DNA FINGERPRINTING OF *Mycobacterium tuberculosis* ISOLATES IN NEPAL USING PCR-LABELED IS*6110* PROBE

A

Dissertation

Submitted to the Central Department of Microbiology

Tribhuvan University

In Partial Fulfillment of the Requirements for the Award of the Degree of

Master of Science in Microbiology

(Medical)

by

Saraswoti Khadge

Central Department of Microbiology

Tribhuvan University

Kirtipur, Kathmandu, Nepal

2007

RECOMMENDATION

This is to certify that Ms. Saraswoti Khadge has completed this dissertation work entitled "**DNA fingerprinting of *Mycobacterium tuberculosis* isolates in Nepal by using PCR-labeled IS*6110* probe**" as a partial fulfillment of M. Sc. degree in Microbiology under our supervision. To our knowledge this work has not been submitted for any other degree.

Dr.Prakash Ghimire, Ph.D.
Associate Professor
Central Department of Microbiology
Tribhuvan University
Kirtipur, Kathmandu, Nepal

Dr. Murdo Macdonald, Ph.D.
Head
Mycobacterial Research Laboratory
Anandaban Hospital
Lele, Lalitpur, Nepal

CERTIFICATE OF APPROVAL

On the recommendation of **Dr. Prakash Ghimire** and **Dr. Murdo Macdonald** this dissertation of **Ms. Saraswoti Khadge,** entitled "**DNA fingerprinting of *Mycobacterium tuberculosis* isolates in Nepal by using PCR labeled IS*6110* probe**" has been approved for the examination and is submitted to Tribhuvan University in Partial fulfillment of the requirement for M. Sc. Degree in Microbiology.

Date: ……..

_____

Dr. Anjana Singh, Ph. D.
Head
Central Department of Microbiology
Tribhuvan University
Kirtipur, Kathmandu
Nepal

BOARD OF EXAMINERS

Recommended by:

_____

Dr Prakash Ghimire, Ph.D.

(Supervisor)

_____

Dr. Murdo Macdonald, Ph.D.

(Supervisor)

Approved by:

_____

Dr. Anjana Singh, Ph.D.

(Head of Department)

Examined by:

_____

(Internal Examiner)

_____

(External Examiner)

Date: ……

ACKNOWLEDGEMENT

It gives me immense pleasure to express my profound gratitude to all those respectable personalities who helped me directly or indirectly for such a long period of time to complete this research work

I would like to express my sincere gratefulness to my supervisors **Dr. Murdo Macdonald**, Head, Mycobacterial Research Laboratory, Anandaban Hospital, Lalitpur, and **Dr. Prakash Ghimire**, Associate Professor, Central Department of Microbiology, Tribhuvan University, for their regular supervision, scholastic inspiration, untiring encouragement and expert guidance during my research work.

I am grateful to **Dr. Anjana Singh**, Head of Department, **Prof. Dr. Shital Raj Basnyat**, and **Mr. Binod Lekhak,** Assistant professor of Central Department of Microbiology for their generous guidance.

I am deeply indebted to **Mr. Chaman Ranjit** and **Mr. Bishwa Raj Sapkota,** for their tremendous support, guidance and meticulous care through out the research work. My special thanks go to **Dr. J. Quimpo, Mrs. S. Falbus, Mr. K.D. Neupane, Mr. I. Shrestha, Mr. S.C Silwal, Mr. D.K. Mahat, Ms. S. Pandey** and **Mr. R. Shrestha** for providing me such a support, kind co-operation and homely environment during the period of lab works. I could not forget the time spent with **Ms. Rachael E. Morling.**

I am extremely thankful to **Dr. Bhawana Shrestha**, **Mr. Bhagwan Maharjan** and all the staffs of GENETUP and **Mr. Dhruba K. Khadka** and all the staffs of NTC, for their co-operation and support during sample collection.

I am much obliged to **Dr. Prasit Palittapongarnpim**, Mahidol University for his kind co-operation and expert guidance. I am equally grateful to **Dr. D.R. Bastola,** Nebraska University and **Dr. G. Lamechane**, John Hopkings University for their valuable guidance for this work.

I would like to thank my friend **Mr. Ajay Poudel** for his support during sample collection. Similarly I am thankful to my friends **Gyan sundar Shahukhal, Niroj man Amatya, Shishir Sharma, Sangeeta Shrestha, Rajani Shrestha** and **Deepa Shrestha.**

Finally, I would like to express my eternal gratitude to my **family members** and my friend **Laxmi** for their constant attention encouragement and moral support throughout my life.

Date……. …….. ……..

Saraswoti Khadge

ABSTRACT

Restriction Fragment Length Polymorphism (RFLP) using IS*6110* probe has been taken as gold standard DNA fingerprinting technology and used all over the world quite successfully to characterize *Mycobacterium tuberculosis* strains. The present study has been carried out to study genomic polymorphism among *M. tuberculosis* isolates collected from patients attending clinics at two different tuberculosis centers of Kathmandu.

M. tuberculosis isolates were collected from pure cultured specimen on Ogawa slants, in sterilized TE buffer pH 8.0, heat killed and brought to MRL, Lele. DNA from these isolates were extracted and purified by physiochemical method, restricted with PvuII enzyme and hybridized with PCR amplified and DIG labeled 245bp IS*6110* probe. Fingerprinting patterns were inspected visually. Among 59 isolates analyzed, 4 isolates were observed to have no copy, 3 had single copy, 9 had 2-5 copies, 23 had 6-12 copies and 20 had 13-17 copies of IS*6110* in their genome. For the purpose of analysis, the patients were divided in to Aryans and Mongols. The low copy numbered strains (≤5 copies) were more common among Aryans while opposite was true for the Mongols. Excluding isolates with one copy or loss of IS*6110*, 23 % of isolates were clustered. All together 48 different fingerprinting patterns were observed showing 80 % of genomic variations amongst isolates. The isolates of the largest cluster were found more commonly in the younger age groups, females, and in patients from the central region. DNA fingerprinting using IS*6110* probe was found to be quite discriminating for molecular typing of most (73 %) of strains which harboured ≥6 copies of IS*6110*. A larger number of isolates from defined geographical area need to be studied to understand molecular epidemiology of M. tuberculosis in Nepal.

Key words - RFLP, IS6110, M. tuberculosis, PvuII, Nepal

TABLE OF CONTENTS

	Page no.
Title page	i
Recommendation	ii
Certificate of approval	iii-iv
Acknowledgement	v
Abstract	vi
Table of contents	vii-x
List of abbreviations	xi-xii
List of tables	xiii
List of figures	xiv
List of photographs	xv
List of appendices	xvi
CHAPTER I: INTRODUCTION	**1-3**
CHAPTER II: OBJECTIVES	**4**
CHAPTTER III: LITERATURE REVIEW	**5-50**
3.1 Tuberculosis	5
3.1.1 Transmission	5
3.1.2 Risk factors	6
3.1.3 TB infection Vs TB disease	7
3.1.4 Host defense mechanism	7
3.1.5 Virulence mechanism	8
3.2 Pathogenesis	9
3.2.1 Primary infection	9
3.2.2 Post primary TB	11
3.3 The etiological agent	11
3.4 Diagnosis	12
3.4.1 Clinical diagnosis	12
3.4.2 Radiological diagnosis	12

3.4.3 Tuberculin skin test 13
3.4.4 Lab diagnosis 13
3.4.4.1 Microscopy 13
3.4.4.2 Culture 14
3.4.4.3 Molecular methods 15
3.5 Treatment and control 15
3.6 Epidemiology of Tuberculosis 15
3.6.1 Tuberculosis: Global aspect 15
3.6.2 Tuberculosis in SAARC countries 16
3.6.3 Tuberculosis in Nepal 16
3.7 Molecular epidemiology 17
3.7.1 Conventional Typing 17
3.7.2 Molecular typing 18
3.7.2.1 Genome of Mycobacteria 18
3.7.2.2 Repetitive elements 19
3.7.2.2.1 Insertion sequence (IS) elements 19
3.7.2.2.1.1 IS*6110* as tool for strain differentiation 22
3.7.2.2.1.2 IS*1081* *25*
3.7.2.2.2 Short repeats 25
3.7.2.2.2.1 Direct repeats 25
3.7.2.2.2.2 Polymorphic GC-rich repetitive sequence 26
3.7.2.2.2.3 Major Polymorphic Tandem Repeats 26
3.7.2.3 Genotyping techniques 27
3.7.2.3.1 Criteria of typing system 27
3.7.2.3.2 IS*6110*-RFLP fingerprinting 29
3.7.2.3.3 PCR based fingerprinting 30
3.7.2.3.3.1 IS*6110* based PCR methods 30
3.7.2.3.3.2 16S and 23S rRNA based methods 31
3.7.2.3.3.3 DR region based methods 31
3.7.2.3.3.4 Minisatelite- based methods 32
3.7.2.4 RFLP as a molecular tool for different studies 32
3.7.2.4.1 Outbreak study and transmission dynamics 32

3.7.2.4.2 Exogenous infection Vs Endogenous reactivation 34
3.7.2.4.3 Geographical distribution & dissemination of tuberculosis 35
3.7.2.5 Well characterized clinical isolates 41
3.7.2.6 Principles of TB-RFLP 43

CHAPTER IV: MATERIAL AND METHODS 51-63
4.1 Materials 51
4.2 Methodology I (NTC/ GENETUP) 51
4.2.1 Sample collection sites 51
4.2.2 Patient enrollment 51
4.2.3 Sample collection and processing 52
4.2.4 Culture 52
4.2.5 Transfer and heat kill of cells 53
4.3 Methodology II (MRL) 53
4.3.1 DNA extraction 53
4.3.2 DNA quantification 54
4.3.3 Digestion of chromosomal DNA by PvuII 54
4.3.4 Separation of digested DNA fragments 55
4.3.4.1 Gel preparation 55
4.3.4.2 Electrophoresis 55
4.3.4.3 Staining and visualization 56
4.3.5 Southern blotting 56
4.3.5.1 Treatment of gel before blotting 56
4.3.5.2 Paper ware set up 57
4.3.5.3 Blotting 57
4.3.5.4 Post blotting membrane treatment 57
4.3.6 Preparation and labeling of DNA probe by PCR 58
4.3.6.1 Optimization of IS*6110*-PCR 58
4.3.6.2 Visualization of amplified DIG labeled IS*6110* probe 60
4.3.6.3 Quantification of PCR probe 61
4.3.7 Southern hybridization 61
4.3.7.1 Prehybridization 61

4.3.7.2 Preparation of working probe 61
4.3.7.3 Hybridization 62
4.3.8 Detection 62
4.3.8.1 Membrane washing 62
4.3.8.2 Membrane DIG-detection 62
4.3.8.3 Autoradiography of membrane 62
4.3.8.4 Stripping and storage of membrane 63
4.3.9 Result analysis 63

CHAPTER V: RESULTS **64-73**
5.1 Sample analysis 64
5.1.1 Sample analysis for part I study 64
5.1.2 Sample analysis for part II study 65
5.2 Result analysis 65

CHAPTER VI: DISCUSSION **74-88**
6.1 Optimization of standard operating procedure (SOP) 75
6.1.1 Optimization of PCR and DIG labeling of probe 76
6.2 Result analysis 78
6.2.1 Analysis of fingerprints by number of copies of IS*6110* *78*
6.2.2 Analysis of fingerprints with respect of gender 82
6.2.3 Analysis of fingerprints with in different age groups 83
6.2.4 Analysis of fingerprints in different geographical regions 84
6.2.5 Analysis of Distribution of fingerprints between races 85
6.2.6 Analysis of clustered samples 86

CHAPTER VII: SUMMARY AND RECOMMENDATION **89-90**
7.1 Summary 89
7.2 Recommendations 90

