for *M. smegmatis*, as we were unable to obtain a genetic knock out mutant in the absence of an additional functional copy of the gene.

2.1.2 INTRODUCTION

2.1.2.1 Cell wall of mycobacteria

The cell envelope of mycobacterial sps are composed of plasma membrane and a unique cell wall which is specific to Family Mycobacteriaceae. The cell wall is mainly made up of peptidoglycan, complex polysaccharides like arabinogalactan (AG) and

unique lipids (mycolic acids). These complex molecules make the mycobacterial cell a tough cell envelope. Mycolic acids are long chain α-alkyl β-hydroxy fatty acids (C70-C90) and are essential to mycobacteria (3, 29, 32). The mycolic acids are attached to the nonreducing ends of the arabinogalactan (42). Whereas the reducing ends of AG is covalently attached to the highly cross-linked peptidoglycan via a linker unit which is phosphoryl-*N*-acetylglucosaminosyl-rhamnosyl (P-GlcNAc-Rha) (41). This massive structure, the mycolate-arabinogalactan-peptidoglycan-complex (mAGP), is the basis of many of the physiological and pathogenic features of *M. tuberculosis* and the site of susceptibility and resistance to many of the cell wall targeting anti-tuberculosis drugs like ethambutol, isoniazid, ethionamide (Fig. 2.1.1). Lipoarabinomannan (LAM) is another major macromolecule in the mycobacterial cell envelope, which modulates host immune response. It is mainly a heterogeneous mixture of lipoglycans.

2.1.2.2 Peptidoglycan

The peptidoglycan in mycobacterial sps consists of alternating units of N-acetyl-α-D-glucosamine (GlcNAc) and N-glycolyl-muramic acid (MurNGlyc). The tetrapeptide chain (L-alaninyl-D-isoglutaminyl-meso-diaminopimelyl-D-alanine) is attached to the carboxylic acid group of MurNGlyc (40). Further two residues of diaminopimelic acid or diaminopimelic acid and D-alanine residue are cross linked between two chains of peptidoglycan (11, 40).

2.1.2.3Arabinogalacatan

The structure of AG is shown in Fig. 2.1.2 (14). The biosynthesis of AG is initiated with the transfer of GlcNAc-1-P and Rha from their nucleotide sugar donors

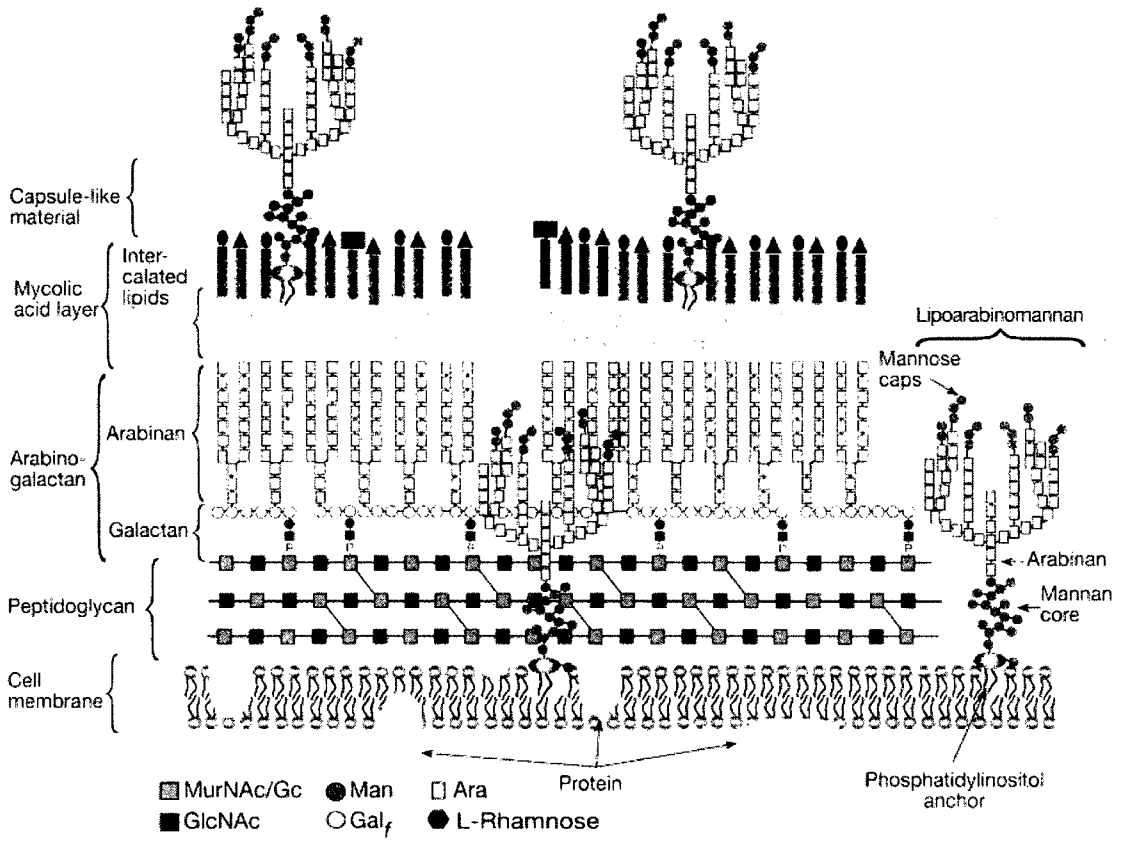


FIGURE 2.1.1 A model representing the cell wall of mycobacteria [taken from Brennan and Crick (11)]

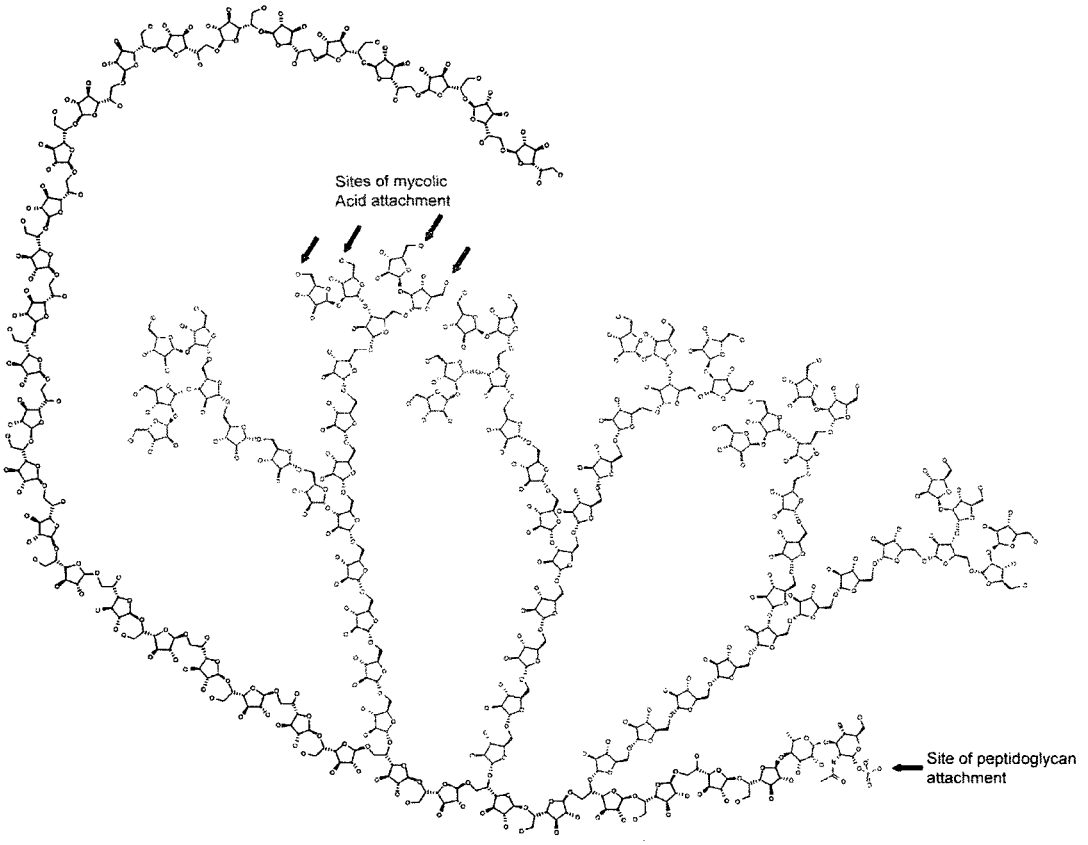


FIGURE 2.1.2 Structure of Arabinogalactan (AG) from *M. tuberculosis* [taken from Crick et al. (14)] Linker unit (Rha-GlcNAc) is shown in purple, galactan chain in black, and the arabinan chain in red color. The terminal hexa-arabinan motif at the non-reducing end of arabinangalactan molecule to which the mycolic acids are attached is shown in blue color.

(UDP-GlcNAc and dTDP-Rhamnose) on to the polyprenol phosphate acceptor with the help of two glycosyltansferases (WecA and WbbL), to form polyprenol-P-P-GlcNAc-Rha (46, 70). Then the initial two Gal*f* (galactofuranose) residues from UDP-Gal*f* are added on to polyprenol-P-P-GlcNAc-Rha to form polyprenol-P-P-GlcNAc-Rha-Gal-Gal by a galactosyltrasferase enzyme, GlfT1 (Rv3782) (44) and later Gal*f* residues are sequentially added with alternating β(1→5) and β(1→6) linkages by a bi-functional galactosyltrasferase enzyme, GlfT2 (Rv3808c) to form a polyprenol-P-P-GlcNAc-Rha-Gal_{30} (6, 33, 48). Further, this polyprenol-P-P-GlcNAc-Rha-Gal_{30} or its intermediates serve as an acceptor for arabinosylation from a decaprenyl-monophosphoryl- β D-arabinose (DPA) donor (69). It was found that the arabinose in the DPA is obtained from 5'-phosphoribosyl-1-phosphate (PRPP) which is produced in non-oxidative pentose phosphate shunt (61). The biosynthetic pathway of DPA from PRPP has also been elucidated (23, 45). Ethambutol (EMB), the frontline drug for anti-tuberculosis was found to inhibit the polymerization of arabinan in the cell wall in both AG and LAM (16, 47). When the EMB resistance region of the DNA from *M. avium* is cloned and expressed in *M. smegmatis* it was found that the minimal inhibitory concentration (MIC) is 10 fold higher than the vector control and this led to the identification of two genes (*embA* and *embB*) which are involved in the polymerization of arabinan in AG (5). Analysis of these *emb* deletion mutants in *M. smegmatis* revealed that the arabinan content of AG was decreased, and in both mutants, the non-reducing Ara_6 motif, [Ara*f*β(1→2)Ara*f*α(1→5)] [Ara*f*β(1→2)Ara*f*α(1→3)] Ara*f*α(1→5) Ara*f*α(1→ which is a template for mycolylation was altered (20). It was shown in a recent study that AftA is the arabinosyltrasferase enzyme that catalyses the addition of the first key Ara*f* residue from DPA on to the

galactan domain of the cell wall, thus 'priming' the galactan for further elaboration by downstream arabinosyltransferases (2, 64). It was also found that the arabinans are added on to the Gal*f* residues at 8, 10 and 12 units of the galactan chain (1). The enzyme which is involved in the α1→5 linkage in *Mycobacterium* sps is still unknown. However in *Corynebacterium glutamicum*, which has only one *emb* gene and was involved in the arabinosylation of α1→5 and α1→3 linkages in AG (1). In a recent structural study of AG, it was reported that it contains three arabinan domains each containing of 31 Ara*f* residues attached to the galactan chain (Gal_{30}) (10). The AG-lipid intermediate (polyprenol-P-P-GlcNAc-Rha-Gal_{30}-Ara_{93}) is then at some point transglycosylated to PG and mycolylated or the non reducing end of AG is first mycolylated and then transglycoslylated on to PG.

The biosynthesis of mycolic acids involves both type I and type II fatty acid synthases (FAS-I and FAS-II). It was found that FAS-I initiates the synthesis of mycolic acids precursors and then elongated by FAS-II system (65). Excellent reviews were published by Takayama et al. and several others on the biosynthesis of mycolic acids (3, 15, 65, 67).

2.1.2.4 Lipoglycans (PIMs, LM amd LAM)

Another important structural component in the mycobacterial cell wall is lipoarabinomannan (LAM), which modulates host immune response (12, 53, 54). LAM is a heterogeneous mixture of molecules in terms of glycosylation and acylation. LAM can be mainly divided into three domains: phophatidylinositol (PI) anchor, arabinomannan and capping domains (Fig. 2.1.3) (8). The PI synthesis in mycobacteria was reported to

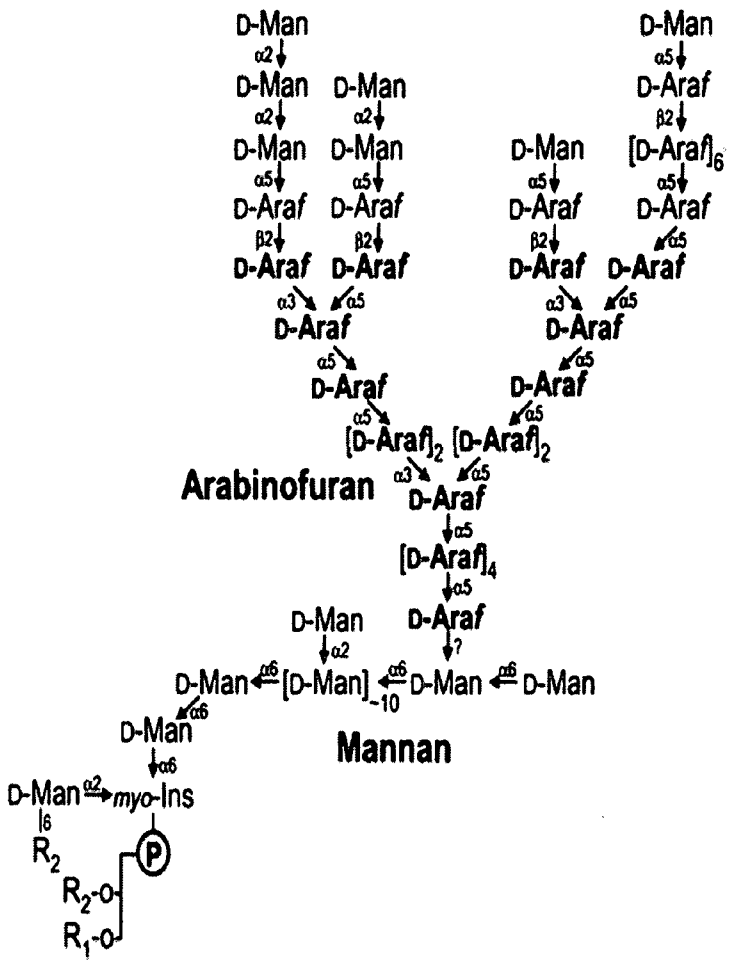


FIGURE 2.1.3 Structure of Lipoarabinomannan (LAM) in *M. tuberculosis* [taken from Berg et al. (8)]

be a *de novo* synthesis, which involves a reaction between cytidinediphosphate - diacylglycerol (CDP-DAG) with free inositol and is catalyzed by CDP-DAG:*myo*-inositol transferase, or PI synthase (*pgsA*, Rv2612c) (24, 58). A α-mannosyltransferase (PimA), catalyzes the reaction in which the Man*p* residue from the GDP-Man*p* is transferred on to the C-2 position of myoinositol in the PI anchor to form PIM_1 (30). The mannose in PIM_1 is acylated by an acyltransferase Rv2611c (31). Recently it was found that PimB' (MSMEG_4253), catalyzes the transfer of a Man*p* residue to the C-6 position of the myoinositol of the PI (22, 35, 49) and the previously identified enzyme (PimB) (60) for this reaction was found to be involved in the biosynthesis of 1,2-di-*O*-$C_{16}/C_{18:1}$-(α-D-Man*p*)-(1→4)-(α-D-Glc*p*-uronic acid)-(1→3)-glycerol ($ManGlcAGroAc_2$) which is a LM like molecule in *Corynebacterium glutamicum* and is now termed *MgtA*. Kremer et al. identified a mannosyltransferase PimC from *M. tuberculosis* CDC1551 strain and showed that this glycosyltransferase mediates the transfer of Man*p* residue from GDP-Man to Ac_3PIM_2 to form Ac_3PIM_3 (34), however homologues for this enzyme are absent in other mycobacterial sps, indicating for the presence of other compensatory genes. The enzyme involved in the conversion of Ac_3PIM_3 to Ac_3PIM_4 is yet to be identified. In formation of polar PIMs recently, a polyprenol monophosphomannose (PPM) dependent mannosyltransferase was shown to be involved in mediating the transfer of fifth α(1→2) Man*p* residue on to the Ac_3PIM_4 (50). Disruption of this gene in *M. smegmatis* led to the accumulation of the tetramannosyl PIM, ($AcPIM_4$) and was defective in $AcPIM_6$ synthesis (50). Then later with the help of an unidentified mannosyltransferase, Ac_3PIM_5 is converted to Ac_3PIM_6, which is the end product of polar PIMs. So the common end point in the biosynthesis of polar PIMs, lipomannan (LM) and LAM is the formation of

Ac_3PIM_4. The subsequent steps in the biosynthesis of LM from Ac_3PIM_4 are not yet clearly known; however it was implicated that PPM dependent mannosyltransferase would be involved in the further steps (8, 51). Recently it was found that Rv2174, a mannosyltransferase is involved in the elongation step in the biosynthesis of LM in the later stages especially from 21-34 Man*p* residues (26). Kaur et al also identified another enzyme (mannosyltransferase, Rv2181) which is responsible for the addition of $\alpha(1\rightarrow2)$ branches to the mannan core of LM (25). This mature branched LM subsequently undergoes arabinosylation to form LAM by a glycosyltransefrase, *embC* (72). The *embC* is found to essential in *M. tuberculosis* (21), whereas it is non-essential in *M. smegmatis* (20, 72). When a point mutation in one of the aspartic acid residues is changed to Alanine (D279A) by site directed mutagenesis in the glycosyltransferase superfamily C (GT-C) motif of EmbC, the enzyme activity is completely lost and arabinan synthesis in LAM is completely abrogated (9). Furthermore, point mutations in three amino acids of the proline motif (W627L, P635S, P641S) of EmbC resulted in marked reduction of arabinan synthesis in LAM and accumulation of an LM (9). Mannose capping (ManLAM) is found to be present in the slow growing mycobacteria like *M. tuberculosis*, *M. leprae*, *M. bovis*, *M. bovis* BCG and *M. avium* whereas phosphoinositide capping (PILAM) is seen in fast growing mycobacteria like *M. smegmatis* and *M. fortuitum* (53). In *M. chelonae* capping was found to be absent, such LAM is called as AraLAM. Dinadayala et al. indentified a mannosyltransferase (Rv1635c) which is involved in addition of the first mannose residue on the non-reducing arabinan end of LAM in the formation of ManLAM (18). Recently it was found that the branching enzyme Rv2181 also adds the 2^{nd} and the 3^{rd}

mannose residues with a α(1→2) linkage of the mannose capping, hence found to have a dual function (27).

2.1.2.5 Hypothesis and Objective

Arabionogalactan and LAM biosynthesis involves several glycosyltransferases (GTFs), several have been discovered recently, but there are more that need to be identified. So in search for the GTFs, we identified one putative GTF (*Rv3805c*) in the cell wall synthesis gene cluster which was described by Belanger and Inamine (4, 68). *Rv3805c* has homology to *ArnT* (arabinosyltransferase) of *Salmonella* and *Escherichia coli* (66). Glycosyltransferases were mainly classified into four super families (GT-A, B, C and D) according to their sequence similarities, [CAZy database (Carbohydrate-Active enZYmes)]. Glycosyltransferases having multiple transmembrane domains and utilizing lipid carriers as sugar donors were grouped under super family C (GT-C). *Rv3805c* belongs to glycosyltransferases super family C (GT-C) and has a conserved DDX, glycosyltransferase motif. During the course of these studies the LAM genes and the GT-C families of GTFs were discovered. A detailed review of these mycobacterial glycosyltransferases (GTF) was reported by Berg et al. (8). So the aim of project is

- To examine if Rv3805c has a role in the cell wall biosynthesis using genetic and biochemical methods.

2.1.3 MATERIALS AND METHODS

2.1.3.1 Bacterial strains and growth conditions

The bacterial strain *E. coli* DH5α used in this study for cloning purpose was grown in Luria-Bertani broth (LB, Difco) at 37°C. *M. smegmatis* mc^2155 was grown in LB broth containing 0.05% Tween80 or on LB agar at 37°C routinely, however based on the experimental requirements *M. smegmatis* was grown at 30° and 42°C. Similarly *Corynebacterium glutamicum* ATCC 13032 strain was normally grown at 30°C in LB broth or brain heart infusion media with sorbitol (BHIS) (19), and when required *C. glutamicum* was grown at 30° and 37°C. When required the antibiotics were added at the following concentrations: 50 μg/ml for hygromycin (Hyg) and kanamycin (Kan), 100 μg/ml for ampicillin (Amp), 5 μg/ml for gentamicin (Gen) and 25 μg/ ml for streptomycin (Strep) for *E. coli* and *M. smegmatis*; 25 μg/ml for Hyg and Kan for *C. glutamicum*. Sucrose was added at a final concentration of 10% and 2% to select for allelic exchange mutants in *M. smegmatis* and *C. glutamicum* respectively.

2.1.3.2 Construction of over-expressing plasmids

The full length *Rv3805c* gene was amplified by PCR using *M. tuberculosis* H37Rv DNA as a template. Two versions were generated by using a common forward primer but with different reverse primers (one reverse primer includes the stop codon of *Rv3805c* and the other does not). The primers contain NdeI and HindIII nucleotide sequences on the forward and reverse primers respectively (Table 2.1.1). This enable us to obtain both 'N' and 'C' terminal six-Histidine tag when cloned into the NdeI and HindIII restriction sites of mycobacterial expression vectors, pVV2 and pVV16 respectively (17). PCR products were initially subcloned into a pGEM T Easy vector (Promega) to obtain pGEMT3805N and pGEMT3805C. The inserts were then verified by sequencing, released by restriction digested with NdeI and HindIII to obtain a 1.8kb

TABLE 2.1.1 Primers used in this study.

Primer name	Oligonucleotide sequence[a]
Rv3805 Fwd-NdeI	GTC*CATATG*GTCCGGGTCAGCTTGT
Rv3805N Rev-HindIII	GTA*AAGCTT*CTCCCGCGGTGGCGGGC
Rv3805C Rev-HindIII	CTC*AAGCTT*GTGATCTTGCTGTCAGTA
ML0096 Fwd-NdeI	TAT*CATATG*CGGCGCCGAATGGTGAGC
ML0096 Rev-HindIII	*AAGCTT*TCATTCGGGCGGGGTAGCA
ML0096 Rev-HindIII	*AAGCTT*ATCGTCGCCGAGTGCTTCGG
MSMEG_6400 Fwd-NdeI	*CATATG*GTGGACATCGACGGGGGCAT
MSMEG_6400 Rev-HindIII	*AAGCTT*CTACGGTCCCGTTGCCGGCAT
MSMEG_6400 Fwd Chromosomal	GCAAGTCGCTGGAGAGTTACAC
MSMEG_6400 Rev Chromosomal	GAGTACACGTCGAGGTACTCGA
MSMEG_6400 Fwd Probe	ACTTTCCGTGCTGGTGCG
MSMEG_6400 Rev Probe	GAGCAGCACCAGCAGCAC
Kan^r Fwd	CATGGATGCTGATTTATATGGGTA
Kan^r Rev	GTTGATGAGAGCTTTGTTGTAGGT
NCgl2780 Fwd-SmaI	TAT*CCCGGG*ATGACGTTTAGCCCCCAGC
NCgl2780Rev-HindIII	GCG*AAGCTT*TTACTGAGAGCTATATAAAGG
Hyg^r Fwd	*TCTAGA*GTAAACTTGGTCTGACAG
Hyg^r Rev	*AAGCTT*TCAGGCGCCGGGGGCGGT

[a] Nucleotides in a restriction site are underlined and italicized.

fragment which was then ligated into NdeI and HindIII sites of pVV2 and pVV16. The resultant plasmids pVV2:3805N and pVV16:3805C were transferred by electroporation into *M. smegmatis*. The transformants selected on plates containing Kan and Hyg and will be hereafter known as TB1 and TB2 respectively (Table 2.2). Similarly full length *ML0096* gene was amplified by PCR from *M. leprae* genomic DNA (obtained from CSU Leprosy contract) using the primers listed in Table 2.1, cloned into pVV2 and pVV16 to obtain pVV2:ML0096N and pVV16:ML0096C and then transformed into *M. smegmatis*. The recombinant strains hereafter will be known as ML1 and ML2 respectively (Table 2.1.2).

2.1.3.3 Construction of *M. smegmatis MSMEG_6400* conditional mutant

MSMEG_6400, the *M. smegmatis* gene corresponding of Rv3805c was amplified by PCR with the primers listed in Table 2.1.1. The 2.076 kbp PCR fragment was subcloned into pGEM T Easy vector to obtain pGEMTMS6400 and then the cloned fragment was verified by sequencing. The plasmid pGEMTMS6400 was digested with BglII and then ligated with a Kan resistance (Kan^r) gene (1264 bp) which was excised from pUC4K by BamHI digestion to obtain pGEMTMS6400::Kan^r. Then the 3.34 kb fragment of MS6400::Kan^r was excised from pGEMTMS6400::Kan^r plasmid with NotI and SpeI, and inserted into the same sites of pPR27-x*ylE* (conditional replication plasmid) (55), yielding pPR27-*x*ylE- MS6400::Kan^r (pSRM1) (Table 2.1.2). Thus the plasmid pSRM1 contains mycobacterial temperature sensitive origin of replication, gentamicin resistance (Gen^r), levansucrase counter selectable marker (*sacB*), catechol-dioxygenase reporter gene (*xylE*) and kanamycin resistance (Kan^r) gene.

TABLE 2.1.2 Bacterial strains and plasmids used in this study.

Plasmids /Strains	Characteristics	Sources/ reference
Plasmids		
pGEM T EASY	Carries Amp^r gene and used for cloning PCR product with a 'A' extension	Promega
pVV2	*E.coli*/mycobacterial shuttle plasmid originated from pMV61 with Kan^r and Hyg^r genes, mycobacterial expression plasmid under hsp60 promoter (Phsp60) and has N terminal six histidine tag.	(7)
pVV16	*E.coli*/mycobacterial shuttle plasmid originated from pMV61 with Kan^r and Hyg^r genes, mycobacterial expression plasmid under hsp60 promoter (Phsp60) and has C terminal six histidine tag.	(7)
pVV2:3805N	pVV2 plasmid carrying Rv3805c with stop codon with an inframe N terminal six histidine fusion tag	This study
pVV16:3805C	pVV16 plasmid carrying Rv3805c without stop codon and has C terminal six Histidine fusion tag	This study
pVV2:ML0096N	pVV2 plasmid carrying ML0096 with stop codon with an inframe N terminal six Histidine tag	This study
pVV16:ML0096C	pVV16 plasmid carrying ML0096 without stop codon and has C terminal six Histidine tag	This study
pPR27-xylE	Carries *E. coli* origin of replication, thermosensitive mycobacterial origin of replication, has *sacB* and *xylE* genes, Gen^r gene	(55)
pUC4k	Carries both Amp^r and Kan^r genes	Amersham biosciences
SRM1	pPR27-*xylE* carrying MSMEG_6400::Kan^r	This study
pET23b-Phsp60	Carries Amp^r gene, hsp60 promoter (Phsp60) exercised from pVV16 and cloned into XbaI and NdeI sites of pET23b	(37-39)
pCG76	Carries *E. coli* origin of replication, thermosensitive mycobacterial origin of replication, $Strep^r$ gene	(55)
SRM2	pCG76carrying MSMEG_6400	This study
pSFKT2	Carries *E. coli* origin of replication, thermosensitive corynebacterial origin of replication	(52)
pPS160-10 (pALB2)	Carries Amp^r and *sacB* genes	(62)
pSFKT2-sacB	pSFKT2 carrying *sacB* gene	This study
pSRM3	pSFKT2-sacB carrying NCgl2780::Hyg^r	This study
Strains		
E. coli DH5α	Used for cloning and plasmid propagation	Invitrogen
M. smegmatis mc^2 155	Non-pathogenic mycobacterial sps	ATCC
C. glutamicum 13032	Non-pathogenic corynebacterial sps	ATCC
TB1	*M. smegmatis* mc^2 155 with pVV2:3805N	This study
TB2	*M. smegmatis* mc^2 155 with pVV16:3805C	This study
ML1	*M. smegmatis* mc^2 155 with pVV2:ML0096N	This study
ML2	*M. smegmatis* mc^2 155 with pVV16:ML0096C	This study
Δ*MSMEG_6400* (RM1)	*M. smegmatis* mc^2 155 which has undergone a second cross over event at locus MSMEG_6400 in the presence of pSRM3	This study
Δ*NCgl2780* (RM2)	*C. glutamicum* 13032 which has undergone a second cross over deletion event at locus NCgl2780	This study

The plasmids pPR27-x*ylE* (control) and pSRM1 were electroporated into *M. smegmatis*, and transformants were selected on LB Kan and LB Kan + Gen agar plates at 30°C respectively. Transformants obtained from the plasmids pPR27-x*ylE* and pSRM1 were inoculated into LB Gen and LB Kan + Gen broth respectively and incubated at 30°C, which is a permissive temperature for plasmid replication. Then the cells were plated on LB Gen and LB Kan + Gen agar plates and incubated at 42°C, a nonpermissive temperature for plasmid replication and at which the plasmid pSRM1 should integrate at the *MSMEG_6400* gene region of *M. smegmatis* genome to obtain a legitimate single crossover (SCO) recombinant event. Five putative SCO colonies were selected and analyzed by PCR and Southern blot analysis of chromosomal DNA. After conformation of SCO event, the SCOs were plated on LB Kan and 10% sucrose plates to select for mutants that undergo second homologous recombination event resulting in an allelic exchange of the *MSMEG_6400* gene with the disrupted Kan^r gene.

For the construction of a rescue plasmid, pGEMTMS6400 was digested with NdeI and HindIII to obtain the full length *MSMEG_6400* gene, which was then ligated into the same sites of pET23b-Phsp60 to obtain pET23b-Phsp60-MS6400 (37-39). The plasmid pET23b-Phsp60-MS6400 was digested with XbaI and HindIII to get a Phsp60-MS6400 fragment and filled in by klenow enzyme (New England Biolabs). The plasmid pCG76, which has a mycobacterial thermosensitive origin of replication and Streptomycin resistance (Strep^r) was digested with BamHI and also filled using klenow enzyme. The Phsp60-MS6400 fragment was ligated into the linear pCG76 to get a rescue plasmid [pCG76-Phsp60-MS6400 (pSRM2)] with a functional copy of *MSMEG_6400* gene (Table 2.1.2).