CHAPTER VIII: REFERENCES **91-108**

ABBREVIATION

AIDS	Acquired Immuno Deficiency syndrome
BPB	Bromophenol blue
CSPD	Disodium 3-(4-methoxyspiro {1, 2-Dioxetane-3, 2' (5'-chloro) tricycle [3.3.1.13, 7] decan}-4-yl) phenyl Phosphate
CTAB	N-cetyl-N,N,N,-trimethyl-ammonium Bromide
dATP	Deoxyadenosine 5'-triphosphate
dCTP	Deoxycytidine 5'-triphosphate
dH2O	Distilled water
dGTP	Deoxyguanosine 5'-triphosphate
DIG	Digoxigenin
DIG-dUTP	Digoxigenin-11-dUTP
DNA	Deoxyribose nucleic acid
dNTPs	Deoxynucleoside 5'-triphosphates
DR	Direct repeat
dUTP	Deoxyuridine 5'-triphosphate
EDTA	Ethylene diamine tetra acetate
EtBr	Ethidium bromide
EtOH	Ethanol
GENETUP	German Nepal Tuberculosis Project
HCS	High Copy Numbered Strains
IS	Insertion Sequence
LCS	Low Copy Numbered strains
L-J	Lowenstein-Jensen medium
MPTR	Major Polymorphic Tandem Repeat
MIRU	Mycobacterial Interspersed Repetitive unit
MRL	Mycobacterial Research laboratory
Min	Minute
ng	nano gram(s) = 10^{-9} gram

NTC	National Tuberculosis Center
OD	Optical density
PCR	Polymerase Chain Reaction
PGRS	Polymorphic Guanine Cytocine rich Sequence
PPD	Purified Protein Derivative
PTB	Pulmonary Tuberculosis
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
RT	Room Temperature
RXN	Reaction(s)
SDS	Sodium Dodecyl Sulphate
Sec	Second(s)
SSC	Saline Sodium Citrate
TB	Tuberculosis
TBE	Tris-Borate EDTA
TE	Tris-EDTA (buffer)
TUTH	Tribhuvan University Teaching Hospital
U	Unit(s) of enzyme
UV	Ultraviolet
V	Volts
VNTR	Variable Number Tandem Repeats

LIST OF TABLES

		Page no.
Table 1:	Repetitive DNA sequences in *M. tuberculosis*	20
Table 2:	Components of digestion mixture	55
Table 3:	Preparation working nucleotides solution	59
Table 4:	Optimization of reaction mixture for IS*6110*-PCR using DIG-11-dUTP	60
Table 5:	Optimization of thermocycler conditions for IS*6110*-PCR	60
Table 6:	Distribution of samples in different parts of the study	64
Table 7:	Gender wise variation of IS*6110* copy numbers of *M. tuberculosis*	67
Table 8:	Distribution of IS*6110* copy numbers among various age groups	67
Table 9:	Distribution of IS*6110* copy numbers in various regions of the country	68
Table 10:	Distribution of IS*6110* copy number between racial groups	69
Table 11:	Distribution of caste and ethnicity of patients in two racial groups studied	70
Table 12:	The profile of clustered and non-clustered cases of *M. tuberculosis* isolates	71
Table13:	The profile of clustered and non-clustered cases of *M. tuberculosis* isolates	72

LIST OF FIGURES

Figure 1: Structure of the DIG-11-dUTP molecule

Figure 2: Flow chart of methodology I (NTC)

Figure 3: Flow Chart of Methodology II (MRL)

Figure 4: Genetic variation of M. tuberculosis on the basis of no. of copies of IS6110

Figure 5: Gender wise distribution of IS6110 copies

Figure 6: Distribution of IS6110 copy numbers in various age groups

Figure 7: Distribution of different number of copies of IS6110-RFLP in various geographical regions.

Figure 8: Distribution of number of copies of IS6110 between racial groups.

LIST OF PHOTOGRAPHS

Photograph 1: Colonies of *M. tuberculosis* on LJ media

Photograph 2: Processing of PCR in the laboratory of MRL

Photograph 3: PvuII digested DNA samples of *M. tuberculosis* isolates on 0.8% agarose

Photograph 4: Capillary transfer set-up for southern blotting of the gel to N^+ membrane

Photograph 5: PCR amplified DIG-11-UTP labeled IS*6110* probe

Photograph 6: IS*6110*-RFLP banding pattern of *M. tuberculosis* isolates

Photograph 7: Matched IS6110- RFLP pattern of clustered isolates of *M. tuberculosis*

LIST OF APPENDECIES

	Page no.
APPENDIX – I: Equiptments and Reagents used for bacterial cultures	i-iv
APPENDIX – II: Equiptments used for IS*6110*-RFLP	v
APPENDIX – III: Reagents and disposables used for IS*6110*-RFLP	vi-vii
APPENDIX – IV: Solutions used for IS*6110*-RFLP	viii-xi
APPENDIX - V: Classification of Mycobacteria	xii
APPENDIX – VI: Treatment regimens for tuberculosis	xiii
APPENDIX – VII: Reading and reporting of laboratory results for cultures	xiv

CHAPTER- I

1. INTRODUCTION

M. tuberculosis, the causative agent of Tuberculosis (TB), is one of the most successful bacterial pathogens in the history of mankind. Despite the availability of effective anti tuberculosis chemotherapy for over 50 years, TB remains a major global health problem. Each year, there are 2 million deaths occur due to tuberculosis, making TB the world's leading cause of mortality due to a single infectious agent. As the rates of TB infection have fallen dramatically in industrialized countries in the past century, resource-poor countries now bear over 90% of all cases globally. In fact, there are more cases of TB today than ever recorded (Mathema *et al.,* 2006).

M. tuberculosis can attack any part of the body, but usually infects the lungs. TB of the brain, kidneys, lymph nodes and spine are non-infectious where as that of the lungs and/or throat is infectious. The bacteria causing TB is spread as airborne particles in the form of droplet nuclei released from the lungs of patients with pulmonary or laryngeal TB (Mathema *et al.,* 2006). Once the bacteria enter alveoli, they resist destruction by the alveolar macrophages and multiply, forming the primary lesions or tubercles; they then spread to regional lymph nodes, enter the circulation, and reseed the lungs. Tissue destruction results from cell-mediated hypersensitivity and is responsible for some major symptoms of the disease like chronic productive cough, low-grade fever, night sweats, and weight loss. A diagnosis of active disease is based on clinical manifestations, an abnormal chest radiograph, acid-fast bacilli in sputum or bronchoscopic specimens and recovery of the organism. Diagnostic assays based upon amplification of mycobacterial genes in clinical specimens are currently being tested. But, Mantoux skin test with purified protein derivative (PPD) is the only widely used method for the diagnosis of recent infection with *M tuberculosis* (McMurray, 2001).

Control of TB can be focused from various angles. One of the major priorities for the success of a control program should therefore be identification and treatment of acid-fast bacillus from smear positive cases and this can be better achieved by understanding

the factors which foster transmission of *M. tuberculosis*. Transmission of TB is influenced by a large number of risk factors and transmission dynamics are therefore different in distinct geographic regions. Dynamics of the disease will only be better understood upon study of the spreading of individual strains of *M. tuberculosis* within a certain population. Until 1990, detection of person to person transfer of *M. tuberculosis* clones was a difficult task; only subgroups of strains could be identified through phage typing and occurrence of some special drug-resistance profiles (Suffy *et al.,* 1997).

Molecular epidemiology is the integration of molecular techniques to track specific strains of pathogens with conventional epidemiological approaches to understand the distribution of disease in population (Small and Moss, 1993; van Embden *et al*., 1992). In the epidemiological studies of *M. tuberculosis,* this discipline began with the identification of IS*6110*, a novel mycobacterial insertion sequence which formed the basis of a reproducible genotyping technique for *M. tuberculosis* (Bergos and Pym, 2002). These technologies have revolutionized our understanding of the transmission of *M. tuberculosis* and enabled us to clarify fundamental questions about the epidemiology and pathogenesis of tuberculosis which were previously obscure (Lillebæk, 2005).

Genetic polymorphisms can be visualized by various subtyping techniques, often referred to as DNA fingerprinting. Specific strains of the *M. tuberculosis* complex can be characterized on the basis of their DNA patterns allowing a high degree of discrimination between different strains (Befani *et al.,* 2002)). The most widely applied DNA fingerprinting technique for *M. tuberculosis* is Restriction Fragment Length Polymorphism (RFLP) typing using the insertion sequence IS*6110* as a probe for strain differentiation (Lillebæk, 2005). IS*6110* is a member of the IS3 family of transposable elements, which is specific to the *M. tuberculosis* complex. The IS*6110*-RFLP method is based on differences in the IS*6110* copy numbers per strain, ranging from 0 to about 25, and variability in the chromosomal positions of these IS*6110* insertion sequences (Cole *et al.,* 2005). The numbers and positions can be visualized and compared using internationally standardized protocols (van Embden *et al.,* 1993).

IS*6110* fingerprinting has proven useful for investigating nosocomial transmission, studying outbreaks, confirming instances of laboratory cross contamination, differentiating relapse caused by endogenous reactivation from re-infection by an exogenous strain and studying TB transmission in large populations (van Soolingen *et al.,* 1993). Furthermore, many studies have shown that epidemiologically unrelated strains have different patterns, whereas those of related strains generally have identical patterns. Therefore, the majority of such studies assume that the strains with identical patterns are connected by recent transmission (Bloom, 1994).

This study was carried out in Nepal for determination of genomic polymorphisms among M. *tuberculosis* isolates of Nepalese patients. Nepal is one of the tuberculosis endemic countries, geographically located between China and India. The copy number of IS*6110* reported from these countries shows extensive variation amongst clinical isolates of *M. tuberculosis* between these two countries. This study aims to determine the distribution patterns of IS*6110* among the *M. tuberculosis* isolates of Nepal.

For the analysis of genomic polymorphisms among *M. tuberculosis* isolate on the basis of IS*6110* copy number, chromosomal DNA was extracted from bacterial culture obtained from sputum samples of different TB patients. IS*6110* based RFLP was carried out according to standardized protocol. Results were interpreted on the basis of IS*6110* copy number and their chromosomal position. Genotyping of *M. tuberculosis* in Nepal would be of great importance for proper identification and characterization of strain types; their correlation with geographical location; understanding transmission patterns and molecular clustering among the isolates in population; distinguishing relapses and reinfection; and most importantly for understanding the transmission dynamics of multi drug resistant strains.

CHAPTER-II

2. OBJECTIVES

General Objective

To study the strain variability of *M. tuberculosis* in clinical isolates of Nepal

Specific Objectives

a) To optimize Restriction Fragment Length Polymorphism (RFLP) as a primary tool of strain differentiation

b) To assess the strain variability of *M tuberculosis* in clinical isolates.

CHAPTER-III

3. LITERATURE REVIEW

3.1 Tuberculosis

Tuberculosis is a chronic infectious disease caused by the bacterium *M. tuberculosis*. Based on site of involvement of infection and host response, tuberculosis (TB) is of different types; pulmonary tuberculosis (PTB), extra pulmonary tuberculosis and disseminated tuberculosis. The clinical expression of infection with *M. tuberculosis* are quite varied and depend upon a number of identified factors like host and microbe related characteristics as well as the consequences of their interactions, which influences the manifestations of tuberculosis infection (Cole *et al.,* 2005).

Before the beginning of the epidemic of HIV infection, approximately 85% of the reported tuberculosis cases were limited to the lungs, with the remaining 15% involving only non-pulmonary sites or both pulmonary and non-pulmonary sites. This proportional distribution is substantially different among persons with HIV infection. The multiplicity of sites is due to an immune system that is limited in its ability to contain infections with *M. tuberculosis*. Included in this category are infants, the elderly and persons with primary and secondary immunodeficiency states resulting from coexisting disease including HIV or malnutrition ((Cole *et al*., 2005).