The plasmids pCG76 (control) and pSRM2 were electroporated into SCOs and were selected on LB Strep and LB Kan + Strep agar plates respectively and incubated at 30°C. Transformants carrying the plasmids pCG76 and pSRM2 were inoculated into LB Strep and LB Kan + Strep broth respectively and incubated at 30°C. Then the cells were plated on LB Strep and LB Kan + Strep agar plates containing 10% sucrose and incubated at 30°C.

M. smegmatis genomic DNA was isolated as described before (37, 38). A 200bp region in the middle of *MSMEG_6400* gene (100bp on either side of BglII, where the Kan^r cassette was inserted) was amplified using the primers (Table 2.1.1) by PCR and was used as a DNA probe. The DNA probe was labeled with digoxigenin, and Southern blot analyses were performed as described for DIG High Prime DNA labeling and detection starter kit I (Roche).

2.1.3.4 Genomic southern analysis

The *M. smegmatis* wild type strain (WT), SCO and Δ*MSMEG_6400* conditional mutants (double cross over, DCO) were inoculated in 20 ml LB broth (Lennox) containing 0.05% Tween 80 with appropriate antibiotics and incubated at both 30°C and 42°C. The growth of cells (OD_{600} value) was measured at regular intervals.

2.1.3.5 Construction of *C. glutamicum NCgl2780* knock out mutant

NCgl2780, the full length *C. glutamicum* gene was amplified by PCR with the primers listed in Table 2.1. The 2.061 Kbp PCR fragment was subcloned into pGEM T Easy vector to obtain pGEMTCG2780 and then the cloned fragment was verified by sequencing. The plasmid pGEMTCG2780 was digested with KpnI and BamHI, a 350bp

fragment was discarded. The Hyg resistance (Hyg^r) gene (1300 bp) was amplified from pVV16 by using primers in Table 2.1. Then both the linear plasmid pGEMTCG2780 obtained from KpnI and BamHI digestion and the PCR product of Hyg^r gene was ligated to obtain pGEMTCG2780:: Hyg^r.

The *Corynebacterium/E. coli* shuttle plasmid, pSFKT2 contains corynebacterial thermosensitive origin of replication, thus it can replicate at 25°C but not ≥34°C and was a gift from Dr. Nakamura, Fermentation and Biotechnology Laboratories, Ajinomoto CO., Inc., Kanagawa, Japan (52). The plasmid pALB2 (62) (obtained from Dr. Herbert Schweizer, Dept. of Microbiology, Immunology and Pathology, CSU, Fort Collins, USA) was digested with XbaI and SphI to get a 2.6 Kbp fragment of *sacB* counter selectable marker. Over hangs were blunt-ended with klenow enzyme and then ligated in the StuI site of pSFKT2 to obtain pSFKT2sacB (Table 2.1.2).

The 3.3 kb fragment of CG2780:: Hyg^r was excised from pGEMTCG2780:: Hyg^r plasmid with SmaI and HindIII, and inserted into the same sites of pSFKT2sacB to get pSFKT2sacB- CG2780:: Hyg^r (pSRM3) (55) (Table 2.1.2). Thus the plasmid pSRM3 contains corynebacterial thermosensitive origin of replication, levansucrase counter selectable marker (*sacB*), Hyg^r and Kan^r genes. Electrocompetent *C. glutamicum* cells were prepared from the protocol described (19). The plasmids pSFKT2sacB (empty vector) and pSRM3 were electroporated to *C. glutamicum* cells by setting the voltage of Gene Pulser (Bio-Rad) to 2.5 kV, capacity to 25 μF, and resistance to 200 Ω. The cells were transferred immediately into 4 ml prewarmed BHIS medium and incubated for 5 min at 46°C, then allowed to regenerate for 1 hr at 25°C and then selected on BHIS agar plate containing Kan and Kan + Hyg at 25°C. Transformants obtained from the plasmids

pSFKT2sacB and pSRM3 were inoculated into BHIS medium with Kan and Kan + Hyg broth respectively and incubated at 35°C until midlog phage. Then the cells were plated on BHIS agar plates with or without 2% sucrose, and incubated at 35°C. The allelic replacement of *NCgl2780* with the disrupted copy (*NCgl2780::*Hyg^r) in the chromosomal DNA of the DCOs was confirmed by PCR and Southern blot analysis. The full length *NCgl2780* gene PCR product was used as a DNA probe for Southern hybridization. The DNA probe labeling and Southern blot analyses were performed as described above (Section 2.1.3.4).

2.1.3.6 Arabinosyltransferase assay

M. smegmatis WT and the other overexpression strains like TB1, TB2, ML1 and ML2 were grown in 2L LB broth to mid-log phase (OD 0.5–0.7). Cells were harvested by centrifugation, washed, and resuspended in buffer B containing 50 mM MOPS (pH 8.0), 5 mM β-mercaptoethanol and 10 mM $MgCl_2$. Cells were disrupted in six cycles by French press at 1500 psi. The cell lysate was centrifuged at 27,000 × g for 60 min at 4°C. The supernatant of 27,000 × g was recentrifuged at 100,000 × g for 2 hr at 4°C. The 100,000 × g supernatant was discarded and the pellet (membrane fraction) was resuspended in ~300 µl of buffer B. The 27,000 × g pellet was suspended in 10 ml of buffer B and Percoll (Thermo Scientific Pierce) (~15ml) was added to achieve a 60% suspension. The suspension was mixed and centrifuged at 27,000 × g for 60 min at 4°C. The particulate upper white layer was collected and washed three times with buffer B at 27,000 × g and finally resuspended in ~500 µl to of buffer B to obtain a cell wall enriched

fraction (P60) (28, 36). Protein concentration of the membrane and the P60 fraction were measured by bicinchoninic acid (BCA) assay (Thermo Scientific Pierce).

The precursor ^{14}C labeled donor of arabinofuranose (Ara*f*) i.e., phosphoribose pyrophosphate (p[14C]Rpp) was prepared by the procedure the described elsewhere (61). In the arabinosyltransferase assay reaction mixture contained: acceptor (0.3mM), ATP (1 mM), p[^{14}C]Rpp (500,000 dpm), membrane fraction (0.5 mg) and P60 fraction (0.3 mg) in a total volume of 200 µl. Different types of acceptors were used in the arabinosyltransferase assay. These acceptors and the p[^{14}C]Rpp were obtained from Dr. Chatterjee (Dept. of Microbiology, Immunology and Pathology, CSU, Fort Collins, USA) (28, 71). The reaction mixtures were incubated at 37°C for 2 hr and then terminated by adding 200 µl of ethanol. The resulting mixture was centrifuged at 14,000 × g, and the supernatants were loaded onto prepacked strong anion exchange (SAX) columns (Burdick and Jackson). The columns were eluted sequentially with 1 ml of 50% ethanol and then 1 ml of water. The eluate was evaporated to dryness and partitioned between the two phases (1:1) of water saturated butanol and water. The butanol fractions were measured for radioactive incorporation by liquid scintillation counter. Equal volume of the radiolabeled material from the WT and other overexpression strains were subjected to thin layer chromatography (TLC) analysis using silica gel plates (Merck) developed in $CHCl_3$:CH_3OH:1 M NH_4OAc:NH_4OH:H_2O (180:140:9:9:23) solvent system (28, 71). Autoradiograms of the TLC plates were obtained by exposure to X-ray film. The original acceptor compound was visualized using α-naphthol spray reagent.

For sugar composition analysis of the radiolabeled product 5000 dpm of the butanol fraction was dried and hydrolyzed in 200 µl of 2 M trifluoroacetic acid (TFA) at 120°C for 2 hr. The TFA was removed by drying, and the hydrolysate was analyzed on a TLC plate developed in pyridine:ethyl acetate:acetic acid:water (5:5:1:3) solvent system followed by autoradiography. Radioactive spots were identified by visualizing against the standard sugars which were developed using α-naphthol spray.

2.1.3.7 Whole cell radiolabeling

For whole cell radiolabeling experiments of the *MSMEG_6400* conditional mutant, several conditions were tested. Initially the conditional mutant and the *M. smegmatis* WT strains were grown at 30°C in LB Lennox medium with 0.05% Tween 80 and Kan until midlog phase (OD_{600} = 0.5). This is used as a starter culture. The midlog culture was inoculated (1/200) into two flasks containing fresh medium (100ml) and ^{14}C-glucose, one flask was placed at 30°C and the other at 42°C. Since no growth at 42°C culture, another experiment condition was designed, in which the midlog culture was inoculated into two flasks of fresh LB Lennox medium. Both the cultures were incubated at 30°C until the cell growth reached an OD_{600} value of 0.05 (~ 8 hr). Then one flask was transferred to 42°C, while the other remained at 30°C. ^{14}C-glucose was added into each of the flasks at this point. Twenty milliliter from 30°C and 42°C cultures were collected at different time periods. Cells were harvested and washed with phosphate-buffered saline (PBS). Lipids were extracted sequential by using non-polar and polar solvent systems like CH_3Cl:CH_3OH (2:1), and CH_3Cl:CH_3OH:H_2O (10:10:3). Later LAM and AG were extracted from the remaining pellet (24, 25, 30, 47, 72). The labeled and the unlabeled

lipids were resolved by TLC on different solvent systems. LAM was analyzed by SDS-PAGE or on commercially available Tricine SDS-PAGE gels (Invitrogen) followed by periodic acid Schiff staining (57). The radiolabeled LAM was analyzed by SDS-PAGE, electropheretically transferred onto nitrocellulose membrane and then the nitrocellulose membrane was exposed to an X-ray film for autoradiography.

2.1.3.8 Digestion with endoarabinanase

The *Cellulomonas* endoarabinanase digestion pattern was found to be specific. The Ara_6, Ara_2 and cyclic Gal_4 products were obtained from endoarabinanase digestion of AG, whereas Ara_6, Ara_4 and Ara_2 products were obtained of LAM endoarabinanase digestion (43, 72). Radiolabeled products from the *in vitro* arabinosyltransferase assay was digested with endoarabinanase (43) and the digestion products were analyzed on TLC in pyridine:ethyl acetate:acetic acid:water (5:5:1:3) solvent system followed by autoradiography. The unlabeled AG from different strains was also digested with endoarabinanase and the digestion products were analyzed on high pH anion exchange chromatography (HPAEC, Dionex) as described previously (13, 72).

2.1.4 RESULTS

2.1.4.1 Identification of *Rv3805c* as a putative glycosyltransferase by bioinformatics

Rv3805c was identified by insilico genome analysis of conserved cell wall synthesis gene cluster (Fig. 2.1.4) (4). *Rv3805c* is located between the gene *Rv3806c* which is involved in the synthesis of decaprenylphosphoryl-D-arabinose (DPA), (donor of arabinose in mycobacteria) (23, 69) and *fbpA* which is involved in addition with

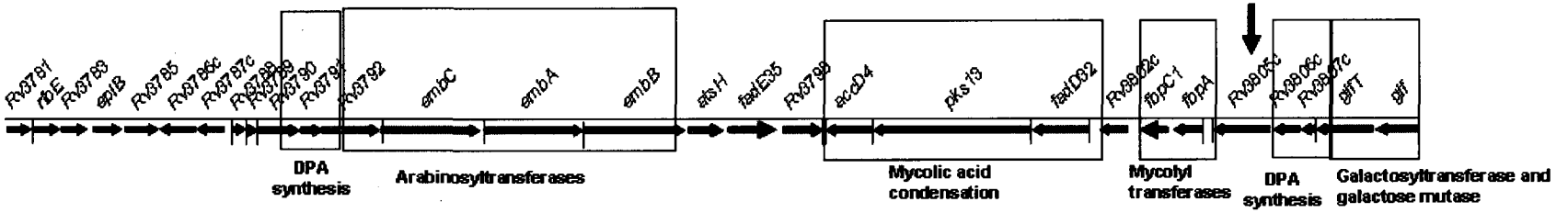


FIGURE 2.1.4 Cell wall biosynthetic gene cluster of *M. tuberculosis* as described by Belanger and Inamine (4). [taken from Vissa and Brennan (65)]. The location of *Rv3805c* is indicated by a vertical arrow. Genes that were characterized in the synthesis of various cell wall macromolecules when this study was initiated are shown in boxes.

mycolic acids on the terminal and penultimate arabinose residues of hexa-Ara (Ara_6) motif of arabinogalactan (AG) of mycobacteria (1, 7, 20). Based on the NCBI BLAST algorithm and conserved domain analysis, *Rv3805c* was found to have some similarity with the enzyme ArnT that catalyzes the transfer of arabinose from decaprenyl-phosphoryl 4-amino 4-deoxy arabinose, in Kdo-lipid A synthesis in *E. coli* and *S. typhimurium* (66). With the homology searches Rv3805c was found to be conserved in mycobacteria and corynebacteria (Fig. 2.1.5). The predicted topology of Rv3805c has ten N terminal transmembrane segments followed by a C terminal hydrophilic domain as seen in EmbA, EmbB and EmbC proteins (Fig. 2.1.6A). Sequence alignments of *Rv3805c* with Emb proteins, and other bacterial proteins revealed presence of aspartic acid containing motif (DDX) in the first cytoplasmic loop which is a signature of GT-C superfamily of glycosyltransferases (8) and conserved prolines within a proline rich motif which was shown to be present in Emb proteins (Fig. 2.1.6B) (9). Taken together *Rv3805c* was identified as a putative arabionosyltransferase, involved in LAM and/or AG biosynthesis.