3.1.1 Transmission

Tuberculosis is spread from person to person through the air by droplet nuclei, particles 1 to 5 μ m in diameter that contain *M. tuberculosis*. Droplet nuclei are produced when persons with pulmonary or laryngeal tuberculosis cough, sneeze, speak, or sing. They also may be produced by aerosol treatments, sputum induction and aerosolization during bronchoscopy, and through manipulation of lesions or processing of tissue or secretions in the hospital or laboratory. Droplet nuclei, containing two to three *M. tuberculosis* organisms, are so small that air currents normally present in any indoor space can keep

them airborne for long periods of time; however, direct sunlight quickly kills the tubercle bacilli. Droplet nuclei are small enough to reach the alveoli within the lungs, where the organisms replicate (American Thoracic Society, 2000). One cough can produce approximately 3000 droplet nuclei. Most patients become non-infectious within 2 weeks after the institution of appropriate chemotherapy because of decrease in the number of organisms excreted and the decrease in frequency of cough (Daniel, 1994). Two factors determine an individual's risk of exposure; the concentration of droplet nuclei in contaminated air and the length of time breathing that air (Harries *et al.*, 1998).

3.1.2 Risk factors

Certain groups of people are at an increased risk of getting TB. These include:

- People sharing the same breathing space with someone who has the TB disease (i.e. family, friends, and co-workers).
- Medically underserved populations.
- The poor and the homeless
- People from countries with high TB incidence (including Latin America, the Caribbean, Africa and Asia (excluding Japan).
- Residents and employees of long-term institutional settings of nursing homes and prisons.
- People who became infected with TB bacteria in the last 2 years.

As immunity wanes, the dormant bacteria reactivate, causing an outbreak of disease often many decades after the initial infection (Cole *et al.,* 1998). The following are at greater risk of TB because they are immunocompromised.

- Substance abusers: alcoholics and intravenous drug users.
- People with medical conditions (i.e. diabetes mellitus, carcinoma of the head and neck, being underweight by greater than or equal to 10% of their ideal body weight, severe kidney disease, silicosis, gastrectomy, jejunoileal bypass, leukemia, lymphomas or Hodgkin's disease).

- People under certain medical treatments (i.e. corticosteroids) or other immunosuppressive therapy (i.e. bone marrow and organ transplants).
- People with Human immunodeficiency virus (HIV) infection, the virus that causes acquired immunodeficiency syndrome (AIDS)
- Babies and young children and old age.

3.1.3 TB infection versus TB disease

Those with TB infection will not necessarily develop TB disease in the future. Infection is not synonymous with disease because infection does not always lead to injury of the host, even if the pathogen is potentially virulent. In a diseased state the host is harmed in some way, whereas infection refers to any situation in which a microorganism is established and growing in a host, whether or not a host is harmed. A disease is a result of the interaction between a host and pathogen that leads to damage to the host (Madigan *et al,* 1997).

Among generally healthy persons, infection with *M. tuberculosis* is highly likely to be asymptomatic and cannot infect others. They have an inactive form of the TB bacteria in the body, which are kept in check by the immune system. While the bacteria are inactive, it remains alive in the body and can become active later. Data from a variety of sources suggest that the lifetime risk of developing clinically evident tuberculosis after being infected is approximately 10%, with 90% like hood of latent infection (Harries *et al.,* 1998). Only a positive tuberculin skin test indicates the presence of the organism in persons with latent infection. The risk of disease development is increased for those in high-risk groups, especially those with weak immune system. The chance of developing disease is greatest shortly after infection and steadily lessens as time goes on (Harries *et al.,* 1998).

3.1.4 Host defense mechanism

Innate susceptibility to pulmonary infection with *M. tuberculosis* is closely influenced by genetic and ethnic variables that have not been defined. Acquired immunity

following mycobacterial infection usually develops within four to six weeks and is associated temporarily with the onset of delayed hypersensitivity to mycobacterial antigens such as purified protein derivatives (PPD). Successful acquired resistance is mediated by T lymphocytes, which lyses infected macrophages directly or activates them through soluble mediators (e.g. γ interferon) to destroy intercellular bacilli (McMurray, 2001). A role of antibodies in either protection against *M. tuberculosis* or infection or in the pathogenesis of tuberculosis has been generally discounted (Anderson, 1997). However, the available data strongly suggests that the certain antibodies can favorably affect the course of infection if they are present at the time of and site of mycobacterial infection (Chambers *et al.*, 2004). This indicates that it is feasible to generate protective antibodies to *M. tuberculosis.*

It has been long postulated that TB prefers areas with high oxygen tension. When host immune response begins to control TB, the microenvironment that TB is exposed, begins to change. It is thought that the oxygen tension drops in this environment, causing the organism to shift down into a non-replicating stage. It is believed that this ability of organism to survive in this unfavorable environment may be responsible for the ability of the organism to be latent in the host for long periods with the capacity to revive and activate in future (Lauzoro *et al.*, 2000).

3.1.5 Virulence mechanism

The virulence of the tubercle bacilli is due to its resistance to cells and fluids rather than to the production of toxic substances (Grange, 1990). The ability of virulent mycobacteria to survive with in phagocytes and macrophages justifies the designation as facultative intracellular pathogen. Mechanisms of intracellular survival are not clear and may vary from species to species. There are some evidences that *M. tuberculosis* can prevent phagosome lysosome fusion. Other studies have demonstrated that virulent mycobacteria can prevent acidification of phagolysozome, perhaps by modulating the activity of membrane proton pump. In addition some component of mycobacterial cell wall (e.g. cord factors, trihalose dimycolate) may be directly cytotoxic to macrophages.

Although haemolysin and lipase are produced, their role in escape of tubercle bacilli from phagosome and importance of extra vascular organism in pathogenesis are unknown (Mc Murray, 2001).

The important characteristic features of the tubercle bacilli for its virulence include slow growth dormancy, complex cell envelope, intracellular pathogenesis and genetic homogenecity. Novel biosynthesis pathways generate cell wall components such as mycolic acid, mycocerosic acid, phenolthiocerol, lipoarabinomannan, arabinogalactan and several of these may contribute to mycobacterial longevity, trigger inflammatory host reactions and act in pathogenesis (Cole *et al*., 1998).

3.2 Pathogenesis

3.2.1 Primary infection

During airborne transmission, primary infection occurs on first exposure to tubercle bacilli present in a droplet nucleus, generated by an index case. Inhaled droplet nuclei are so small that they avoid the mucobacillary defenses of bronchi and lodge in the terminal alveoli of the lungs.

a) This droplet nucleus is ingested by an alveolar macrophage at the site of implantation, and after a lag period of a few days, the bacillus multiplies intracellularly. The macrophage eventually dies and the bacilli are released and ingested by other macrophages. As this process continues, a primary lesion forms which can be identified after calcification. As the primary lesion enlarges, some organisms are transported to the lymph nodes draining the area containing the primary lesion. The lymph nodes enlarge as the bacilli multiply intracellularly, creating a situation in which bacilli escape from the leaky, swollen lymph node. The term progressive primary tuberculosis is often used to describe disease arising directly from either the parenchymal or the lymph node component of the primary complex, disease which commonly occurs within 3-8 months after tuberculin conversion.

b) Even though the macrophages do not kill the bacteria, they are able to process mycobacterial antigens and present them to circulating T-lymphocytes.

c) The T-cells proliferate and produce lymphokines (especially γ interferon) which effectively enhance the intracellular killing of microorganisms. This immune response kills most of bacilli leading to the formation of granulomas (granuloma is an indefinite term applied to the nodular inflammatory lesions usually small or granular firm persistent and containing compactly grouped mononuclear phagocytes). It is at this point the person will be test positive to the tuberculin skin test. The normal healing process occurs over 6-10 weeks, and results in calcification of the granulomas thereby preventing further spread of the bacilli. The person is then said to have non-infectious TB.

d) In case of susceptible host or when infecting dose is high, the immune system becomes overwhelmed. The lymphocytes will produce cytotoxic substances which along with hydrolytic enzymes from living and dead macrophages will cause caseation necrosis (caseation refers to necrosis with conversion of the damaged tissue into a soft, cheesy substance). When this occurs in the lungs, the caseous material will liquefy, forming cavities in the lung tissue where the bacilli can continue to replicate and greatly multiply in number. This is the first time when bacteria multiply extracellularly to tremendous number, which may not be controlled even in highly immune hosts.

e) The caseum can be discharged into the airways, producing the signs and symptoms of acute pneumonia that are seen in tuberculosis patients. When the bacilli are not contained by the normal healing process, cytokines released from the lymphocytes continue to cause a severe inflammatory response leading to fever, loss of appetite, weight loss, etc. that are characteristic of the "Wasting" typical of advanced tuberculosis disease.

(http://wupa.wustl.edu/record/archive/1997/09-25-97/4066.html)

(http://www.rockefeller.edu/pubinfo/mckinney081700.nr.html)

3.2.2 Post primary TB

Post primary TB occurs following a latent period of months or years, after primary infection. Approximately 10% of individuals with latent infection, progress to active disease. About half of these groups do so in the first 2 years after the initial infection and the remaining half develop active disease at less predictable times during the remaining years of life (Hornick, 1993). Post-primary TB can be caused either by reactivation of bacteria remaining from the initial infection or by failure to control a subsequent reinfection. In most cases, reactivation may be in response to a trigger, such as weakening of immune system. Post-primary TB is predominantly a pulmonary disease. The characteristic features are the extensive lung destruction with cavitations, positive sputum smear, upper lobe involvement and usually no intrathoracic lymphadenopathy (Harries et *al*., 1998).

3.3 The etiological agent

The etiological agent of tuberculosis, *M. tuberculosis* belongs to the genus Mycobacterium. They have traditionally been considered to be rather separate and are thus usually treated as a family Mycobacteriacae. But there is considerable evidence that they are closely allied to the genera *Corynebacterium, Nocardia* and *Rhodococcus* (Wayne and Kubica, 1986). They have been placed in order Actinomycetales and the class Actinomycetes (Rastogi *et al*., 2001).

Mycobacteria are aerobic, non-motile, slow growing, straight or slightly curved shaped bacteria, which are characteristically acid and alcohol fast. Sometime they show branching filamentous form resembling fungal mycelium and hence bear the name Mycobacteria i.e. fungus-like bacteria (Ananthanarayan and Paniker, 1996)

Clinically mycobacteria are classified into two group; those associated with tuberculosis are collectively known as *M. tuberculosis* complex (*M. tuberculosis, M. africanum, M. bovis, M. microtii*). Other mycobacteria that may be associated with human disease are collectively known by different names such as atypical, anonymous, non-tuberculosis,

tuberculoid, opportunistic and mycobacteria other than tuberculosis bacilli, usually abbreviated to MOTT (APPENDIX- V). Many MOTTs are found in the environment but they can also colonize in man, as in the part of a previously damaged respiratory tract and cause clinical infection (Watt *et al.,* 1996).

M. tuberculosis is a non-sporing, non-capsulated, straight or slightly curved rod measuring 1-4 x 0.2-0.5 µm (Cheesbrough, 1989). In sputum and other clinical specimens, they may occur singly or in small clumps, and in liquid cultures, they often grow as twisted rope like colonies termed serpentine cords (Greenwood *et al*., 2002**)**. Tubercle bacilli are aerobes, grow slowly with generation time 12-24 hours, colonies usually appears in 2-3 weeks and may sometimes requires 8 weeks incubation, optimum temperature 37°C and pH 6.4-7.0 (Chakraborty, 2003). The tubercle bacilli are able to grow on a wide range of enriched culture media, but Lowenstein-Jensen (LJ) medium is widely used. The colonies of *M. tuberculosis* on LJ medium are dry, rough, and creamy or buff colored (Cheesbrough, 1989).

3.4 Diagnosis

Diagnosis of active disease includes clinical suspicion, chest radiographs, staining for acid-fast bacilli, culture for mycobacteria, and more recently, nucleic acid amplification assays (Foulds and O'Brien, 1998).

3.4.1 Clinical diagnosis

Clinically, pulmonary tuberculosis is chiefly present with persistent cough for three or more weeks, haemoptysis, shortness of breath and chest pain, loss of appetite and loss of weight, malaise and fatigue, night sweats and fever (Enarson *et al.*, 2000). Symptoms of extra-pulmonary tuberculosis depend on the organ involved (Enarson *et al.*, 1996).

3.4.2 Radiological diagnosis

Chest X-rays (radiology) can also help in the detection of pulmonary tuberculosis but they do not allow etiological diagnosis. There is a radiologic difference between

primary and secondary tuberculosis. Primary disease is usually characterized by a single lesion in the middle or lower right lobe with enlargement of the draining lymph nodes. Endogenous reactivation is often accompanied by a single (cavitary) lesion in the apical region, with unremarkable lymph nodes and multiple secondary tubercles (McMurray, 2001).

3.4.3 Tuberculin skin test

Infection in an asymptomatic individual can be diagnosed with the help of the intradermal PPD skin test. Intradermal introduction of PPD into a previously infected, hypersensitive person results in the delayed (48-72 hr) appearance of an indurated (raised, hard) reaction with or without erythema. It is impossible to distinguish between present and past infection based on a positive tuberculin test (McMurray, 2001). The tuberculin skin test also fails to detect a substantial proportion of persons co infected with HIV and of person with advanced tuberculosis (ATS- CDC, 1981; CDC, 1991; Huebner *et al*., 1993). Its sensitivity for active disease varies considerably from 65% to 94 % (Lalvani *et al*., 2001).

3.4.4 Laboratory diagnosis

3.4.4.1 Microscopy

Microscopy is carried out by utilizing 'Acid Fast' property of mycobacteria. The acid-fast staining procedure depends on the ability of mycobacteria to retain dye even when treated with mineral acid or an acid-alcohol solution (Bloom, 1994**)**. There are several methods of determining the acid-fast nature of mycobacteria. In the carbol-fuchsin (Ziehl-Neelsen) procedure, acid-fast organisms appear red against a blue background, while in the fluorochrome procedures (Auramine-O, Auramine-Rhodamine), the acid fast organisms appear as fluorescent rods, yellow to orange (the colour may vary with filter system used) against a pale yellow or orange background (WHO, 1998). Because fluorochrome-stained smears are viewed under high dry magnification rather than the oil immersion magnification required by fuchsin-stained smear, they may be examined more rapidly and efficiently. Furthermore, with a fluorochrome dye, mycobacteria seen

as staining bright against a dark background are easier to detect, hence smears examined by this method been found to have a greater sensitivity than those stained with fuschin (Strumpf *et al*.,1979). The microscopy method detects 5000-10000 bacteria per ml with sensitivity range between 46-78%. Specificity is virtually 100% depending on the source of the sample, and the *Mycobacterium* involved (Kox, 1996).

3.4.4.2 Culture

Till date, culture is taken as "Gold Standard" method for diagnosis of tuberculosis, requiring only 10 to 100 organisms to detect *M. tuberculosis*. The sensitivity of culture is excellent, ranging from 80% to 93 % and the specificity is quite high at 98 % (ATS-CDC, 2000). Different types of culture media solid media, which include egg-based media (Lowenstein-Jensen, Ogawa medium) and agar based media (Middlebrook 7H10 and 7H11), and liquid media (Middlebrook 7H12 and other broths) for growth of the bacteria. However, solid media are generally in practice as it allows visualization of colony morphology and pigmentation, which is useful diagnostically for distinguishing colonies of *M. tuberculosis* from those of some non tuberculous mycobacteria. Furthermore, identified distinct colonies will be an excellent source for high quality and quantity DNA for standard molecular typing techniques. L-J medium is popular for isolating human strains of *M. tuberculosis* and most other mycobacteria. It contains eggs, glycerol, asparagines, mineral salts and malachite green dye (that inhibits certain contaminating bacteria) (Salfinger *et al.*, 1990). The presence of glycerol in L-J medium improves the growth of *M. tuberculosis*, but not that of *M. bovis*. Ogawa's medium that is also in common use, contains egg yolk instead of whole eggs (Grange, 1990).

M. tuberculosis is an obligate aerobe, grows optimally at 37°C (range: 25°C to 40°C) and pH 6.4 -7.0. It is a slow growing organism with generation time of 14 -15 hours. On solid medium *M. tuberculosis* forms a dry, rough, raised, irregular colony with wrinkled surface. The colonies are creamy white initially, becoming yellowish or buff colored later and tough when picked off. They are tenacious and not easily emulsified. Until, recently, identification of *M. tuberculosis* from positive cultures depend on biochemical

tests for niacin, aryl sulphatase, neutral red, catalase-peroxidase, amidase, and nitrate reductase after the incubation of 2 to 3 weeks (Forbes *et al*., 1998).

3.4.4.3 Molecular methods

Molecular methods rely on extraction, desired/targeted nucleic acid amplification and detection of conserved gene sequences of *M. tuberculosis*. These methods provide rapid detection, identification and characterization of the *M. tuberculosis* strains. Different target sequences have used to confirm the diagnosis of TB (Niemann *et al*., 2000). Several molecular procedures useful for diagnosis of mycobacterial diseases include strand displacement amplification (SDA), polymerase chain reaction (PCR) amplification, transcription-mediated amplification (TMA), reporter phase systems, oligonucleotide ligation amplification and Q-beta replicase amplification (Bloom,1994). More recently, Loop Mediated Isothermal Amplification (LAMP) has been standardized for rapid reaction, simple operation, and easy detection of mycobacteria directly in clinical material within 24 hours or less of specimen receipt (Iwamoto *et al*., 2003).

3.5 Treatment and control

Tuberculosis therapy generally consists of a 6 to 9 month course of isoniazid, rifampicin, streptomycin, thiocetazone, pyrazinamide and ethambutol (Appendix VI). There are three main properties of anti-TB drugs: bactericidal ability, sterilizing ability and the ability to prevent resistance (Maher *et al.*, 1997).

3.6 Epidemiology of tuberculosis

3.6.1 Tuberculosis: global aspect

Tuberculosis is the world's second most common cause of death from infectious disease, after HIV/AIDS. Nearly one-third of the global population (2 billion persons) is infected with *M. tuberculosis* bacilli and is at risk of developing active clinical TB. Worldwide more than 16 million people are suffering from active TB disease (WHO, 2003). There were 8.8 million estimated new cases of TB (all types) in 2002, of which 3.9 million

were smear-positive (infectious type) (WHO, 2004a). Sub-Saharan Africa has the highest incidence rate (290 per 100000 populations), but the most populous countries of Asia have the largest numbers of cases: India, China, Indonesia, Bangladesh, and Pakistan together account for more than half the global burden. Eighty of new cases occur in 22 high-burden countries. Everyday more than 5000 people (approximately 2 million per year) are dying from the disease (WHO, 2003; WHO, 2004b).

3.6.2 Tuberculosis in SAARC countries

Tuberculosis is one of the major public health problems in the SAARC region with immense socio-economic impacts. Almost 50% the adult population of this region has already been infected with *M. tuberculosis* and is at risk of developing tuberculosis disease. In the year 2002, an estimated 2.4 million people newly developed TB disease of which about 1.1 million were smear positive and capable to spread the disease to others. According to this estimate SAARC region was bearing 27.4% of the total global new TB cases (with 22% of population share). India, Bangladesh, and Pakistan are occupying the $1^{st,}$ 5^{th} and 6^{th} position in the list of 22 high burden nations (According To Estimated Incidence of TB: High Burden Countries, 2002) with India revealing the highest (20%) global absolute burden of TB. These three SAARC nations account for 26.7% of the total global new TB cases. Every year about 600,000 people are dying due to this disease. More than 75% of these cases and deaths occur among 15-54 years age group, economically the most productive age group. As a result the social and economic loses due to TB are huge (SAARC 2004).

3.6.3 Tuberculosis in Nepal

Tuberculosis is one of the major public health problems in Nepal. About 45% of total population is infected with TB, out of which 60% are in adult age group. Every year, 40000 people develop active TB, of whom 20000 have infectious pulmonary diseases. Introduction of treatment by Directly Observed Treatment Short course (DOTS) have

already reduced the numbers of deaths, however 5000-7000 people continue to die every year from this diseases (NTCP, 2004/2005).

3.7 Molecular epidemiology

Molecular epidemiology is the integration of molecular techniques to track specific strains of pathogens with conventional epidemiological approaches to understand the distribution of disease in population (Small and Moss, 1993; van Embden *et al*., 1992). In essence, molecular epidemiology focuses on the role of genetic and environmental risk factors, at the molecular/cellular or biochemical level, in disease etiology and distribution among populations. More specifically to infectious diseases, molecular epidemiology attempts to utilize a multidisciplinary approach to identify factors that determine disease causation, propagation/dissemination, and distribution (in time and space). This is primarily achieved by associating epidemiologic characteristics with the biologic properties of clinical isolates recovered from symptomatic individuals (Mathema *et al.,* 2006).

Strain identification can be used as an additional tool in epidemiological investigation in order to gain a better understanding of factors that influence TB transmission, for identification of risk factors of TB transmission in communities and for evaluation of regional control programs permitting a rational design of more adequate control measures. It can help to address important epidemiological questions such as the origin of an infection in a patient's household or community and the spread and early detection of organism with acquired antibiotic resistance (Maguire *et al*., 2002).

3.7.1 Conventional typing

The earliest methods of strain comparison relied upon antibiotic susceptibility patterns. But this approach is very limited since *M. tuberculosis* isolates are quite homogeneous and cannot be differentiated by the routine biochemical tests performed in clinical laboratories or by serological methods. Multiple drug resistance provides reproducible markers for specific strains but this is not generally usable. Until recently, the only available typing method was phage typing (Crawford and Bates, 1984). Phage typing is

cumbersome and lacked sensitivity because of limited number of *Mycobacterium* phage types available. However, the technique has proved to be useful in typing *M. tuberculosis* strains from outbreaks (Snider *et al*., 1984).

More recently, phenotypic markers have been replaced with molecular techniques such as DNA fingerprinting. Use of these novel techniques in conjunction with conventional investigation is providing increased insight in the current epidemiology of tuberculosis (Bloom, 1994).

3.7.2 Molecular typing

3.7.2.1 Genome of *M. tuberculosis*

The establishment of the genome sequence of *M. tuberculosis* H37Rv in 1998 paved the way for major breakthroughs in understanding the biology of tubercle bacilli particulary and mycobacteria in general (Cole *et al*,, 1998). The prokaryotic world now accounts for more than 200 sequences, with the *Mycobacterium* genus being represented by four complete genome sequences, those of *M. tuberculosi*s strains H37Rv and CDC1551, *M. bovis* AF2122/97 and *M. leprae*. {http://www.pasteur.fr/recherche/unites/lgmb. (Cole *et al.,* 2005)}

The sequence of the first complete genome of *M. tuberculosis,* that of the paadism laboratory strain, H37Rv, comprises 4411532 bp (Cole *et al.*, 1998). Its circular chromosome is about the size of that of *Escherichia coli* and is about 10 times larger than the smallest bacterial one, that of *Mycoplasma genitolium* (Thomson *et al.,* 2001). Unlike some environmental mycobacteria, no plasmid has yet been identified among the different numbers of the *M. tuberculosis* complex. It has high G+C content of 65.6%, a parameter associated with an aerobic lifestyle. The last reannotation identified 4043 genes thought to encode 3993 proteins and 50 stable RNAs (Cole *et al.,* 2005)

The most striking feature is the abundance of genes involved in lipid metabolism, 233 to date. Twenty cytochrome p-450 monooxygenases, which may modify lipids, are also

present, none are found in *E. coli.* Besides 56 insertion elements and two prophages, several large multigene families were uncovered (Tekaia *et al.*, 1999). Comparative genomics of attenuated strain *M. tuberculosis* H37Ra, a close relative of H37Rv, identified several regions absent from virulent strain, termed RvD1 to RvD5, that result from IS*6110*-mediated deletion events (Brosch *et al*., 1999) .