2.1.4.2 *in vitro* arabinosyltransferase assay using recombinant overexpression strains

We attempted to express Rv3805c and ML009c using mycobacterial expression vectors pVV2 and pVV16 in *M. smegmatis*. We could only detect the recombinant Rv3805c from TB1 and TB2 strains on a western dot blot analysis with anti-His monoclonal antibody, but not in the WT or *M. smegmatis* with pVV16 (data not shown). This is probably due to the highly hydrophobic nature of the protein and limited expression. However, we utilized these recombinant strains (TB1, TB2, ML1 and ML2)

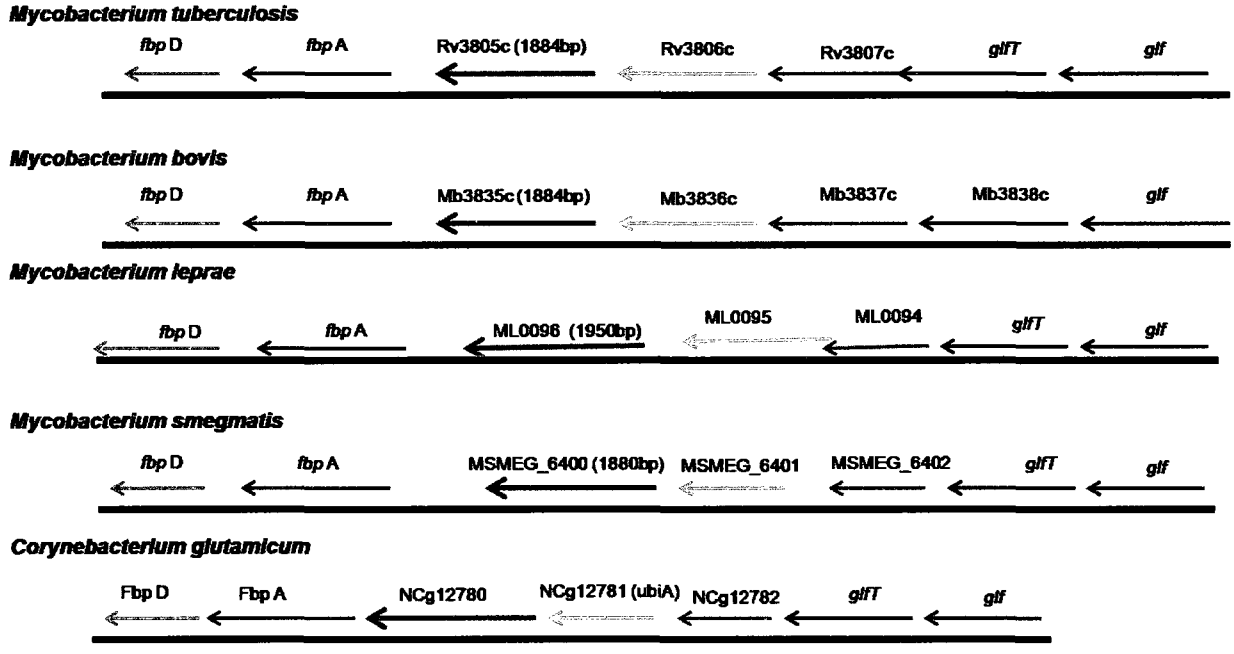


FIGURE 2.1.5 Genetic organization of *Rv3805c* orthologs in mycobacteria and corynebacteria

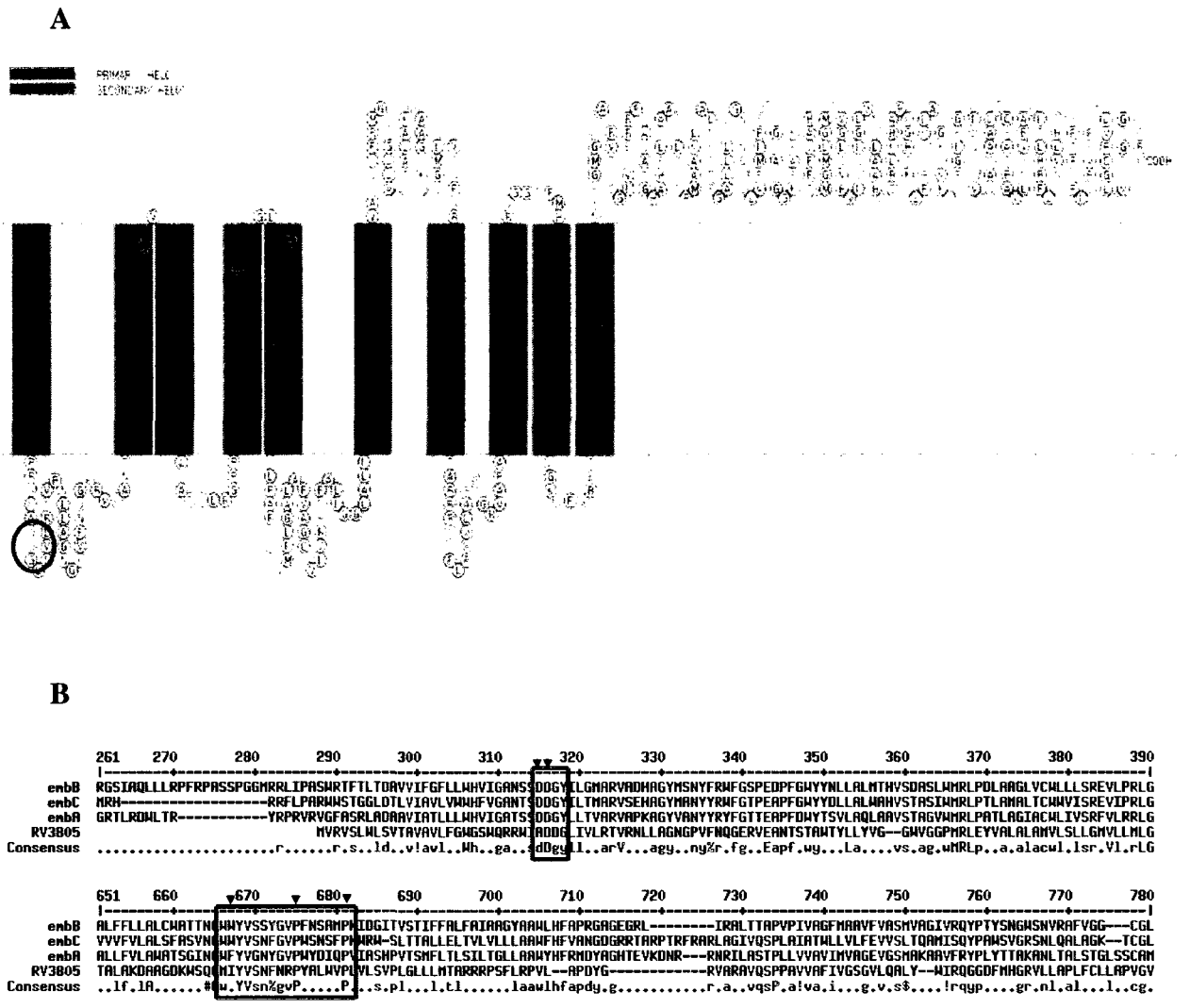


FIGURE 2.1.6 Membrane topology and multiple sequence alignment of Rv3805c protein and Emb proteins. **A**.Topology of Rv3805c as predicted by SOSUI software and the circle indicates the GT-C glycosyltransferase motif (DDX). **B**. Sequence alignment of Rv3805c and EMB proteins by Multialign software, indicates that the GT-C and the proline rich motif (shown in the red box) are shared characteristic feature. The amino acid positions were shown on the top of the dotted line.

and their vector controls (*M. smegmatis* WT with pVV2/pVV16) in a previously developed arabinosyltransferase assay (28, 36). We utilized four different acceptors namely α-D-Ara*f*-(1→5)-α-D-Ara*f*-O-C_8H_{17} (Di-Ara-octyl), α-D-Ara*f*-(1→5)-α-D-Ara*f*-O-C_5H_{10} (Di-Ara-pentenyl), α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O-C_5H_{10} (Tri-Ara-pentenyl), α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-O-C_5H_{10} (Tetra-Ara-pentenyl) and p[^{14}C]Rpp as the Ara*f* donor (Fig. 2.1.7). In the butanol fractions the radiolabeled products were 2-3 fold higher with the overexpressed Rv3805c enzyme (TB2 strain), when compared to the control strain (Fig. 2.1.8A). Similar results were obtained using other overexpression strains like TB1, ML1 and ML2 (data not shown). When the butanol layer products were analyzed by thin layer chromatography (TLC), it was observed that radiolabeled products were not formed in the reactions with out synthetic acceptor (lanes 1 and 2 in Fig. 2.1.8B). TLC and autoradiography analysis of the radiolabeled products obtained from different acceptors indicate that only a single [^{14}C]Ara residue was transferred to each acceptor based on the retention factor (R*f*) (Fig. 2.1.8 C). Since membrane and cell wall fractions are used as enzyme source and p[^{14}C]Rpp as the Ara*f* donor; endogenous products were also formed and eluted along with the acceptor dependent radiolabeled products, so some non-specific bands were also observed (Fig. 2.1.8 C).

Acid hydrolysis followed by TLC against sugar standards established that the radiolabeled products (butanol fractions) contain [^{14}C]Ara which originates from p[14C]Rpp as shown by (Fig. 2.1.9A). Endoarabinanase digestion of the radiolabeled products (equal volumes) followed by TLC analysis with Di-Ara-pentenyl acceptor and enzyme fractions from both control and TB2 strains resulted in two products. One of the

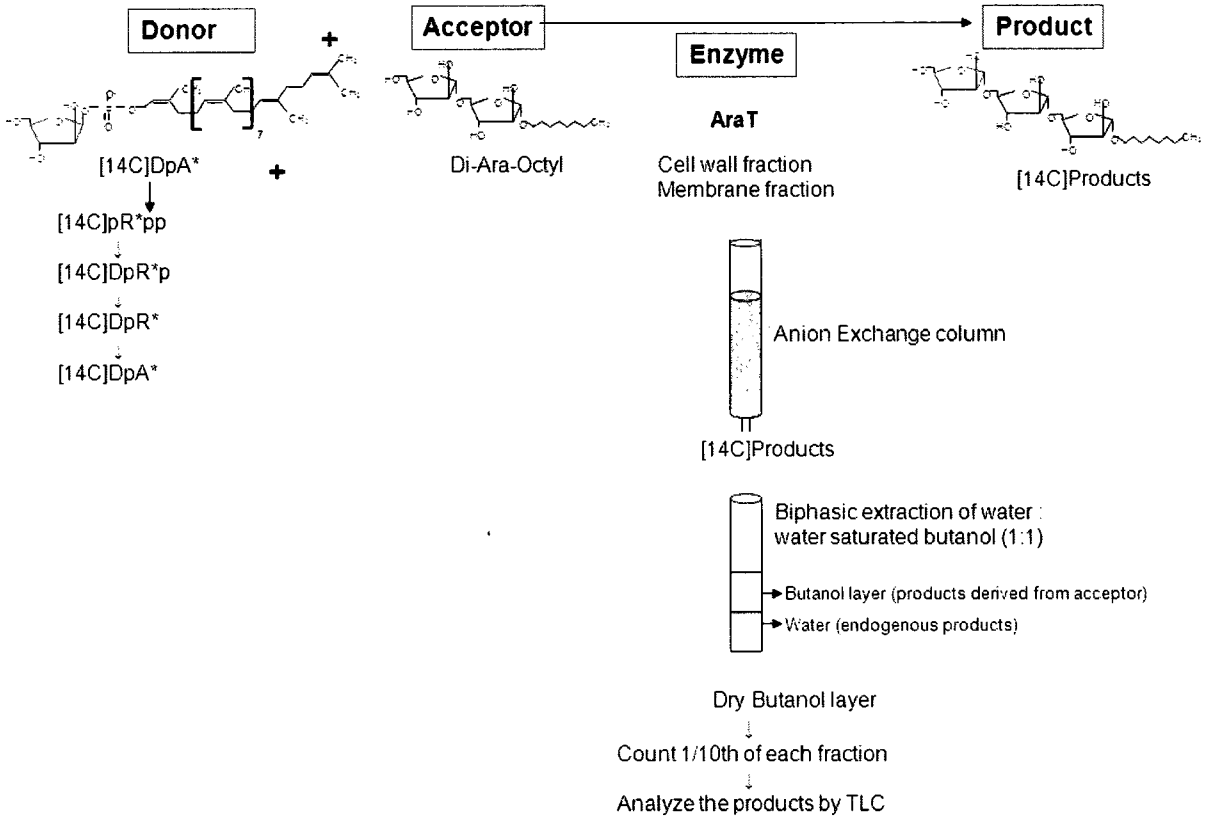


FIGURE 2.1.7 Schematic diagram showing the arabinosyltransferase assay with various synthetic arabinose acceptors.

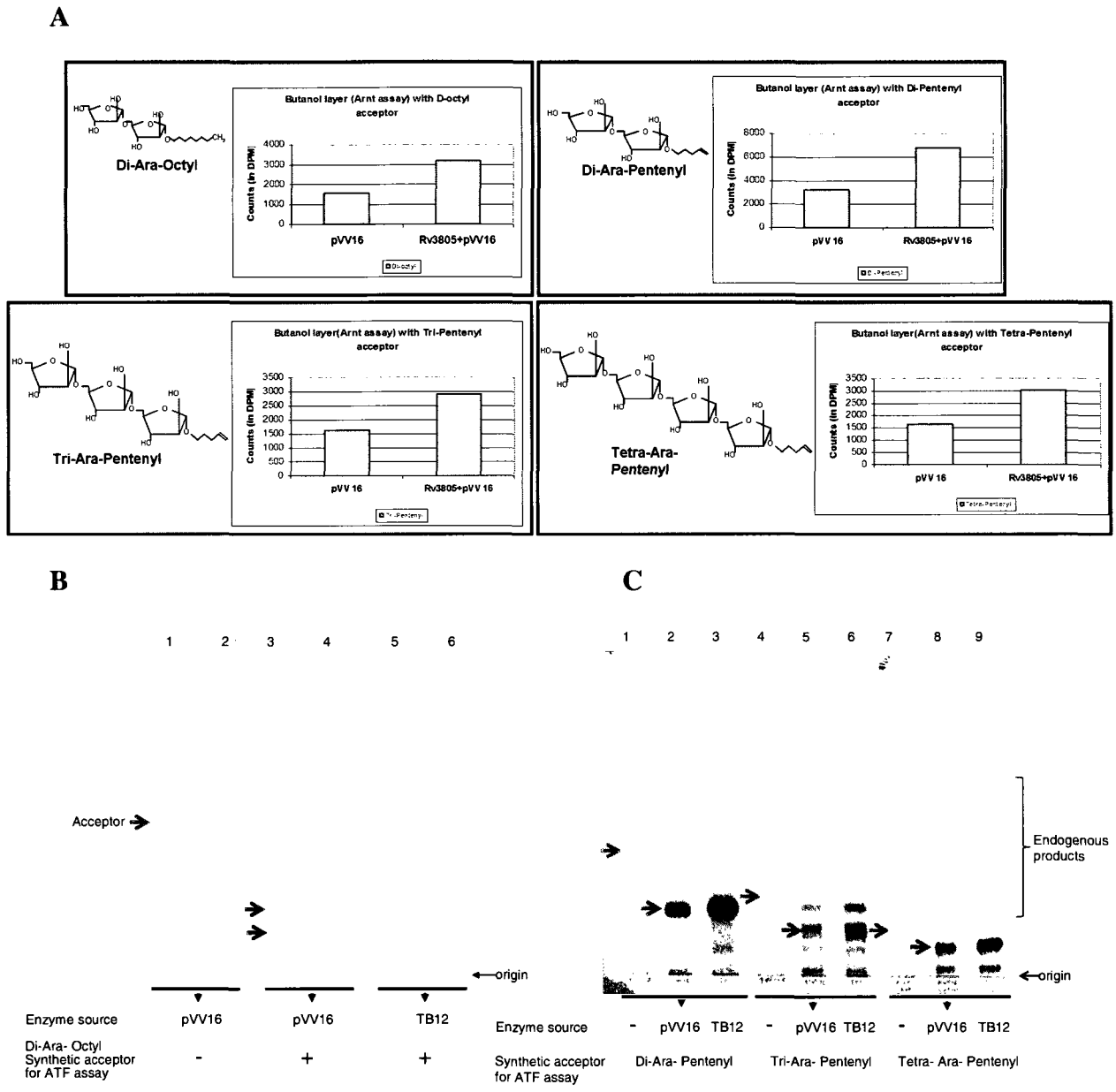


FIGURE 2.1.8 Effect of overexpression of Rv3805c in the arabinosyltransferase assay. **A.** Arabinosyltransferase activity of the overexpressed Rv3805c when compared to pVV16 (control) with various synthetic arabinose acceptors (radioactive products measured in dpm) **B and C**. TLC and autoradiography analysis of the radioactive products obtained with overexpressed Rv3805c (TB12) and pVV16 control using various synthetic arabinose acceptors in the arabinosyltranseferase assay. The acceptor in each assay is indicated by green arrow, while the products were indicated in red colored arrows. The original acceptors are visualized by spraying (lanes 1, 4 and 7 in panel C) with α-napthol.

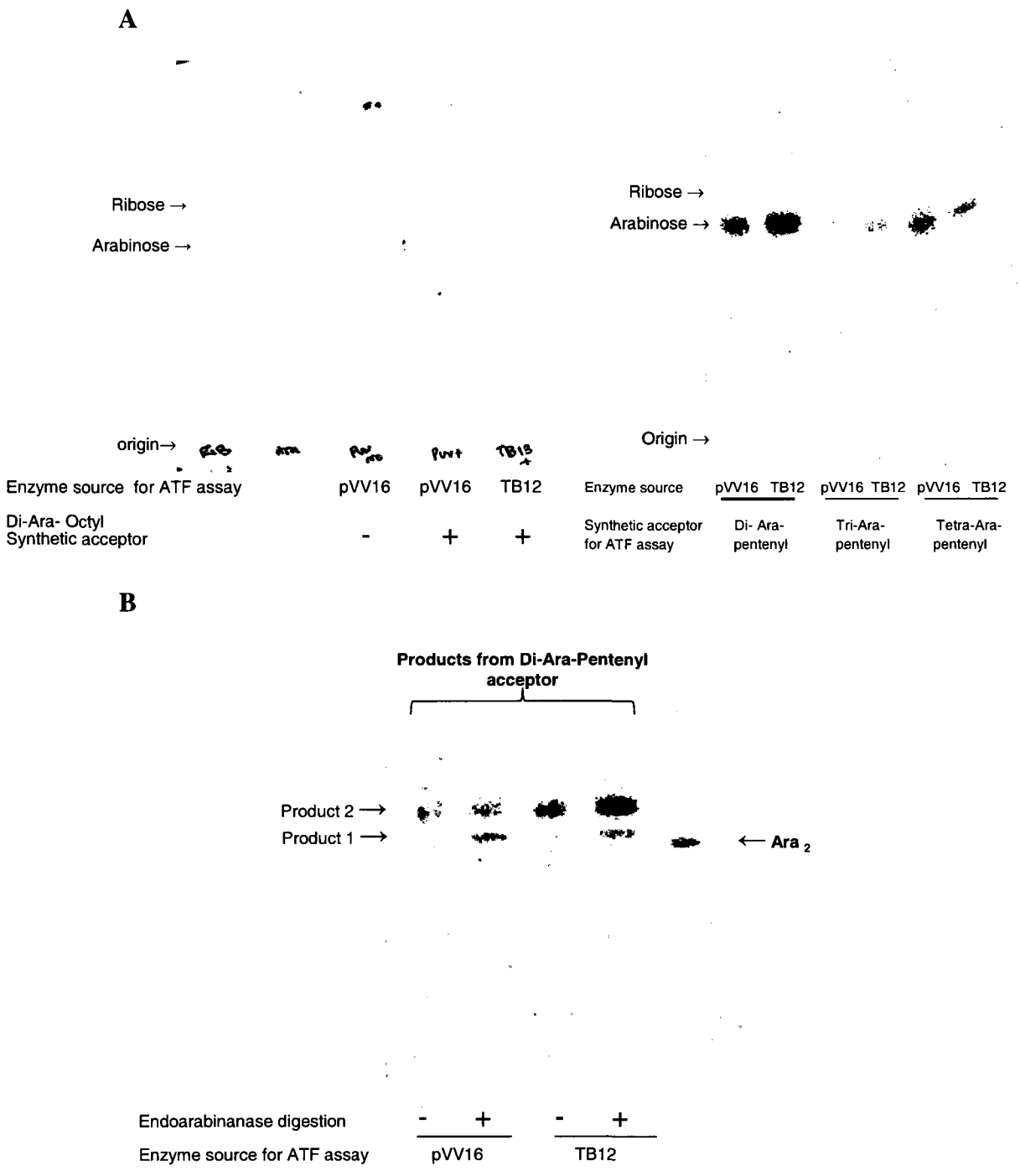


FIGURE 2.1.9 Characterization of the products (butanol layer fraction) from the arabinosyltransferase assay. **A**. Acid hydrolysis (2 M TFA) of the radioactive products obtained from arabinosyltransferase assay and applied to TLC plate and subjected to autoradiogram for monosaccharide analysis. Neutral sugars are visualized by spraying α-napthol. **B**. Radioactive products obtained from arabinosyltransferase assay were digested with endoarabinanase and the digested products are applied to TLC plate and followed by autoradiography.

products (product 1) with equal intensities was comigrating with Ara_2 standard. The product 2 is the original product formed from arabinosyltransferase assay which is resistant to endoarabinanase digestion. Product 2 was found to be higher in TB12 compared to that of control (pVV16) (Fig. 2.1.9B). It was known from the earlier studies that digestion pattern of endoarabinanase is specific and cuts the (1→5) linear arabinan yielding to Ara_2 units (43). Therefore the product 1 could be [^{14}C]Ara-(1→5)-Ara (similar R*f* of Ara_2 standard) and the endoarabinanase resistant product 2 could be either [^{14}C]Ara-(1→3)-Ara-(1→5)-Ara-OC_5H_{10} or [^{14}C]Ara-(1→2)-Ara-(1→5)-Ara-OC_5H_{10}. Moreever from earlier studies by GC-MS analysis, it was found that an Ara_3 [Ara-(1→2)-Ara-(1→5)-Ara (major product) and Ara-(1→5)-Ara-(1→5)-Ara (minor product)] was formed from Di-Ara acceptors (36). Therefore from these results it can be concluded that Rv3805c is an arabinosyltransfearse with (1→2) activity.

2.1.4.3 Phenotypic analysis of *MSMEG_6400* conditional mutant

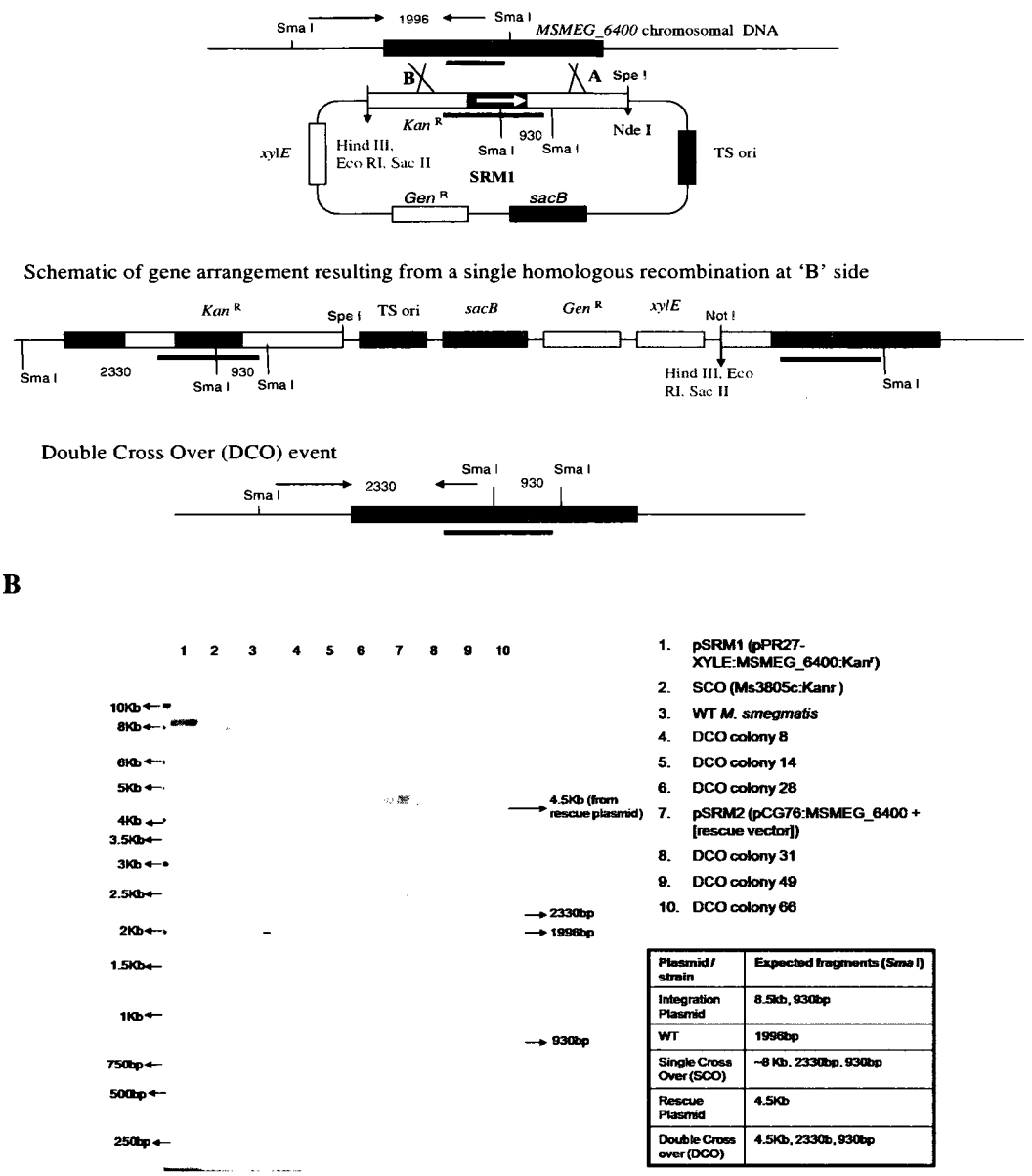
To determine the *in vivo* role of Rv3805c in AG and/or LAM biosynthesis, attempts to generate a genetic knock out mutant of *MSMEG_6400* failed and it was concluded that *MSMEG_6400* is essential to *M. smegmatis*. Therefore it was necessary to construct a conditional mutant. The experimental approach includes a traditional two step homologous recombination procedure (55). In the first step, a conditional replication plasmid (pSRM1) was constructed. The plasmids pPR27-*xylE* (vector control) and pSRM1 were electroporated into *M. smegmatis* and then transformants selected at 30°C on Gen^r and Kan^r + Gen^r respectively. When the transformants were grown on 42°C, colonies were not observed in the control plates. Hence, the Kan^r + Gen^r and XylE positive colonies from the pSRM1 plasmid selected at 42°C were candidates in which

single cross over (SCO) may have occurred. Southern blot analysis on the nine SCO candidates indicated that seven resulted from homologous recombination with gene arrangement from B event (data not shown), while the other two arose from illegitimate recombination. SCOs were grown in LB kan + Gen broth and then plated on LB Kan and 10% sucrose plates to select for mutants that would undergo intrachromosomal allelic exchange. These should be able to grow on sucrose, Kan and appear as white colonies (x*ylE* negative) when sprayed with catechol. However there were thousands of sucrose resistant colonies but none of them were white. All of them were yellow colonies (*xylE* positive) indicating *sacB* mutants, so this suggested that *MSMEG_6400* gene is essential for *M. smegmatis*.

In the second step, the vector control (pCG76) and the rescue plasmid (pSRM2) were transformed into *M. smegmatis*, transformants were selected at 30°C on the LB plates containing 10% sucrose with Strep and Kan + Strep respectively. All the colonies in the vector control plates were yellow (*xylE* positive) indicating there was no intrachromosomal allelic exchange. However, ~60% of the colonies with the rescue plasmid on LB Kan + Strep + sucrose plates were white in color (x*ylE* negative), thus are the candidates for DCO events. Southern blot analysis of chromosomal DNA showed that all the DCO candidates had undergone intrachromosomal allelic exchange (Fig. 2.1.10 A and B). Hence, allelic exchange was possible only in the presence of the rescue plasmid.

The growth characteristics of the conditional mutant of *MSMEG_6400* (RM1) and *M. smegmatis* WT were tested at 30° and 42°C (Fig. 2.1.11). As expected the RM1 exhibited a reduced growth characteristics to that of *M. smegmatis* WT at 30°C; however

A



Plasmid / strain	Expected fragments (Sma I)
Integration Plasmid	8.5kb, 930bp
WT	1998bp
Single Cross Over (SCO)	~8 Kb, 2330bp, 930bp
Rescue Plasmid	4.5Kb
Double Cross over (DCO)	4.5Kb, 2330b, 930bp

FIGURE 2.1.10 Schematic representation of homologous recombination method for construction of *MSMEG_6400* conditional mutant. The expected restriction patterns and Southern blot analysis resulting from various plasmids and strains were also shown. Gen^r: gentamycin resistance, Kan^r: kanamycin resistance, TS ori: temperature sensitive origin of replication, *SacB*: levansucrase gene (counter selectable marker), xylE: catechol dioxigenase gene (colored marker) **A**. Genetic organization, partial restriction map and expected hybridization profiles of *MSMEG_6400* chromosomal region of *M. smegmatis* WT, single cross over (SCO) event and double cross over event. Probe used for the Southern blot (**B**) is indicated by a thick black line. **B**. Southern blot analysis of pSRM1, WT, WT+pSRM1, SCO and conditional mutants with the rescue plasmid pSRM2 (DCO). The lane assignments are indicated on the right. The expected fragment sizes for each of the strains and the plasmids are given in the table next to the blot.

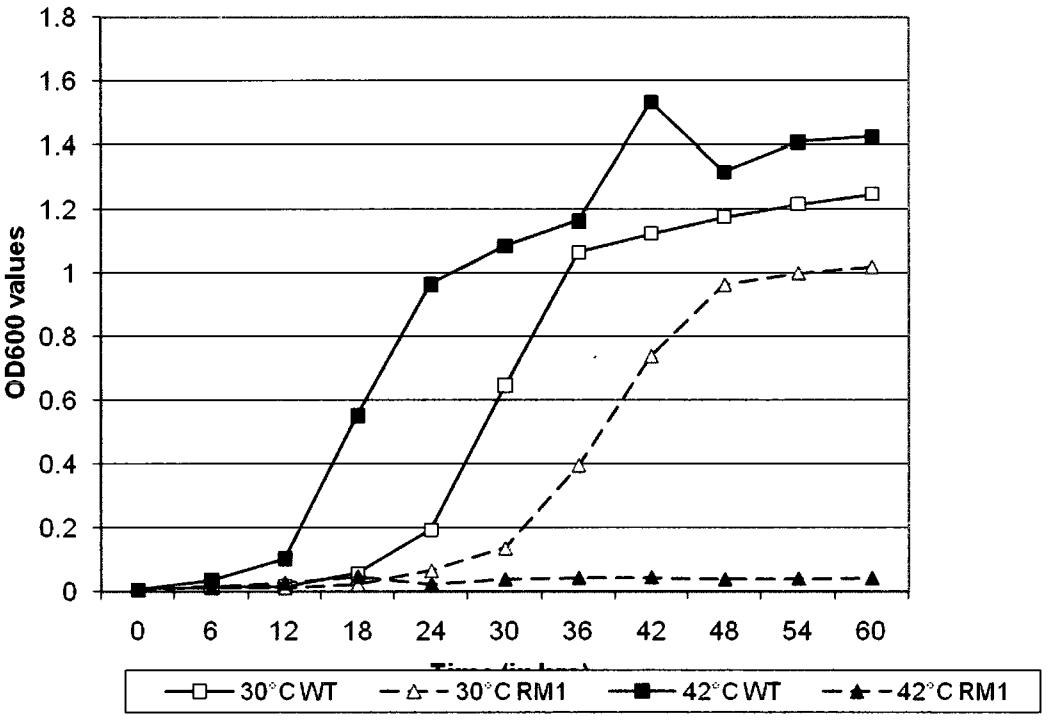


FIGURE 2.1.11 Growth characteristics of *M. smegmatis* WT (solid line) and RM1 (dotted line) at 30° and 42°C. Shown are the growth curves of *M. smegmatis* WT at 30°C (□),*M. smegmatis* WT at 42°C (■), RM1 at 30°C (△) and RM1 at 42°C (▲) cultivated in LB Lennox with Tween 80 and LB Lennox with Tween 80 + Kan respectively.