3.7.2.2 Repetitive elements

Bacterial isolates belonging to the *M. tuberculosis* complex show an unusually high degree of conservation in housekeeping genes. So the tuberculosis genome is remarkably homogeneous throughout the world with very few silent nucleotide substitutions. Because the mutation frequency of *M. tuberculosis* is similar to that of other bacteria, it is speculated that the *M. tuberculosis* complex isolates presently seen have diverged from a common ancestor not more than 10000 to 15000 years ago. Much more DNA polymorphism in *M. tuberculosis* complex bacteria has been found to be associated with repetitive DNA, such as transposable elements (transposons) and short perfect or imperfect repeats (van Embden *et al.*, 2000). Repetitive DNA is universally present in bacterial and eukaryotic genomes with two major classes, usually distinguished as interspersed repeats, such as mobile genetic elements, and tandem repeats (TRs). Furthermore, in molecular epidemiology of tuberculosis, repetitive DNA elements as genetic markers are categorized into two classes.

3.7.2.2.1 Insertion sequence (IS) elements

Insertion sequences (IS) are simple transposons or interspersed repeated sequences usually less than 2.5kb in size that are widely distributed in most bacterial genomes. These are inherently unstable and have the potential to move within the genome causing many types of rearrangements such as transposition, deletion, inversion and duplication. IS elements are commonly defined as carrying only the genetic information related to their transposition and regulation, unlike transposons, which can also carry genes that encode phenotypic markers(e.g., antibiotic resistance). Transposition of IS elements often causes gene disruptions that can have strong polar effects and in other cases can

lead to the activation or alteration of expression of adjacent genes due to the regulatory sequences, including promoters and protein-binding sequences (Chandler and Mahillon, 2002). From an evolutionary perspective, there are at least two distinct hypotheses explaining the role of IS elements in genomes. One regards the elements as genomic parasites that, on balance, harm their hosts (i.e., bacteria) (Charlesworth *et al.,* 1994). In contrast, others postulate that IS elements are important to their hosts for adaptive evolution, which is maintained by selection of occasional advantageous IS-derived mutations (Blot, 1994).

Table 1 Repetitive DNA sequence in *M. tuberculosis*[a].

Repeated sequence	Host range	Copy number	Polymorphism
IS*6110* (IS 986, IS 987)	*M. tuberculosis*	0-20	Yes
	M. africanum	0-20	Yes
	M. bovis	1-20	Yes
	M. bovis- BCG	1-2	No
S*1081*	*M. tuberculosis*	5-6	Small
	M. africanum	5-6	Small
	M. bovis	5-6	Small
	M. bovis- BCG	5-6	Small
DR cluster	*M. tuberculosis*	1	Yes
	M. africanum	1	Yes
	M. bovis	1	Yes
	M. bovis- BCG	1	No
MPTR[b]	Tuberculosis complex	+/-80	Small
PGRS[c]	Tuberculosis complex	26-30	Yes

a: modified from Poulet *et al.,*1995b
b: also described in *M. kansasii* and *M. gordonae*
c: also described in *M. kansasii*, *M. gordonae*, *M. marinum, M. microti, M. gastri and M. szulgai*

IS can be grouped in families on the basis of genetic structure, similarity between encoded transposable enzymes and fate of target sites (Mahillon and Chandler, 1998). Analysis of *M. tuberculosis* H37Rv genome has revealed 56 loci with similarity to IS elements occupying approximately 77kb. They are classified into the major IS families such as IS3, IS5 or IS21. The most abundant IS element in the genome of *M. tuberculosis* is IS*6110*, belonging to IS3 family. Their number varies from 0 to 25 (Cole *et al.,* 2005).

Depending on the organism in which it was characterized, the sequence is called IS*6110* or IS986 in *M. tuberculosis* and IS987 in *M. bovis*. In *M. tuberculosis*, the IS was independently isolated and sequenced by Gicquel and coworker and named IS*6110*, and simultaneously by Zainuddin and Dale (1989) and designated IS986 (Thierry *et al.,* 1990). Both are the same element with few base pair differences. But as the description of IS*6110* was published first and also recommended by van Embden *et al.* 1993, it is the preferred designation in molecular epidemiology of *M. tuberculosis*. Subsequently insertion sequence in *M. bovis BCG* was sequenced and termed IS987 (Hermans *et al.,* 1991).

IS*6110* is a 1361 bp long sequence that was detected in members of the *M. tuberculosis* complex and difference of only a few nucleotides have been detected between the sequence copies. This sequence is characterized by presence of direct repeats (DR) separated by transposage gene. The number of IS*6110* copies present in the genome is species and strain dependent. Most strains of *M. tuberculosis* carry 8 to 15 copies in different position of the genome, although single copy strains are common (Kanduma *et al.*, 2003).

Insertion sequences have well-defined preferential sites for integration into the mycobacterial genome, so-called hot spots. Recent data indicate that the transposition of the IS*6110* element is non-random and has certain hotspots for integration in the genome. A few hotspots have been identified; three studies have reported the DR region to be one. Three other preferential integration regions for IS*6110* have also been

described, the *ipl* locus, the *dnaA–dnaN* intergenic region and the region between Rv1754c and Rv1762c (Mostrom *et al*., 2002).

3.7.2.2.1.1 IS*6110* as a tool for strain differentiation

As an insertion sequence, IS*6110* has theoretical capacity to transpose. It is to be expected that fingerprint patterns based on this element change over time. In fact, the duplication of 3bp in DNA flanking the element suggests that such transposition is the source of IS*6110*-based RFLP (Mc Adams *et al*., 1990). Laboratory investigation suggests that IS*6110* patterns are relatively stable, even during 6 months of serial passage in liquid media (van Soolingen *et al*., 1991). The vaccine strain BCG has been propagated separately in different laboratories for over decades, nonetheless all 35 strains tested have been found to carry the IS element designated IS987, invariably at a unique chromosomal locus. Only three of these strains have one additional element, all of them at a common chromosomal position (van Soolingen *et al*., 1993). In, an *in vivo* study, when the strains were serially passaged in guinea pigs for 2 months, no changes were detected in pattern, similarly serial isolates from patients with chronic tuberculosis were found to be unchanged or to have patterns that differed by only the addition or loss of a single band (Otal *et al*.,1991; Small *et al*.,1993b).

The changes in fingerprint patterns that have been observed can usually be accounted for by a single transposition event. However, the transposition is low, time dependent and has not been demonstrated in the laboratory. The most important fact is that a spontaneous mutation resulting in drug resistance does not change the fingerprint pattern (Otal *et al*., 1991). Markers which change too rapidly obscure epidemiological links, where those that are too stable infer direct links where they do not exist. Cave *et al.,* 1994 observed the fingerprint pattern for a given patient remained identical or nearly identical despite recovery of the isolates during intervals which ranged from 8 to 54 months.

The IS*6110* transposition half-life (*t*1/2) (the period over which the IS-specific hybridization pattern does not change), taken from sequentially positive culture with sampling intervals ranging from days to months, was estimated to be between 3 and 4 years (de Boer *et al.,*1999). Warren *et al.,* (2002), investigated the stability of IS*6110* banding patterns in serial *M. tuberculosis* isolates collected from patients living in areas of high TB incidence and noted a half-life of 8.74 years when a constant rate of change was assumed. The authors also noted that the high rate of change seen during the early disease phase (*t*1/2-0.57 years), when the mycobacterial replication rate is presumably high, and the lower rate in the late disease phase (*t*1/2-10.69 years), when bacterial doubling times are longer during or after treatment. Therefore, they concluded that the observed IS*6110* stability is strongly influenced by the time between onset of disease and sample collection.

Another investigation of serial patient isolates used deterministic and stochastic simulation models to estimate an IS half-life of 2.4 years for a strain that has 10 IS*6110* copies (Rosenberg *et al.,* 2003). Indeed, IS*6110* transposition, which is a replicative process and half-life, may be heavily dependent on strain-specific in vivo replication rates, host-pathogen interactions, or anatomical properties. It can be assumed that the half-life of the DNA fingerprint depends on regional factors, such as the efficiency of the therapy, the time interval between manifestation of the disease and the start of treatment and possibly also the regional dominance of stable or less stable pathogen strains. Niemann *et al.,* 2000 refer to the stability of the DNA fingerprint in chains of infection which had been established on the basis of medical histories. This approach focuses on the stability of the DNA fingerprint in recently transmitted tuberculosis and *not* on the "intracorporal" constancy of the DNA fingerprint in patients with repeated positive tuberculosis cultures, as in the studies described above. This study revealed a rate of change of only 1.9% per transmission Nonetheless, IS*6110*-based RFLP patterns seem to be sufficiently stable (and polymorphic) for studying TB transmission dynamics at the local or population level and over time. For instance, Lillebaek *et al.,* 2002 used IS*6110* genotyping to demonstrate endogenous reactivation of TB after over 30 years of

latency. However, one study found that banding pattern of isolates collected 90 days apart from same patient were different by 29% of cases (Burgos, 2002).

The variability in fingerprint pattern results from the variable copy number and location of insertion. The results to date indicate that the number of patterns is very large and because IS*6110* is an active insertion element, the number is continuously increasing. The study results indicate that small set of random isolates (<50) from a given community will likely all have unique fingerprint patterns; thus outbreak associated isolates can easily be identified. However, there has not been a sufficiently large study to define the approximate number of patterns in any geographical region. Data of van Soolingen *et al.*, (1991) indicate that the variability of isolates may be greater in the areas that have low incidence if tuberculosis than in areas where tuberculosis is still epidemic. This reflects the reactivation of disease in epidemiologically unrelated individuals without subsequent transmission (low incidence) versus continued spread of a limited number of strains (high incidence). The specificity of the fingerprint pattern is dependent upon the copy number of IS*6110*, i.e., the higher the number, the greater the certainty that two matching patterns represent epidemiologically related strains.

Identical fingerprints from different isolates are usually evidence of recent person to person transmission or endogenous reactivation. As transposition of IS*6110* is a time dependent process, the degree of IS*6110* polymorphism among the descendents of a particular clone in a population is a reflection of the time that has elapsed since their divergence. Thus, analysis of *M. tuberculosis* population structure by IS*6110* associated RFLP may provide information about the evolutionary history and the dissemination of particular clones in a given geographic region (van Soolingen *et al.,* 1995).

However, IS*6110* is the most widely used epidemiological tool in molecular epidemiology. This may not be true for all patterns. Strains (particularly from Asia and Africa) have only single copy of IS*6110*, generally located on a similarly sized Pvu II restriction fragment, thus limiting strain differentiation based on this element. In addition, some strains lack any copies of IS*6110* (van Soolingen *et al.*, 1993). In some

communities low copy number strains (<5 copies) make up to 25% of strains. The band position of low copy number strains show less polymorphism than high copy number and this coupled with the fact there are fewer bands for similarity calculation means that IS*6110* typing is less discriminatory when applied to these strains. In addition, some mycobacteria other than tuberculosis posses multiple copies of sequences that hybridize with the IS*6110* probe and this will produce a pattern (Kanduma *et al.,* 2003). Thus, an alternative secondary typing should be required for such strains.