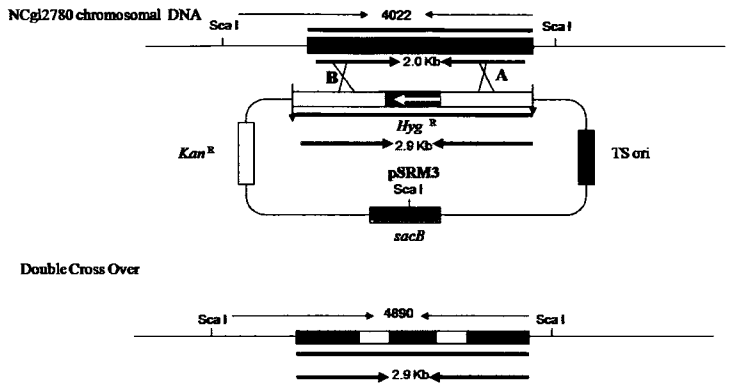
RM1 was unable to grow at 42°C while the *M. smegmatis* WT continued to grow exponentially. This further suggests that *MSMEG_6400* is essential for *M. smegmatis*.

Lipid, LAM and AG analysis of RM1 strain placed at 42°C for different times points with or without ^{14}C label (whole cell labeling) experiments did not yield any significant results. The technical challenges were very few cells from 42°C culture or ^{14}C labeling was not uniform since cells started to lyse.

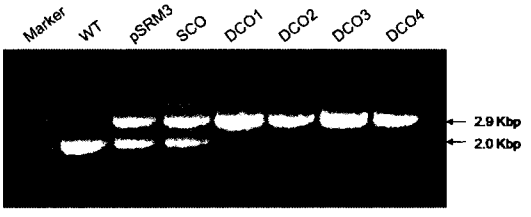
2.1.4.4 Construction of *Coynebacterium glutamicum NCgl2780* mutant

Since the question about the role of Rv3805c in AG and/or LAM biosynthesis was unattained, the ortholog in *C. glutamicum* (*NCgl2780*) was targeted, since *C. glutamicum* was shown to be able to survive even without AG and corynomycolic acids (2, 56). In *C. glutamicum* a LM like molecule with a single arabinose extension on the mannan has been described (19). In an attempt to obtain a deletion mutant of *NCgl2780*, a conditional replication plasmid pSRM3 was constructed which has corynebacterial temperature sensitive origin of replication, Kan^r and *sacB* counterselectable markers. The plasmids pSFKT2sacB (empty vector) and pSRM3 were introduced into *C. glutamicum*, and the transformants were selected on Kan and Kan + Hyg plates at 35°C respectively. Colonies obtained with the plasmid pSRM3 were possible candidates of single homologous recombination wherein the plasmid was integrated into the chromosome at *NCgl2780* locus (SCO). SCOs were then selected on 2% sucrose plates to obtain the mutants that had undergone intrachromosomal allelic exchange (DCO). PCR and Southern blot analysis showed that all the DCOs had undergone intrachromosomal allelic exchange (Fig. 2.1.12 A, B and C).

A



B



C

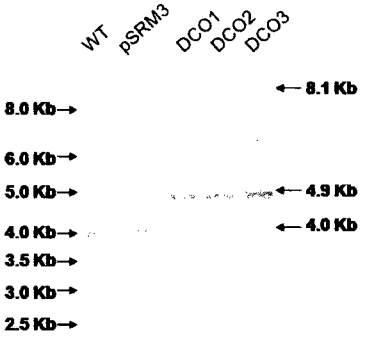


FIGURE 2.1.12 Schematic of expected restriction pattern, PCR and southern blot analysis the *C. glutamicum NCgl2780* mutant. **A**. Genetic organization, partial restriction map and expected hydridizatiom profiles of *NCgl2780* chromosomal region of *C. glutamicum* WT and in case of allelic exchange event. The expected PCR fragments for the plasmid pSRM3 and strains (WT and DCO) are shown in red colored letters. **B**. PCR analysis of the chromosomal DNA obtained from *C. glutamicum* WT, single cross over mutant and double cross over mutant using *NCgl2780* gene primers listed in Table 2.1.1. **C**. Southern blot analysis of pSRM1, WT, WT+pSRM1, SCO and conditional mutants with the rescue plasmid pSRM2 (DCO). PCR product of full length *NCgl2780* gene was used as the probe for the southern blot which is indicated by a thick black line in the FIG 2.1.12A.

2.1.4.5 Analysis of arabinan motifs of AG from *NCgl2780* mutant

Endoarabinanase digestion of the AG from Δ*NCgl2780* mutant (RM2) and analysis by HPAEC yielded a profile that was significantly different from that of AG from *C. glutamicum* WT (Fig. 2.1.13). There was almost complete loss of Ara_6 and presence of a new peak from the AG of RM2. However the new peak matches with the profile obtained from Acceptor 1 in (71) (branched Penta-Ara acceptor) after endoarabinanase digestion and HPAEC analysis and thus indicated that the new peak might be a branched Ara_4 ([α-D-Ara*f*]$_2$-3,5-α-D-Ara*f*-(1→5)-α-D-Ara*f*) (Fig. 2.12). This also indirectly indicates that the terminal Ara residues were lost in the Ara_6 motif of the RM2, which are linked (β(1→2)). So these results indicate that *NCgl2780* is involved in the transfer of β(1→2) Ara*f* unit in the Ara_6 motif of AG.

2.1.5 DISCUSSION

Tuberculosis (TB) remains as a major concern of morbidity and mortality throughout the world with the emergence of the muiltidrug resistant TB (MDR-TB) and extremely drug resistant TB (XDR-TB), therefore is urgent to develop more new anti tuberculosis drugs. Mycobacterial sps have a distinct cell wall which is neither completely of Gram positive nor Gram negative types, due to the presence of two major unusual molecules, MAGP and LAM, so these are obviously potential drug targets. The most successful drugs isoniazid, ethionamide and ethambutol administered against tuberculosis are directed to the cell walls of *M. tuberculosis*. Earlier studies illustrated the importance of three proteins EmbA, EmbB and EmbC in arabinan synthesis and these data indicated that

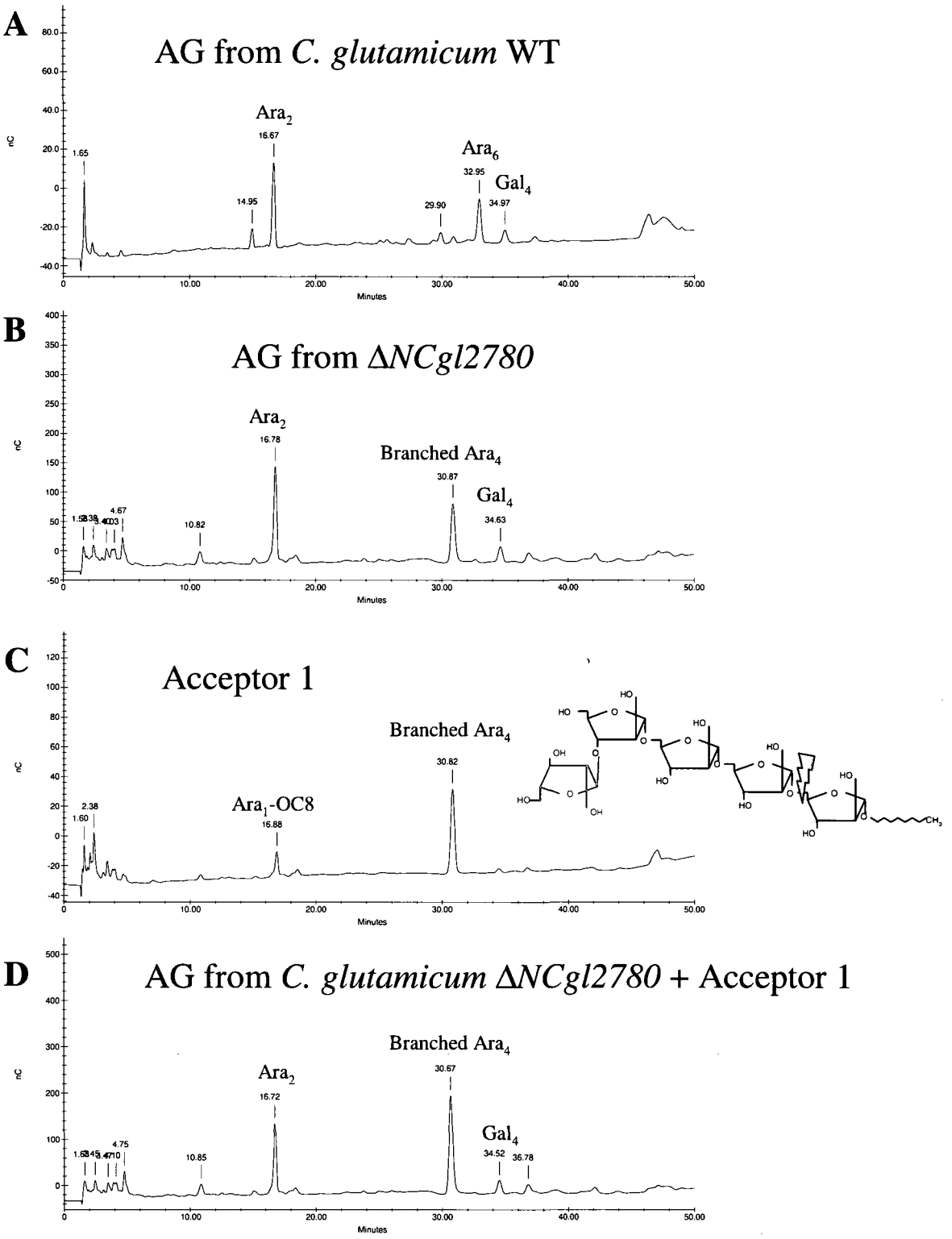


FIGURE 2.1.13 Dionex-HPAEC profiles of the endoarabinanase digestion products of AG from C. *glutamicum*WT (A), ΔNCgl2780 (B) synthetic branched penta-arabinose acceptor (Acceptor 1) (68) (C) and ΔNCgl2780 + branched penta-arabinose acceptor (D).

embA and *embB* are involved in α(1→3) branching in AG, while *embC* is required for extension of the lipomannan (LM) to lipoarabinomannan (LAM) (5, 20, 72). Initially, a gene cluster from *M. avium* comprising *embA*, *embB* and *embC* conferring a higher level of ethambutol resistance was identified and later found to be conserved in all the mycobacteria (5). Belanger et al, identified a putative cell wall synthesis cluster around these *emb* genes (4). Several genes in this cluster were uncharacterized at the time we bagan our studies, but were thought to be potential targets of genes in the cell wall biosynthesis.

In the present study through our *in vitro* arabinosyltransferases assays, we provide evidence that Rv3805c could transfer a single Ara*f* residue on to different linear arabinose acceptors. Furthermore two products were formed from di-Ara-Pentenyl acceptor which was also supported by earlier studies (36) and also from unpublished data by Angala et al. The two products formed were α-D-Ara*f*-(1→5)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-OC_5 (minor product) and α-D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-OC_5 (major product). Furhthermore, the radiolabeled product D-Ara*f*-(1→2)-α-D-Ara*f*-(1→5)-α-D-Ara*f*-OC_5 was found to be more in Rv3805c overexpression than compared to that of the control, suggesting Rv3805c is a β (1→2) arabinosyltransefrase.

The ortholog of *Rv3805c* in *M. smegmatis* (MSMEG_6400) is essential, since we were unable to generate a deletion mutant, a functional copy of this gene was needed for *M. smegmatis* survival before the chromosomal copy could be disrupted. This data is consistent with the results obtained by sassetti et al (59), who predicted that the *Rv3805c* could be an essential gene by their screen transposon of mutants in *M. tuberculosis*. In contrast the ortholog, *NCgl2780* is not essential to *C. glutamicum*. Moreover, when the

endoarabinanase enzyme was used to digest AG of the *ΔNCgl2780*, branched Ara_4 was obtained, the loss of terminal arabinose residues in non reducing ends of hexa arabinose motif (Ara_6) of AG.

During the period these studies were conducted in our laboratory, similar observations were obtained by Seidel et al. and published (63). These authors showed that Rv3805c (named as *AftB*) is involved in the β(1→2) glycosylation of the terminal non reducing arabinose and is non essential in *C. glutamicum*. It was also observed from our results and the other study (63) that both the β-Ara*f* residues were lost in Δ*aftB* of *C. glutamicum* on the Ara_6 motif of AG, indicating that the enzyme is involved in the glycosylation of both the terminal β-Ara*f* residues. In *M. tuberculosis* mycolic acids are attached to both the terminal β-Ara*f* and penultimate α(1→2)-Ara*f* residue of the Ara_6 motif of AG. In *C. glutamicum ΔNCgl2780*, it was observed that even in the absence of terminal β-Ara*f* residues, mycolic acids are still bound to the penultimate α(1 →2)-Araf residue of the Ara_6 motif of AG (63). As indicated by Seidel et al. a same gene (Rv3805c /*AftB*) or two genes might be involved in the glycosylation of two terminal β-Ara*f* residues of AG, where in which *AftB* is required to add one β-Araf residue before a second new arabinosyltransferase adds on the second β-Ara*f* residue (63). Otherwise, it could be that the mycolylation first starts on the terminal β-Ara*f* residues and then occurs on the penultimate α(1 →2) Ara*f* residue of the Ara_6 motif of AG in mycobacteria, which may not be needed in the case of *C. glutamicum*. If that's the case then this *AftB* gene is essential in mycobacteria, because cells cannot survive without mycolic acids. It should be also noted that *C. glutamicum* can survive even without mycolic acids (56). Nevertheless, further analysis should be done to prove these hypotheses. However, since

Rv3805c is an essential gene in mycobacterial sps, it is a good target for mycobacterial drug development.

2.1.6 REFERENCES

1. **Alderwick, L. J., E. Radmacher, M. Seidel, R. Gande, P. G. Hitchen, H. R. Morris, A. Dell, H. Sahm, L. Eggeling, and G. S. Besra.** 2005. Deletion of Cg-emb in corynebacterianeae leads to a novel truncated cell wall arabinogalactan, whereas inactivation of Cg-ubiA results in an arabinan-deficient mutant with a cell wall galactan core. J Biol Chem **280:**32362-71.

2. **Alderwick, L. J., M. Seidel, H. Sahm, G. S. Besra, and L. Eggeling.** 2006. Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in *Mycobacterium tuberculosis*. J Biol Chem **281:**15653-61.

3. **Asselineau, C., J. Asselineau, G. Laneelle, and M. A. Laneelle.** 2002. The biosynthesis of mycolic acids by Mycobacteria: current and alternative hypotheses. Prog Lipid Res **41:**501-23.

4. **Belanger, A. E., and Inamine, J. M. .** 2000. Genetics of cell wall biosynthesis. ASM Press, Washington, DC.

5. **Belanger, A. E., G. S. Besra, M. E. Ford, K. Mikusova, J. T. Belisle, P. J. Brennan, and J. M. Inamine.** 1996. The embAB genes of Mycobacterium avium encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. Proceedings of the National Academy of Sciences of the United States of America **93:**11919-11924.

6. **Belanova, M., P. Dianiskova, P. J. Brennan, G. C. Completo, N. L. Rose, T. L. Lowary, and K. Mikusova.** 2008. Galactosyl transferases in mycobacterial cell wall synthesis. J Bacteriol **190:**1141-5.