3.7.2.2.1.2 IS*1081*

The only other transposable element known in *M. tuberculosis* is IS*1081*, a 1324bp sequence discovered by Collins while attempting to clone on *M. bovis* specific DNA sequence. The element is related to the IS256 family of Insertion element originally found in *Staphylococcus aureus*. It is found in all species of *M. tuberculosis* complex as well as in *M. xenopsi.* In contrast to IS*6110*, IS*1081* is distributed rather homogeneously in different *M. tuberculosis* strains with 5 to 7 copies frequently present in the same sites. Therefore, IS1081 is not a useful marker for distinguishing between strains of *M. tuberculosis*. However, because IS*1081* is present in characteristic PvuII fragments in *M. tuberculosis*, it may be used to differentiate BCG from other strains of *M. tuberculosis* complex (van Soolingen *et al.*, 1992).

3.7.2.2.2 Short repeats

These are short perfect or imperfect repeats, varying in size from three to 36 base pairs.

3.7.2.2.2.1 Direct repeats (DR)

The DR region is a polymorphic locus in the genome of *M. tuberculosis*, which comprises a cluster of directly repeating sequences of 36 bp, separated by unique spacer sequences of 36 to 41 bp. The number of copies of the DR sequence has been determined to be approximately 10 to 50 in a variety of *M. tuberculosis* complex strains. One repeat sequence and the following spacer sequence together have been termed a

DVR (Direct Variable Repeat). The number of DVRs varies from strain to strain, allowing this locus to be used as a genetic marker to differentiate strains. Two mechanisms have been proposed for the polymorphisms at this locus; homologous recombination between adjacent or distant DVRs may lead to deletion of one or more discrete DVRs and IS*6110* transposition, as the DR region in *M. tuberculosis* has been identified as a hot spot of integration of the IS*6110* (Groenen *et al.,*1993; Hermans *et al*.,1991). Because there appears to be considerable polymorphism within small parts of the chromosome, this region is well suited for PCR-based fingerprinting techniques (Bloom, 1994).

3.7.2.2.2.2 Polymorphic GC-rich repetitive sequence (PGRS)

The polymorphic GC-rich repetitive sequence (PGRS) is a short sequence present in multiple chromosomal clusters composed of many nonperfect repeats, like the major polymorphic tandem repeats (MPTRs). Though initially identified as *M. tuberculosis* complex, its host range is now known to include *M. kansasii, M. gastr*i and *M. szulgae*, a host range that resembles that of the MPTR (Doran *et al.,* 1992). In *M. tuberculosis*, PGRS appears to be present in at least 26 copies, varying in number and distribution from strain to strain. Polymorphism in PGRS has been harnessed for typing and a recombinant plasmid pTBN12 containing the GC-rich consensus sequence as a probe has been used for secondary fingerprinting of *M. tuberculosis* with absent or low copies of IS*6110* (Kanduma,2003).

7.2.2.2.3 Major polymorphic tandem repeats (MPTRs)

Major polymorphic tandem repeats (MPTRs) are 10-bp repeating sequences invariably separated by unique 5-bp spacer sequences. These sequences have been identified in *M. tuberculosis* complex as well as in a number of other mycobacteria (Hermans *et al.,* 1992). The 10-bp repeat is heterogeneous in sequence, with at least five bases conserved. Self-hybridization experiments suggest that this sequence is the major repetitive DNA in *M. tuberculosis*. Because this MPTR sequence is partially homologous to the repetitive extragenic palindromic sequence and the recombination

signal Chi of *Escherichia coli*, it is speculated that the MPTR sequence might play a role in genetic rearrangements and gene regulation. There are over 80 copies of this element in the genome of *M. tuberculosis* and not restricted to organism belonging to *M. tuberculosis* complex (Hermans *et al.,* 1990). MPTR can be used as an epidemiological marker for pathogenic mycobacteria species (Kanduma *et al.,* 2003), but has limited polymorphism.

3.7.2.3 Genotyping techniques

The discovery of transposable elements and other repetitive DNA elements in *M. tuberculosis* complex strains in early 1990s has lead to various DNA fingerprinting methods to differentiate strains belonging to complex. Some of these methods are based on the amplification of repetitive DNA sequences using the polymerase chain reaction (PCR). Other methods visualize restriction fragment containing particular repetitive DNA elements like IS elements. A comparison between PCR-based-and RFLP-typing methods revealed that IS*6110* RFLP is the most discriminative and reproducible typing method currently available. Because IS*6110* RFLP typing has been internationally standardized with regard to the choice of restriction enzyme, probe, and size markers, interlaboratory comparison is enabled by computerization of DNA fingerprints (van Soolingen *et al.,* 1999).

3.7.2.3.1 Criteria of typing system

Several criteria are proposed for evaluating the performance of typing systems (Struelens, 1998). These criteria include typeability, reproducibility, stability, discriminatory power, and epidemiologic concordance.

Typeability refers to the proportion of isolates that can be scored in the typing system and assigned a type, ideally all isolates.

Reproducibility refers to the ability of the typing system to assign the same type on repeat testing of the same strain.

Stability is the biological feature of clonally derived isolates to express constant markers over time and generations. The stability of markers may be acceptable even in the presence of variation, provided that the typing system enables recognition of clonal relatedness and does not lead to misclassification of subclonal variants as epidemiologically unrelated.

Discriminatory power is a key characteristic of typing systems, because it conditions the probability that isolates sharing identical or closely related types are truly clonal and part of the same chain of transmission. Discriminating power can be calculated based on Simpson's index of diversity. Ideally, the index, based on testing a large number of epidemiologically unrelated isolates, should equal one. In other words, each independent isolate should be sufficiently different to be assigned to a distinct clone. In practice, a typing system or combination of systems, displaying a discrimination index greater than 0.95 is acceptable. This level of discrimination corresponds to a 5% probability of erroneously assigning independent isolates to the same clone.

Epidemiologic concordance is the capacity of a typing system to correctly classify into the same clone all epidemiologically related isolates from a well-described outbreak. Additional comparative studies are needed to establish the relative value of systems currently used for typing microbial pathogens (Struelens, 1998).

Kremer *et al.*, (2004) compared 5 different methods of RFLP typing which employed IS*6110*, IS1081, PGRS, the DR and the GTG5 repeat as probes. Of the PCR- based methods compared, VNTR typing, mixed-linker PCR and spoligotyping were highly reproducible between different laboratories. The double repetitive PCR (DRE-PCR), IS*6110* inverse PCR, IS*6110* ampliprinting and arbitrarily primed PCR were not reproducible. Despite the development of different typing methods, RFLP using IS*6110* is being widely used and considered the Gold standard to which other methods are compared (Bifani *et al*., 2002). Thus, implementation of multiple molecular techniques in a single study provided better discrimination between strains and insight for phylogenetic groupings (van Soolingen *et al.,* 1991). Today, most of the molecular

epidemiologic studies rely on IS*6110* RFLP typing and a secondary typing method such as PGRS or spoligotyping for isolates with less than six bands in the IS*6110* RFLP band pattern.

3.7.2.3.2 IS*6110* RFLP-fingerprinting

A restriction fragment length polymorphism **(**RFLP) is defined as the variation(s) in the length of DNA fragments produced by a specific restriction endonuclease from genomic DNAs of two or more individuals of a species. RFLPs are visualized by digesting DNA from different isolates with a restriction enzyme, followed by gel electrophoresis to separate fragments according to size, then blotting and hybridization to a labeled probe that identifies the locus under investigation. RFLPs were the predominant form of DNA variation used for linkage analysis until the advent of PCR. Even now, in the PCR age, RFLPs provide a convenient means for turning an uncharacterized DNA clone into a reagent for the detection of a genetic marker. The main advantage of RFLP analysis over PCR-based protocols is that no prior sequence information, nor oligonucleotide synthesis, is required.

Good resolution of genomic RFLP analysis is obtained by either transfer of restriction fragments onto membranes, followed by Southern-blot hybridization with DNA probes or use of endonucleases that have infrequent (<30) recognition sites in the chromosome, followed by separation of these macrorestriction fragments by pulsed-field gel electrophoresis (PFGE). Different types of nucleic acid probes are used for typing: (i) genes encoding metabolic, virulence or resistance functions (ii) multicopy elements, including insertion sequences and transposons, and (iii)rRNA or rDNA sequences (ribotyping). Careful selection and optimization of probe sequence, restriction endonucleases digestion, electrophoresis and hybridization conditions need to be developed for each species or pathovar to be typed. These techniques are not rapid and required specialized reagents and expertise. International standardization of technique, reagents, type strains and nomenclature have been established by public health reference laboratories for IS*6110* RFLP-fingerprinting of *M. tuberculosis*, which integrates

standard computer analysis of patterns and a common database, and is now widely applied for large-scale surveillance of tuberculosis (Bauer *et al.* 1998).

3.7.2.3.3 PCR-based fingerprinting

The RFLP typing requires a well grown culture for DNA extraction. The time lag between isolation of *M. tuberculosis* and growth of mycobacterial culture is often too long to inform patient care or outbreak investigation. PCR-based methods have the advantage of typing *M. tuberculosis* directly in clinical samples, increasing the speed of identification of the organism. They can also be used for non-viable isolates. Some of theses methods, however, lack reproducibility or have less discriminatory power than IS*6110*-RFLP (Kremer *et al.,* 1999).

3.7.2.3.3.1 IS6110-based PCR methods

Many PCR based typing assays have been developed in the recent past based on IS*6110* as the target. Ligation mediated PCR, mixed linker PCR, Hemi-nested inverse PCR, IS*6110* inverse PCR, IS*6110* ampliprinting and double-repetitive (DR) element PCR are among the techniques developed to date. Amplityping is based on the used of oligonucleotide primers hybridizing with ends of IS*6110* and generating a PCR reaction directed away from the insertion sequence. The method is not suitable for comparison of large number of non-specific amplification but can be used for suspected outbreaks (Yuen *et al*., 1995). The mixed linker PCR typing uses a primer complementary to IS*6110* and a second primer complementary to a linker ligated to genomic DNA digested with a restriction enzyme. This typing generates more bands and can be applied directly on smear positive clinical specimens. In interlaboratory comparison of discriminatory power and reproducibility, it performed well (Kremer *et al.,* 1999). Hemi-nested inverse PCR analysis of IS*6110* integration sites based on amplification of a part of the IS*6110* sequences together with its flanking sequences has been developed (Kanduma *et al*., 2003). It is technically simple and has excellent discriminatory power comparable with standard RFLP methods.

Double repetitive element PCR based on amplification of IS*6110* and PGRS generating a banding polymorphism because of distance between these elements has been introduced by Friedman *et al*., (1995). It has predictive value of 96% and DNA patterns seem to be sufficiently stable to use the method for epidemiology. As these methods are based on IS*6110*, their application is limited to the strains containing more than five copies of IS*6110*.

3.7.2.3.3.2 16S- and 23S rRNA-based methods

Amplification of the spacer region between the genes coding for 16S and 23S rRNA and posterior digestion of the amplicon with restriction enzymes has also been also performed for differentiation of *M. tuberculosis* strains (Abed *et al.,* 1995a). Using this amplification system a better discrimination was obtained upon random amplified polymorphic DNA analysis (RAPD) of the amplified product (Abed *et al*., 1995b). This generates patterns that can be easily analyzed and seem to have high discriminatory power but reproducibility and the final discriminatory power of RAPD-based method was found to be limited (Frothingam 1995; Glennon and Smith1995).