7. **Belisle, J. T., V. D. Vissa, T. Sievert, K. Takayama, P. J. Brennan, and G. S. Besra.** 1997. Role of the major antigen of *Mycobacterium tuberculosis* in cell wall biogenesis. Science **276:**1420-2.

8. **Berg, S., D. Kaur, M. Jackson, and P. J. Brennan.** 2007. The glycosyltransferases of *Mycobacterium tuberculosis* - roles in the synthesis of arabinogalactan, lipoarabinomannan, and other glycoconjugates. Glycobiology **17:**35-56R.

9. **Berg, S., J. Starbuck, J. B. Torrelles, V. D. Vissa, D. C. Crick, D. Chatterjee, and P. J. Brennan.** 2005. Roles of conserved proline and glycosyltransferase motifs of EmbC in biosynthesis of lipoarabinomannan. J Biol Chem **280:**5651-63.

10. **Bhamidi, S., M. S. Scherman, C. D. Rithner, J. E. Prenni, D. Chatterjee, K. H. Khoo, and M. R. McNeil.** 2008. The identification and location of succinyl residues and the characterization of the interior arabinan region allow for a model of the complete primary structure of *Mycobacterium tuberculosis* mycolyl arabinogalactan. J Biol Chem **283:**12992-3000.

11. **Brennan, P. J., and D. C. Crick.** 2007. The cell-wall core of *Mycobacterium tuberculosis* in the context of drug discovery. Curr Top Med Chem **7:**475-88.

12. **Briken, V., S. A. Porcelli, G. S. Besra, and L. Kremer.** 2004. Mycobacterial lipoarabinomannan and related lipoglycans: from biogenesis to modulation of the immune response. Mol Microbiol **53:**391-403.

13. **Chatterjee, D., and K. H. Khoo.** 1998. Mycobacterial lipoarabinomannan: an extraordinary lipoheteroglycan with profound physiological effects. Glycobiology **8:**113-20.

14. **Crick, D. C., S. Mahapatra, and P. J. Brennan.** 2001. Biosynthesis of the arabinogalactan-peptidoglycan complex of *Mycobacterium tuberculosis*. Glycobiology **11:**107R-118R.

15. **de Souza, M. V., L. Ferreira Mde, A. C. Pinheiro, M. F. Saraiva, M. V. de Almeida, and M. S. Valle.** 2008. Synthesis and biological aspects of mycolic acids: an important target against *Mycobacterium tuberculosis*. ScientificWorldJournal **8:**720-51.

16. **Deng, L., K. Mikusova, K. G. Robuck, M. Scherman, P. J. Brennan, and M. R. McNeil.** 1995. Recognition of multiple effects of ethambutol on metabolism of mycobacterial cell envelope. Antimicrob Agents Chemother **39:**694-701.

17. **Dhiman, R. K., M. C. Schulbach, S. Mahapatra, A. R. Baulard, V. Vissa, P. J. Brennan, and D. C. Crick.** 2004. Identification of a novel class of omega,E,E-farnesyl diphosphate synthase from *Mycobacterium tuberculosis*. J Lipid Res **45:**1140-7.

18. **Dinadayala, P., D. Kaur, S. Berg, A. G. Amin, V. D. Vissa, D. Chatterjee, P. J. Brennan, and D. C. Crick.** 2006. Genetic basis for the synthesis of the immunomodulatory mannose caps of lipoarabinomannan in *Mycobacterium tuberculosis*. J Biol Chem **281:**20027-35.

19. **Eggeling, L., and Bott, M.,.** 2005. Handbook of Corynebacterium glutamicum. Taylor & Francis, CRC press New york.

20. **Escuyer, V. E., M. A. Lety, J. B. Torrelles, K. H. Khoo, J. B. Tang, C. D. Rithner, C. Frehel, M. R. McNeil, P. J. Brennan, and D. Chatterjee.** 2001. The role of the embA and embB gene products in the biosynthesis of the terminal hexaarabinofuranosyl motif of Mycobacterium smegmatis arabinogalactan. J Biol Chem **276:**48854-62.

21. **Goude, R., A. G. Amin, D. Chatterjee, and T. Parish.** 2008. The critical role of embC in *Mycobacterium tuberculosis*. J Bacteriol **190:**4335-41.

22. **Guerin, M. E., D. Kaur, B. S. Somashekar, S. Gibbs, P. Gest, D. Chatterjee, P. J. Brennan, and M. Jackson.** 2009. New insights into the early steps of

phosphatidylinositol mannosides biosynthesis in mycobacteria. PIMB is an essential enzyme of mycobacterium smegmatis. J Biol Chem.

23. **Huang, H., M. S. Scherman, W. D'Haeze, D. Vereecke, M. Holsters, D. C. Crick, and M. R. McNeil.** 2005. Identification and active expression of the *Mycobacterium tuberculosis* gene encoding 5-phospho-{alpha}-d-ribose-1-diphosphate: decaprenyl-phosphate 5-phosphoribosyltransferase, the first enzyme committed to decaprenylphosphoryl-d-arabinose synthesis. J Biol Chem **280:**24539-43.

24. **Jackson, M., D. C. Crick, and P. J. Brennan.** 2000. Phosphatidylinositol is an essential phospholipid of mycobacteria. J Biol Chem **275:**30092-9.

25. **Kaur, D., S. Berg, P. Dinadayala, B. Gicquel, D. Chatterjee, M. R. McNeil, V. D. Vissa, D. C. Crick, M. Jackson, and P. J. Brennan.** 2006. Biosynthesis of mycobacterial lipoarabinomannan: role of a branching mannosyltransferase. Proc Natl Acad Sci U S A **103:**13664-9.

26. **Kaur, D., M. R. McNeil, K. H. Khoo, D. Chatterjee, D. C. Crick, M. Jackson, and P. J. Brennan.** 2007. New insights into the biosynthesis of mycobacterial lipomannan arising from deletion of a conserved gene. J Biol Chem **282:**27133-40.

27. **Kaur, D., A. Obregon-Henao, H. Pham, D. Chatterjee, P. J. Brennan, and M. Jackson.** 2008. Lipoarabinomannan of Mycobacterium: mannose capping by a multifunctional terminal mannosyltransferase. Proc Natl Acad Sci U S A **105:**17973-7.

28. **Khasnobis, S., J. Zhang, S. K. Angala, A. G. Amin, M. R. McNeil, D. C. Crick, and D. Chatterjee.** 2006. Characterization of a specific arabinosyltransferase activity involved in mycobacterial arabinan biosynthesis. Chem Biol **13:**787-95.

29. **Kolattukudy, P. E., N. D. Fernandes, A. K. Azad, A. M. Fitzmaurice, and T. D. Sirakova.** 1997. Biochemistry and molecular genetics of cell-wall lipid biosynthesis in mycobacteria. Mol Microbiol **24:**263-70.

30. **Kordulakova, J., M. Gilleron, K. Mikusova, G. Puzo, P. J. Brennan, B. Gicquel, and M. Jackson.** 2002. Definition of the first mannosylation step in phosphatidylinositol mannoside synthesis. PimA is essential for growth of mycobacteria. J Biol Chem **277:**31335-44.

31. **Kordulakova, J., M. Gilleron, G. Puzo, P. J. Brennan, B. Gicquel, K. Mikusova, and M. Jackson.** 2003. Identification of the required acyltransferase step in the biosynthesis of the phosphatidylinositol mannosides of mycobacterium species. J Biol Chem **278:**36285-95.

32. **Kremer, L., Baulard, A. R., and Besra.G. S., .** 2000. Genetics of mycolic acid biosynthesis. . ASM Press, Washington, D.C. USA.

33. **Kremer, L., L. G. Dover, C. Morehouse, P. Hitchin, M. Everett, H. R. Morris, A. Dell, P. J. Brennan, M. R. McNeil, C. Flaherty, K. Duncan, and G. S. Besra.** 2001. Galactan biosynthesis in *Mycobacterium tuberculosis*. Identification of a bifunctional UDP-galactofuranosyltransferase. J Biol Chem **276:**26430-40.

34. **Kremer, L., S. S. Gurcha, P. Bifani, P. G. Hitchen, A. Baulard, H. R. Morris, A. Dell, P. J. Brennan, and G. S. Besra.** 2002. Characterization of a putative alpha-mannosyltransferase involved in phosphatidylinositol trimannoside biosynthesis in *Mycobacterium tuberculosis*. Biochem J **363:**437-47.

35. **Lea-Smith, D. J., K. L. Martin, J. S. Pyke, D. Tull, M. J. McConville, R. L. Coppel, and P. K. Crellin.** 2008. Analysis of a new mannosyltransferase required for the synthesis of phosphatidylinositol mannosides and lipoarbinomannan reveals two lipomannan pools in corynebacterineae. J Biol Chem **283:**6773-82.

36. **Lee, R. E., P. J. Brennan, and G. S. Besra.** 1997. Mycobacterial arabinan biosynthesis: the use of synthetic arabinoside acceptors in the development of an arabinosyl transfer assay. Glycobiology **7:**1121-8.

37. **Li, W., Y. Xin, M. R. McNeil, and Y. Ma.** 2006. rmlB and rmlC genes are essential for growth of mycobacteria. Biochem Biophys Res Commun **342:**170-8.

38. **Ma, Y., F. Pan, and M. McNeil.** 2002. Formation of dTDP-rhamnose is essential for growth of mycobacteria. J Bacteriol **184:**3392-5.

39. **Ma, Y., R. J. Stern, M. S. Scherman, V. D. Vissa, W. Yan, V. C. Jones, F. Zhang, S. G. Franzblau, W. H. Lewis, and M. R. McNeil.** 2001. Drug targeting *Mycobacterium tuberculosis* cell wall synthesis: genetics of dTDP-rhamnose synthetic enzymes and development of a microtiter plate-based screen for inhibitors of conversion of dTDP-glucose to dTDP-rhamnose. Antimicrob Agents Chemother **45:**1407-16.

40. **Mahapatra, S., D. C. Crick, M. R. McNeil, and P. J. Brennan.** 2008. Unique structural features of the peptidoglycan of Mycobacterium leprae. J Bacteriol **190:**655-61.

41. **McNeil, M., M. Daffe, and P. J. Brennan.** 1990. Evidence for the nature of the link between the arabinogalactan and peptidoglycan of mycobacterial cell walls. J Biol Chem **265:**18200-6.

42. **McNeil, M., M. Daffe, and P. J. Brennan.** 1991. Location of the mycolyl ester substituents in the cell walls of mycobacteria. J Biol Chem **266:**13217-23.

43. **McNeil, M. R., K. G. Robuck, M. Harter, and P. J. Brennan.** 1994. Enzymatic evidence for the presence of a critical terminal hexa-arabinoside in the cell walls of *Mycobacterium tuberculosis*. Glycobiology **4:**165-73.

44. **Mikusova, K., M. Belanova, J. Kordulakova, K. Honda, M. R. McNeil, S. Mahapatra, D. C. Crick, and P. J. Brennan.** 2006. Identification of a novel galactosyl transferase involved in biosynthesis of the mycobacterial cell wall. J Bacteriol **188:**6592-8.

45. **Mikusova, K., H. Huang, T. Yagi, M. Holsters, D. Vereecke, W. D'Haeze, M. S. Scherman, P. J. Brennan, M. R. McNeil, and D. C. Crick.** 2005. Decaprenylphosphoryl arabinofuranose, the donor of the D-arabinofuranosyl residues of mycobacterial arabinan, is formed via a two-step epimerization of decaprenylphosphoryl ribose. J Bacteriol **187:**8020-5.

46. **Mikusova, K., M. Mikus, G. S. Besra, I. Hancock, and P. J. Brennan.** 1996. Biosynthesis of the linkage region of the mycobacterial cell wall. J Biol Chem **271:**7820-8.

47. **Mikusova, K., R. A. Slayden, G. S. Besra, and P. J. Brennan.** 1995. Biogenesis of the mycobacterial cell wall and the site of action of ethambutol. Antimicrob Agents Chemother **39:**2484-9.

48. **Mikusova, K., T. Yagi, R. Stern, M. R. McNeil, G. S. Besra, D. C. Crick, and P. J. Brennan.** 2000. Biosynthesis of the galactan component of the mycobacterial cell wall. J Biol Chem **275:**33890-7.

49. **Mishra, A. K., S. Batt, K. Krumbach, L. Eggeling, and G. S. Besra.** 2009. Characterization of the Corynebacterium glutamicum deltapimB' deltamgtA double deletion mutant and the role of *Mycobacterium tuberculosis* orthologues Rv2188c and Rv0557 in glycolipid biosynthesis. J Bacteriol **191:**4465-72.

50. **Morita, Y. S., C. B. Sena, R. F. Waller, K. Kurokawa, M. F. Sernee, F. Nakatani, R. E. Haites, H. Billman-Jacobe, M. J. McConville, Y. Maeda, and T. Kinoshita.** 2006. PimE is a polyprenol-phosphate-mannose-dependent mannosyltransferase that transfers the fifth mannose of phosphatidylinositol mannoside in mycobacteria. J Biol Chem **281:**25143-55.

51. **Morita, Y. S., R. Velasquez, E. Taig, R. F. Waller, J. H. Patterson, D. Tull, S. J. Williams, H. Billman-Jacobe, and M. J. McConville.** 2005. Compartmentalization of lipid biosynthesis in mycobacteria. J Biol Chem **280:**21645-52.

52. **Nakamura, J., S. Kanno, E. Kimura, K. Matsui, T. Nakamatsu, and M. Wachi.** 2006. Temperature-sensitive cloning vector for Corynebacterium glutamicum. Plasmid **56:**179-86.

53. **Nigou, J., M. Gilleron, and G. Puzo.** 2003. Lipoarabinomannans: from structure to biosynthesis. Biochimie **85:**153-66.

54. **Nigou, J., T. Vasselon, A. Ray, P. Constant, M. Gilleron, G. S. Besra, I. Sutcliffe, G. Tiraby, and G. Puzo.** 2008. Mannan chain length controls lipoglycans signaling via and binding to TLR2. J Immunol **180:**6696-702.

55. **Pelicic, V., M. Jackson, J. M. Reyrat, W. R. Jacobs, Jr., B. Gicquel, and C. Guilhot.** 1997. Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*. Proc Natl Acad Sci U S A **94:**10955-60.

56. **Portevin, D., C. De Sousa-D'Auria, C. Houssin, C. Grimaldi, M. Chami, M. Daffe, and C. Guilhot.** 2004. A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. Proc Natl Acad Sci U S A **101:**314-9.

57. **Prinzis, S., D. Chatterjee, and P. J. Brennan.** 1993. Structure and antigenicity of lipoarabinomannan from Mycobacterium bovis BCG. J Gen Microbiol **139:**2649-58.

58. **Salman, M., J. T. Lonsdale, G. S. Besra, and P. J. Brennan.** 1999. Phosphatidylinositol synthesis in mycobacteria. Biochim Biophys Acta **1436:**437-50.

59. **Sassetti, C. M., D. H. Boyd, and E. J. Rubin.** 2003. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol **48:**77-84.

60. **Schaeffer, M. L., K. H. Khoo, G. S. Besra, D. Chatterjee, P. J. Brennan, J. T. Belisle, and J. M. Inamine.** 1999. The pimB gene of *Mycobacterium tuberculosis* encodes a mannosyltransferase involved in lipoarabinomannan biosynthesis. J Biol Chem **274:**31625-31.