3.7.2.3.3.3 DR region-based method

DR region-based method is based on detection of DNA polymorphism in the DR cluster and is called "Direct Variable Repeat Polymer Chain Reaction" (DVR-PCR)(Groenen *et al.,* 1993). The method is based on outward amplification of IS*6110* and the DR, generating as such a strain specific banding pattern upon hybridization with a DR-probe. It has good differentiating power when limited numbers of strains are being tested, but stability of DR region is higher than that of IS*6110* thus showing identity in otherwise different strains, differentiated by IS*6110*-RFLP.

Spoligotyping is based on amplification of the DR-region and subsequent differential hybridization of the amplified products with membrane bound oligonucleotides, complementary to the variable spacer regions localized between the DRs (van Soolingen *et al.* 1995). The presence of these spacer sequences is variable in different *M.*

tuberculosis strains and is visualized by a spot on a fixed site of the membrane. The differentiating power of spoligotyping is less than traditional fingerprinting when high IS*6110* copy number strains are analyzed; in the case of strains with few IS*6110* elements, discrimination is better achieved with spoligotyping.

3.7.2.3.3.4 Minisatellite-based methods

Methods based on minisatellites that contain variable numbers of tandem repeats (VNTRs) have been demonstrated to be effective and portable methods for typing *M. tuberculosis.* Supply *et al.,* (2000) have identified 41 such loci in the *M. tuberculosis* genome and termed them mycobacterial interspersed repetitive units (MIRUs). Twelve loci were demonstrated to vary in tandem repeat numbers and, in most, sequence between repeat units. The MIRU-VNTR typing has discrimination similar to that of high IS*6110* copy number strains and better for low copy number strains. The method can be automated for large-scale typing projects using high throughput sequencing apparatus, so also suitable for global epidemiological surveillance of tuberculosis (Kanduma *et al*., 2003).

3.7.2.4 RFLP as a molecular tool for different studies

IS*6110* based RFLP has been used in a number of studies in developed countries. Some of them important for developing or endemic countries are described below.

3.7.2.4.1 An outbreak study and transmission dynamics

Outbreak situation usually involves person-to-person spread or simultaneous infection from a common source. By definition, all isolates involved in outbreak of an infection would be expected to be clonal. Non-clonality, which is often easier to determine, eliminates an isolate from consideration in a specific chain of transmission.

To understand the transmission dynamics of tuberculosis in Tarrant County, Texas, Weis *et al*., (2002) performed a population-based study of 159 patients with culture-

proven tuberculosis, combining RFLP analysis with prospective interviewing to identify epidemiologic links between patients. Patients whose isolates had identical or closely related RFLP patterns were considered a cluster. Of 159 isolates, 83 patients were nonclustered and 76 patients were clustered in 19 clusters, varying in size from two to 32 patients. The clustering was significantly more common in patients 31–40 years old. When they assumed that one patient in each cluster was the index case and that the other cases in the cluster resulted from recent transmission, 36% (57 patients) of the morbidity due to tuberculosis in Tarrant County resulted from recent transmission during the study period.

In a prospective, population-based study in the province of North Holland, The Netherlands, van Deutekom **(**2004), combined molecular methods with highly detailed epidemiologic information to determine why many clustered cases are not detected at an early stage. Of 481 patients, 138 (29%) fell into 43 clusters, suggesting recent transmission in 20%. Of 155 patients in clusters occurring within 2 years before or after the diagnosis of the disease, 21 (14%) had no epidemiologic links with other patients. Of 47 patients with a clear epidemiologic link, 37 (24% of 155) were identified early, e.g., by contact tracing, and 10 (6%) were missed. In 85 (55%) patients, an epidemiologic link was likely but undetected when using conventional contact tracing. Their results indicate that 86% of clustered study patients had epidemiologic links and that opportunities for earlier identification using conventional tuberculosis control strategies are limited.

Narayan *et al*., 2004 used molecular and conventional epidemiologic techniques to study the mechanisms and risk factors for tuberculosis transmission in a rural area with high prevalence in South India. RFLP analysis with IS*6110* and direct repeat probes was performed with 378 *M. tuberculosis* isolates. Forty one percent of the isolates harboured a single copy of IS*6110*. Of 378 patients, 236 had distinct strains; 142 (38%) shared a strain with other patients, indicating recent infection. Their findings suggested that the majority of the tuberculosis cases in South India were due to reactivation.

Vijaya bhanu *et al.,* (2004) used multiple genetic markers, IS*6110*, DR sequence, and PGRS sequence to support the circumstantial epidemiologic link between eight strains of *M. tuberculosis*, suggesting their possible involvement in small-scale, interpersonal transmission of tuberculosis in PTB patients admitted to two adjacent wards of a tuberculosis hospital in Delhi, India. The result of their study emphasizes the need for more comprehensive investigation of high-risk situations for tuberculosis transmission and long-term follow-up analysis for identifying such instances of unsuspected transmission. However, the use of multiple molecular typing markers and techniques unequivocally identified the exact clonality of strains isolated from the hospital.

Palmero *et al*., 2003 studied the initial multidrug-resistant (MDR) tuberculosis (TB) in HIV-negative patients treated at a Buenos Aires referral hospital Argentina from 1991 to 2000 using IS*6110* DNA fingerprinting. Hybridization patterns were compared visually. Isolates from 49 of the 79 patients with initial MDRTB were available for DNA typing. Of these 49, thirty-six fitted in six molecular clusters with RFLP patterns of six or more bands. The M strain was responsible for the largest cluster involving 24 of the cases. All patients in the M cluster lived and were treated in same city area. The second largest cluster occurred in four members of a family who had immigrated to Argentina in 1999. The remaining four clusters consisted of two cases each. The authors concluded that the hospital-related disease attributable to the M strain was detected in 15 cases from seven different health institutions in the metropolitan area. The M strain associated disease was significantly more frequent among hospital-exposed cases when compared with patients who acquired the disease elsewhere.

3.7.2.4.2 Exogenous infection versus endogenous reactivation

Post-primary TB, which occurs many years after a primary infection, may develop as the result of reactivation of the endogenous primary infection or because of a recent exogenous infection. In this era of effective treatment regimens, the notion that multiple episodes of TB in one patient are almost always caused by endogenous reactivation may be questioned. It is now in practice to characterize the genotype of *M. tuberculosis* by

DNA fingerprinting, which can show whether a new episode of the disease is caused by infection with the same strain that caused a previous episode or by a different strain.

RFLP studies conducted in Hong Kong by Das *et al*., (1993) showed that the patterns of 88 % of the isolates from patients with relapses matched those for their pretreatment counterparts indicating a high frequency of occurrence of infections caused by endogenous reactivation of *M. tuberculosis*.

A study conducted at the Tuberculosis Research Centre, Chennai on pre and post treatment isolates by IS*6110*-RFLP analysis indicated 50% of the isolates ere identical by IS*6110* (Das *et al*.,1995). A high degree of endogenous reactivation among patients who have relapses after successful completion of chemotherapy was observed.

3.7.2.4.3 Geographical distribution and dissemination of tuberculosis

A collection of closely related strains derived from a common ancestor can be classified as a group and, over time, progeny of such a group can disseminate into new populations (clonal expansion) giving rise to new groups of related strains. The identification of independent genetic markers common to members of multiple groups provides a framework to classify such groups more broadly into a family structure (Bifani *et al*., 2002). Thus, individual members of a family are closely related on the molecular level. They are believed to be the progeny of a single ancestral *M. tuberculosis* cell that has undergone clonal expansion, and the members form a large branch in the *M. tuberculosis* phylogenetic lineage. Because the changes in the genome of *M. tuberculosis* are time dependent, the degree of polymorphism among progeny of a particular strain is a reflection of the time that has elapsed since their divergence (van Soolingen *et al*., 1995). Analysis of the population structure of strains of *M. tuberculosis* may therefore provide information about the evolutionary history and dissemination of particular strain(s) in the population, or information about the evolutionary history and dissemination of particular strain(s) in populations in different geographical regions in the world (Bifani *et al*., 2002).

In a multinational study performed by Park *et al*., (2000) a total of 422 *M. tuberculosis* isolates from eight countries in the Western Pacific Region were subjected to IS*6110* and IS*1081* DNA fingerprinting to characterize *M. tuberculosis* strains from each country. They followed the standard protocol of fingerprinting and found Chinese, Mongolian, Hong Kong, Filipino, and Korean isolates had comparatively more copies of IS*6110* (proportion with eight or more copies; 95% ± 5%), while Thai, Malaysian, and Vietnamese isolates had fewer copies (proportion with eight or more copies, 60% ± 4%). Filipino isolates had peculiar IS*6110* RFLP patterns, with many common *Pvu*II DNA fragments, especially those of 4.4, 2.8, 2.3, and 2.0 Kb. Similarly a 1.3-Kb single-copy IS*6110* fragment was found only in Vietnamese *M. tuberculosis* isolates. Thus, they concluded *M. tuberculosis* isolates from each country had comparatively similar characteristics depending on the classification factor but each country's isolates showed characteristic DNA fingerprints and differed slightly from the isolates from the other countries in either the mode number of IS*6110* copies or the distribution of IS*1081* types.

A study based on IS typing claimed that *M. tuberculosis* strains from regions in Central Africa, where tuberculosis is highly endemic, are generally related to each other than isolates from the Netherlands, where the transmission rate is slow and the majority of TB cases are presumed to be the result of reactivation of latent tuberculosis infection, LTBI (Casper *et al*., 1996).

Analysis of the population structure of *M. tuberculosis* strains from the People's Republic of China showed that 85% of the isolates were strains with more than 66 % similarity among their IS*6110* RFLP patterns. Because 89% of these strains originated from patients in Beijing, the authors designated that grouping the **"Beijing family"**. (van Soolingen *et al*., 1995).

Ravins *et al.,* 2001 used RFLP analysis with IS*6110* and DR probes to study 69 *M. tuberculosis* isolates obtained from Israeli patients and new immigrants from the former Soviet Union and Ethiopia. DNA fingerprinting identified unique patterns for almost all

isolates, indicating that most patients were infected with a unique strain imported from their country of origin and that their latent infection was reactivated in Israel.

Chauhan *et al*., (2004) studied the IS*6110* polymorphism among isolates of *M. tuberculosis* from Agra region of India, from patients attending the clinics at Agra on the basis of geographical distribution of patients. DNA from those biochemically identified isolates were purified by a physicochemical procedure, restricted with PvuII enzyme and hybridized with PCR amplified and DIG labeled 245 bp IS *6110* probe. Among the 60 isolates taken in study, 5 had no copy of IS *6110*, 8 had 1-4 copies and 47 had multiple copies of IS *6110*. The study discriminated 80% of typed strains with multiple copies. However, as RFLP profiles did not correlate with geographical areas, contacts or the resistance pattern of the strains, the authers recommended that a larger number of isolates from defined geographical areas need to be studied to understand the molecular epidemiology of tuberculosis in that region.

Douglas *et al*., 2003 compared three molecular typing methods, IS*6110-RFLP*, spoligotyping, and DNA sequencing of the *oxyR*, *gyrA*, and *katG* loci, in 48 *M. tuberculosis* strains obtained from patients living in metropolitan Manila, Philippines. The result showed IS*6110* RFLP patterns of the *M. tuberculosis* strains from the area exhibited a high degree of similarity, yet none of them was identical. Of the 40 isolates for which RFLP patterns were available, thirty-eight (95%) showed similarities of 80% or greater among the patterns. Thus, they designated these 38 isolates “The Manila family”. Of the two remaining RFLP patterns, one exhibited more than 75% similarity to those of the Manila family and the other was very distinct from the other patterns, exhibiting only 40% similarity. Comparison with the international database proved that the latter strain had the Beijing genotype of *M. tuberculosis.* The uniqueness of the Manila family was established by comparing the 38 IS*6110* RFLP patterns with an international database.