61. **Scherman, M. S., L. Kalbe-Bournonville, D. Bush, Y. Xin, L. Deng, and M. McNeil.** 1996. Polyprenylphosphate-pentoses in mycobacteria are synthesized from 5-phosphoribose pyrophosphate. J Biol Chem **271:**29652-8.

62. **Schweizer, H. P., and T. T. Hoang.** 1995. An improved system for gene replacement and xylE fusion analysis in Pseudomonas aeruginosa. Gene **158:**15-22.

63. **Seidel, M., L. J. Alderwick, H. L. Birch, H. Sahm, L. Eggeling, and G. S. Besra.** 2007. Identification of a novel arabinofuranosyltransferase AftB involved in a terminal step of cell wall arabinan biosynthesis in Corynebacterianeae, such as Corynebacterium glutamicum and *Mycobacterium tuberculosis*. J Biol Chem **282:**14729-40.

64. **Shi, L., R. Zhou, Z. Liu, T. L. Lowary, P. H. Seeberger, B. L. Stocker, D. C. Crick, K. H. Khoo, and D. Chatterjee.** 2008. Transfer of the first arabinofuranose residue to galactan is essential for Mycobacterium smegmatis viability. J Bacteriol **190:**5248-55.

65. **Takayama, K., C. Wang, and G. S. Besra.** 2005. Pathway to synthesis and processing of mycolic acids in *Mycobacterium tuberculosis*. Clin Microbiol Rev **18:**81-101.

66. **Trent, M. S., A. A. Ribeiro, S. Lin, R. J. Cotter, and C. R. Raetz.** 2001. An inner membrane enzyme in Salmonella and Escherichia coli that transfers 4-amino-4-deoxy-L-arabinose to lipid A: induction on polymyxin-resistant mutants and role of a novel lipid-linked donor. J Biol Chem **276:**43122-31.

67. **Vilcheze, C., and W. R. Jacobs, Jr.** 2007. The mechanism of isoniazid killing: clarity through the scope of genetics. Annu Rev Microbiol **61:**35-50.

68. **Vissa, V. D., and P. J. Brennan.** 2001. The genome of Mycobacterium leprae: a minimal mycobacterial gene set. Genome Biol **2:**REVIEWS1023.

69. **Wolucka, B. A., M. R. McNeil, E. de Hoffmann, T. Chojnacki, and P. J. Brennan.** 1994. Recognition of the lipid intermediate for arabinogalactan/arabinomannan biosynthesis and its relation to the mode of action of ethambutol on mycobacteria. J Biol Chem **269:**23328-35.

70. **Yagi, T., S. Mahapatra, K. Mikusova, D. C. Crick, and P. J. Brennan.** 2003. Polymerization of mycobacterial arabinogalactan and ligation to peptidoglycan. J Biol Chem **278:**26497-504.

71. **Zhang, J., K. H. Khoo, S. W. Wu, and D. Chatterjee.** 2007. Characterization of a distinct arabinofuranosyltransferase in Mycobacterium smegmatis. Journal of the American Chemical Society **129:**9650-9662.

72. **Zhang, N., J. B. Torrelles, M. R. McNeil, V. E. Escuyer, K. H. Khoo, P. J. Brennan, and D. Chatterjee.** 2003. The Emb proteins of mycobacteria direct arabinosylation of lipoarabinomannan and arabinogalactan via an N-terminal recognition region and a C-terminal synthetic region. Mol Microbiol **50:**69-76.

Chapter 2.2

Summary and conclusion for essentiality and characterization of an arabinosyltransferase involved in the cell wall synthesis of mycobacteria

Mycobacterial cell wall is unique in its structure with macromolecules containing polysaccharides such as arabionogalactan (AG) and lipoarabinomannan (LAM) and lipids like mycolic acids. The biosynthetic process of these complex molecules is being elucidated in a day to day manner. Thus AG and LAM biosynthesis involves several glycosyltransferases (GTF). In search of these GTFs; we identified one putative GTF (*Rv3805c*) located in the cell wall synthesis gene cluster (2). It has a conserved glycosyltransferase amino acid motifand belongs to glycosyltransferases super family C (GT-C). Sequence alignments of Rv3805c with Emb proteins revealed presence a proline rich motif which was found to be present in previously characterized Emb proteins. Topology of Rv3805c has ten N terminal transmembrane segments followed by a C terminal hydrophilic domain. In *in vitro* arabinosyltransferase assays using Di-Ara, Tri-Ara and Tetra-Ara synthetic acceptors, and pRpp as donor for arabinose we found that increased product was formed using cell free membrane and particulate fractions prepared from *M. smegmatis* overexpressing Rv3805c compared to vector control strains. We also identified that a single Ara*f* residue is added on the acceptors. Based on several lines of evidence we concluded that Rv3805c catalyzes $\beta(1\rightarrow2)$.

To confirm its biological role we attempted to generate a knock out mutant of this gene in *M. smegmatis*. Allelic exchange at locus *MSMEG_6400* was only possible in the presence of the rescue plasmid and failure of the *MSMEG_6400* conditional mutant to grow at 42°C which indicated that *MSMEG_6400* is essential for *M. smegmatis*.

We were able to obtain a mutant of *NCgl2780* of *C. glutamicum* (ortholog of *Rv3805c*), and the analysis of arabinan motifs in the arabinogalactan of *NCgl2780* mutant revealed

that it lacks β(1→2) Ara*f* residues at the non-reducing ends. During the period of these studies similar observations were obtained by Seidel et al (9).

However, few questions remain unanswered:

Why Rv3805c is essential in *M. smegmatis* and not in *C. glutamicum*: Is this because of the loss of β capping or due to any indirect effect i.e. the affects on the mycolylation in the Ara_6 motif of AG in mycobacteria ? The non-reducing Ara_6 motif, [Ara*f*β(1→2)Ara*f*α(1→5)] [Ara*f*β(1→2)Ara*f*α(1→3)] Ara*f*α(1→5) Ara*f*α(1→ is a template for mycolylation in mycobacteria. If the bacteria lose ability of β capping on AG, the elongation of arabinan chain length may not be impeded, which is a loss of energy and precursors for the bacteria, and may also affect the mycolylation.

It is also interesting to see if *Rv3805c* is involved in the addition of β(1→2) Ara*f* residues at the non-reducing ends of lipoarabinomannan.

Future directions

To address the above questions the following experiments were proposed:

- In the construction of conditional mutant of *MSMEG_640,* we used a rescue plasmid with temperature sensitive origin of replication; the main technical challenge was very few cells were obtained at 42°C, since the cells started to lyse. Therefore we could not get any significant results from the structural analysis of AG and LAM. Alternatively a inducible rescue plasmid such as pVV16 with acetamide promoter, where we can regulate the level of expression of Rv3805c by slowly depleting the

acetamide concentration in the culture medium can be explored. By this approach we might solve various unanswered questions.

- Mycobacterium ssp are not viable without AG and mycolic acids whereas *C. glutamicum* can survive even (1, 6) without them. Even though the cell wall biosynthetic machinery is similar between mycobacterial and corynebacterial ssp, the cell wall dynamics are different. Seidel et al (9) hypothesized that mycolylation of the penultimate Ara*f* residue may occur before the β(1→2)-linked Ara*f* residues is attached, due to the fact that mycolic acids are still bound to the penultimate Ara*f* residues at the non reducing ends of the Ara_6 motif of AG of *C. glutamicum NCgl2780* mutant. On the other hand, ethambutol (EMB), the frontline drug for anti-tuberculosis was found to inhibit the polymerization of arabinan in the cell wall in both AG and LAM (5). Two genes (*embA* and *embB*) were identified in the EMB resistance strains which were involved in the polymerization of arabinan in AG (3). Analysis of these *emb* deletion mutants in *M. smegmatis* revealed that the arabinan content of AG was decreased, and in both mutants, the non-reducing Ara_6 motif was altered. The branching i.e. [βAra*f* (1→2)-αAra*f* (1→3)] in the Ara_6 motif was lost in both the mutants. However, *embA* and *embB* mutants in *M. smegmatis* were still able to survive (4). So if Seidel et al. hypothesis was correct then MSMEG_6400 should not be essential for *M. smegmatis* since mycolylation can still occur on the penultimate Ara*f* residues of Ara_6 motif.

We have shown that MSMEG_6400 is essential for *M. smegmatis*, so we hypothesize that mycolylation first starts on the terminal βAra*f* residues and then occurs on the penultimate αAra*f* residue of the Ara_6 motif of AG in mycobacteria.

This hypothesis can be tested by using an *in vitro* assay. Two synthetic acceptors will be generated, one with [α-D-Ara*f*]$_2$-3,5-α-D-Ara*f*-(1→5)- α -D-Ara*f*-(1→5)- α-D-Ara*f* with a pentenyl chain (A1) similar to that of acceptor 1 in Zhang et. al. 2007 (10) and the other acceptor with both the terminal βAraf residues i.e., [β-D-Ara*f*]$_2$-[α-D-Ara*f*]$_2$-3,5-α-D-Ara*f*-(1→5)- α -D-Ara*f*-(1→5)- α -D-Ara*f* with a pentenyl chain (A2). Then both the acceptors will be biotinylated at the pentenyl portion of the acceptors.

Wild type *M. smegmatis* and *C. glutamicum* strains will be grown in minimal media until early exponential phase and then ^{14}C acetate will be added, to incorporate label into mycolic acids of both strains. Cells will be washed and sonicated. The crude extract will serve as the source of enzyme (for transfer of mycolic acids on to synthetic acceptors) and donor of radiolabeled mycolic acids. Both the acceptors will be incubated for different time points (for e.g. 0 and 1 hr) along with the crude extracts derived from both strains.

After the incubation, the reaction mixture will be passed on a streptavidin column, so that only the biotinylated products will be bound to the column. Later, the streptavidin column-bound products will be eluted. In this manner, other non specific compounds and the unspent reagents will be separated from the products.

Radiolabeled material should be seen in the eluted products in reactions which were incubated for 1 hr, but not in the reactions incubated for 0 hr (negative control) in both the strains. We should see activity with both the acceptors with *C. glutamicum* strain, since the penultimate α(1→2)Ara*f* residue can also be mycolylated even in the absence of βAra*f* residue, as observed in case of *Aft*B mutant in *C. glutamicum* (9).

With respect to *M. smegmatis* strain, we should see activity with only A2 acceptor. The eluted products will be subjected to GC-MS analysis by using the protocol published by Puech et al.(7). This experiment will indirectly prove if the terminal βAra*f* residue is needed for mycolylation in *M. smegmatis*.

- In an alternative approach we could adopt an assay described by Sanki et al. where they showed that antigen 85C (enzyme for mycolylation) could transfer acyl chains from synthetic acyl donors on to the fragments of arabinan (8). Several synthetic arabinose acceptors were tested in the *in vitro* assay. When β-D-Ara*f*-(1→2)-methyl - α-D-Ara*f* (arabinose acceptor 9) was used as an acceptor, *p*-nitrophenyl-6-O-octanonyl-β-D-Glc*p* used as a donor of acyl chain in the *in vitro* assay; octanonyl chains were transferred from *p*-nitrophenyl-6-O-octanonyl-β-D-Glc*p* onto both Ara*f* residues of the acceptor at 5^{th} position mimicking mycolylation of the sugars in AG in the presence of antigen 85C enzyme. However it was not clear if this acyl transferase activity occurs simultaneously on both the sugars or sequentially i.e. first on the β-D-Ara*f* and then α-D-Ara*f* or first on α-D-Ara*f* and then β-D-Ara*f*. So to find out this arabinose acceptor 9 can be modified by blocking C5 of β-D-Ara*f* with a methyl group. Then this modified acceptor can used in the assay (8) with antigen 85C enzyme from both *M. tuberculosis* and *C. glutamicum* wild type strains. We should not see any acyl transferase activity with *M. tuberculosis* antigen 85C if the terminal βAra*f* residue is needed for mycolylation.

- To check if *Rv3805c* is involved in the addition of β(1→2) Ara*f* residues at the non-reducing ends of lipoarabinomannan, an *in vitro* arabinosyltransferase assay will be developed. This includes an acceptor {[α-D-Ara*f*(1→5)]$_6$-(1→5)-α-D-Ara*f* [-α-D-

Ara*f*-(1→3)]α-D-Ara*f*-(1→5)- α -D-Ara*f*-(1→5)- α -D-Ara*f* with a pentenyl/octyl chain (resembles LAM non-reducing end)}, p[^{14}C]Rpp as donor for arabinose and cell free membrane and particulate fractions prepared from Rv3805c overexpressing and vector control strains of *M. smegmatis*. The arabinosyltransferase activity of Rv3805c overexpressing strain will be compared with the vector control strain.

REFERENCES

1. **Alderwick, L. J., M. Seidel, H. Sahm, G. S. Besra, and L. Eggeling.** 2006. Identification of a novel arabinofuranosyltransferase (AftA) involved in cell wall arabinan biosynthesis in Mycobacterium tuberculosis. J Biol Chem **281:**15653-61.

2. **Belanger, A. E., and Inamine, J. M. .** 2000. Genetics of cell wall biosynthesis., vol. ASM Press, Washington, DC.

3. **Belanger, A. E., G. S. Besra, M. E. Ford, K. Mikusova, J. T. Belisle, P. J. Brennan, and J. M. Inamine.** 1996. The embAB genes of Mycobacterium avium encode an arabinosyl transferase involved in cell wall arabinan biosynthesis that is the target for the antimycobacterial drug ethambutol. Proceedings of the National Academy of Sciences of the United States of America **93:**11919-11924.

4. **Escuyer, V. E., M. A. Lety, J. B. Torrelles, K. H. Khoo, J. B. Tang, C. D. Rithner, C. Frehel, M. R. McNeil, P. J. Brennan, and D. Chatterjee.** 2001. The role of the embA and embB gene products in the biosynthesis of the terminal hexaarabinofuranosyl motif of Mycobacterium smegmatis arabinogalactan. J Biol Chem **276:**48854-62.

5. **Mikusova, K., R. A. Slayden, G. S. Besra, and P. J. Brennan.** 1995. Biogenesis of the mycobacterial cell wall and the site of action of ethambutol. Antimicrob Agents Chemother **39:**2484-9.

6. **Portevin, D., C. De Sousa-D'Auria, C. Houssin, C. Grimaldi, M. Chami, M. Daffe, and C. Guilhot.** 2004. A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. Proc Natl Acad Sci U S A **101:**314-9.

7. **Puech, V., C. Guilhot, E. Perez, M. Tropis, L. Y. Armitige, B. Gicquel, and M. Daffe.** 2002. Evidence for a partial redundancy of the fibronectin-binding proteins for the transfer of mycoloyl residues onto the cell wall arabinogalactan termini of Mycobacterium tuberculosis. Mol Microbiol **44:**1109-22.

8. **Sanki, A. K., J. Boucau, D. R. Ronning, and S. J. Sucheck.** 2009. Antigen 85C-mediated acyl-transfer between synthetic acyl donors and fragments of the arabinan. Glycoconj J **26:**589-96.

9. **Seidel, M., L. J. Alderwick, H. L. Birch, H. Sahm, L. Eggeling, and G. S. Besra.** 2007. Identification of a novel arabinofuranosyltransferase AftB involved in a terminal step of cell wall arabinan biosynthesis in Corynebacterianeae, such as Corynebacterium glutamicum and Mycobacterium tuberculosis. J Biol Chem **282:**14729-40.

10. **Zhang, J., K. H. Khoo, S. W. Wu, and D. Chatterjee.** 2007. Characterization of a distinct arabinofuranosyltransferase in Mycobacterium smegmatis. Journal of the American Chemical Society **129:**9650-9662.