Lockman *et al*., (2001) performed a population-based DNA fingerprinting study among TB patients in Botswana, using IS*6110* RFLP. DNA fingerprints with greater than five

bands were considered clustered if they were either identical or differed by at most one band, while those with 5 bands were considered clustered only if they were identical. Patients with matching fingerprints were reinterviewed. TB isolates of 125 (42%) of the 301 patients with completed interviews and DNA fingerprints fell into 20 different clusters of two to 16 patients. HIV status was not associated with clustering. Twenty six (11%) of 243 patients overall and 26 (25%) of 104 clustered patients shared both a DNA fingerprint and strong antecedent epidemiologic link. Most of the increasing TB burden in Botswana may be attributable to reactivation of latent infection.

Blackwood *et al*., (2003) used IS*6110* as molecular epidemiological tool to describe the demographic and geographic distribution of tuberculosis in Manitoba, Canada. From 855 patients with tuberculosis in between 1992-1999 and their contacts, recovered isolates were typed by IS*6110* RFLP. Bivariate and multivariate logistic regression models were used to identify risk factors involved in clustering. The result shows clustering was observed among the Canadian born treaty Aboriginal subgroup in contrast to the foreign born. The dominant type, designated fingerprint type 1, accounted for 26% of total cases and 75% of treaty Aboriginal cases. Statistical models revealed that significant risk factors for acquiring clustered tuberculosis are gender, age, ethnic origin and residence. Those at increased risk are males ($p < 0.05$); those under age 65 ($p < 0.01$ for each age subgroup); treaty Aboriginals ($p < 0.001$), and those living on reserve land ($p < 0.001$).

Nguyen *et al.,* (2003) described the results of an integrated traditional and molecular epidemiology study of all culture-positive *M. tuberculosis* cases in the Arctic Inuit communities of Quebec Canada, from 1990 until 2000, using IS*6110* RFLP. On analysis of the results, 42 out of 46 isolates showed highly conserved patterns, likely derived from the same original epidemic. On performing spoligotyping and MIRUs-VNTR of those isolates, 76% (35 of 46) of TB cases were clustered (six clusters, median size four cases) and estimated that at least 63% of TB cases were due to ongoing transmission. Thus the results suggested a potentially reduced utility of IS*6110* RFLP alone in tracking ongoing transmission. By integrating the epidemiologic and genotyping data,

the authors observed the genotyping clustering results were concordant with recognized epidemiologic links but most notably identified previously unrecognized inter village transmission.

In a study performed by Dolzani *et al*.,(2004) in the northeast of Italy, 51 consecutive isolates of *M. tuberculosis*, collected during a 2-year period, were subjected to IS*6110* RFLP analysis to detect the presence of clusters and assigned to one of the three genotypic groups delineated by single nucleotide polymorphisms in the genes *katG* and *gyrA*. The probe was obtained by amplification of a fragment of the insertion sequence IS*6110* with primers INS-1 and INS-2 using genomic DNA from strain H37Rv as a template, and labelled by incorporation of digoxigenin-11-dUTP during amplification. DRE-PCR was used as a secondary typing method to confirm the composition of the clusters obtained by IS*6110*-RFLP. Fourty two different patterns were identified among the 51 isolates tested. Copy number of IS*6110* ranged between 5 and 15 copies per isolate, with the exception of one isolate, which had only two copies of IS*6110*. Thirty six of 51 isolates (70 %) showed a unique pattern, while the remaining 15 isolates were divided into six clusters (named A-F), which included two to four isolates. DRE-PCR analysis was also in agreement with the results.

Toungoussova *et al*., (2002) collected 119 strains of *M. tuberculosis* isolated from PTB patients in the Archangel Oblast, Russia, in 1998 and 1999, and typed these by IS*6110* - RFLP and spoligotyping. The result demonstrated that 53 (44%) of the strains belonged to the Beijing genotype, with a significantly higher rate of resistance than *M. tuberculosis* strains of other genotypes circulating in the region. In particular, 43% of the strains of the Beijing genotype were multidrug resistant; in contrast, only 11% of the other strains were. Of the strains of the Beijing genotype, 93% were part of a cluster, while only 33% of the remaining strains were clustered. Analysis of the medical records of the patients demonstrated that individuals infected with a strain of the Beijing genotype were significantly more likely to be alcohol abusers and to have chronic obstructive pulmonary disease prior to the tuberculosis diagnosis.

Fujikane *et al.,* (2004) performed RFLP analysis on *M. tuberculosis* strains from patients with pulmonary tuberculosis from April 1999 to March 2002 in the north Hokkaido district, Japan. Of 229 patients, strains from 227 (99%) were available for analysis. Among them, 207 patients had six or more copies of IS*6110*. Of these, 16 (8%) had identical patterns and grouped in eight clusters, 120 (58%) belonged to five groups that had similar patterns (Dice coefficient >0.7) and 80 (39%) belonged to the 35 groups with the most common patterns (Dice coefficient >0.9).

Dale *et al.,* (1999) applied IS*6110-RFLP* typing to *M. tuberculosis* isolates from all parts of Malaysia. Strains related to the Beijing family were found to be common in Peninsular but were less common in Sabah and Sarawak where a distinct group of strains was common, comprising nearly 40% of isolates. Single-copy strains constituted nearly 20% of isolates from the peninsula but were virtually absent in East Malaysia. The degree of clustering was found to be increased with patient age, suggesting that reactivation may contribute to clustering. Identical banding patterns were also obtained for isolates from widely separate regions. Thus, the authors concluded that use of clustering as a measure of recent transmission must be treated with caution.

Palittapongarnpim, *et al.,* (1997) studied 211 isolates of MTB received from hospitals in central Thailand by RFLP with IS*6110* probe. They found one non-hybridized isolate, 43(20%) single banded with hybridized fragment of either 1.45kb or 5.0kb, 80(38%) Beijing family, 29(14%) Nonthaburi group.Of the 58(28%) that had more than one hybridized band, none belonged to other groups and were referred as Heterogenous(H) type. Fifty percent of this group had less than 6 bands. The clustering rate was found to be 27% in the study.

Rijal, (2006) used molecular characterization of MTB isolates from Nepal for better understanding of multi drug resistance patterns in the country. Of the 31 stains characterized by IS*6110*-RFLP, 12(39%) were Beijing type, 8 strains (26%) were single banded, 4 strains (13%) were Heterogeneous (H), 7 strains (23%%) different heterogeneous(h) type which were not matching to previously typed strains and were

labeled heterogeneous (h) Nepal type. Of the 12 Beijing family 5(42%) were MDR.. Similarly of 8 single banded strains 3 (37.5%) and of the 7 heterogeneous (h) Nepal type 2(27%) were MDR. In contrast MDR was not observed in heterogeneous (H) type.

3.7.2.5 Well characterized clinical isolates

Recently, molecular strategies to distinguish naturally occurring clones of MTB have focused attention on strains thought to be highly virulent, such as strain CDC1551, or very prevalent, such as C strain, W or Beijing strain family, subpopulations of the W family–like strain 210, and New Jersey W4 strain (Kato-Maeda, *et al*.,2001).

Strain CDC1551 is a recent clinical isolate which was responsible for an outbreak of TB in a rural community of USA with low risk of tuberculosis between 1994 and 1996. On average, 80% of the contacts of patients with disease caused by this strain had tuberculin skin test reactivity. In addition, evaluation of the growth of CDC 1551 in the lungs of mice 20 days after aerosol infection gave 100-fold higher numbers of bacilli compared to the numbers of bacilli isolated from the lungs of mice infected with the *M. tuberculosis* laboratory strain Erdman. As a result of these findings, CDC 1551 was assumed to be highly virulent. (Betts *et al.,* 2000)

MTB 210 is a widely distributed strain, characterized by a 21-band fingerprint pattern when probed with IS*6110*. Currently, the 210 strain has been shown to account for 215 cases of tuberculosis in five states of the U.S.A. To determine whether the dissemination of this strain correlated with its capacity for replication, the intracellular growth rate of strain 210 in human macrophages was measured. Compared to isolates from other clusters or from non-clustered patients, strain 210 grew significantly faster. The data indicated that the 210 strain should also be considered a W variant (Kato-Maeda,*et al.,* 2001).

The C strain is a drug-susceptible organism prevalent in New York City that was responsible for a large outbreak in homeless men's shelters in Manhattan. This strain caused more than 20% of the drug-susceptible tuberculosis cases in New York and was

very frequently isolated from users of injected drugs. In vitro studies demonstrated that this strain is particularly resistant to reactive nitrogen intermediates (RNI), which arise from nitric oxide (NO) following the induction of cytokines in macrophages. Friedman *et al.* have postulated that injection drug users have high levels of NO and that the prevalence of this clone in such patients results from its capacity to resist NO-mediated clearance (Kato-Maeda, *et al.,* 2001).

W or Beijing strain family

The groups of *M. tuberculosis* strains collectively known as "the Beijing family" is the most dominant family of *M. tuberculosis* strains characterized worldwide (van Soolingen *et al.,* 1995). The "W-strain family", concurrently identified on the American and Asian continents, represents the same genotype family (Kremer, *et al*., 2004).

The traditional view is that different strains of *M. tuberculosis* are equally virulent. The first Beijing report suggested, however, that the Beijing strains may possess selective advantages compared to other strains, and that they might be expanding aggressively to other countries and continents (van Soolingen *et al*., 1995). Subsequently, Beijing strains have been associated with (multi)drug resistance, and with specific pathogenic properties and increased virulence. It has also been suggested that the transmission potential of Beijing strains is likely to be enhanced, as compared with that of other strains, because Beijing strains are more readily aerosolized, can establish infection more effectively, and/or progress more frequently from infection to disease (Lillebæk, 2005).

Individual level data on >29000 patients from 49 studies in 35 countries were combined to assess the Beijing genotype's prevalence worldwide, trends over time and with age, and associations with drug resistance. The European Concerted Action group (CA), (2006) found 4 patterns for Beijing/W genotype tuberculosis (TB).

1) Endemic not associated with drug resistance (high level in most of East Asia, lower level in parts of the United States).
2) Epidemic, associated with drug resistance (high level in Cuba, the former Soviet Union, Vietnam, and South Africa, lower level in parts of Western Europe).
3) Epidemic but drug sensitive (Malawi, Argentina).
4) Very low level or absent (parts of Europe, Africa).

This study confirmed that Beijing/W genotype of MTB is an emerging pathogen in several areas and a predominant endemic strain in others; it is frequently associated with drug resistance (European Concerted Action, 2006).

All members of the W strain family and the Beijing strain belong to principal genetic group 1 of MTB and share characteristic genetic fingerprints. They have an insertion of IS*6110* in the dnaA-dnaN locus. Spoligotyping was taken as the gold standard. Presence of at least three of the last nine spacers (no. 35-43) and none of the spacers 1 to 34 is considered to be 100% specific for the Beijing family (Kremer *et al*., 2004). However, worldwide, the RFLP IS*6110* method is the most widely applied DNA fingerprinting technique for strains of *M. tuberculosis*, and by 2001, on estimated 50000 strains had been genotyped by this method (Bifani *et al*., 2002). Therefore the CA project group also explored how to identify Beijing strains on the basis of their IS *6110* RFLP. A set of 19 "IS*6110* Beijing reference strains" were defined using a representative sample of strains of *M. tuberculosis.* These reference patterns gave a sensitivity of 98.1% and a specificity of 99.7% for identifying Beijing strains (defined by spoligotyping) in an international database of 1084 strains (Kremer *et al*., 2004).

3.7.2.6 Principle of TB RFLP

Isolation of genomic DNA

IS*6110*-RFLP requires a large quantity (>4.5µg) of high quality, high molecular weight (intact) DNA. The DNA extraction protocols standardized for this purpose are based on chemical and enzymatic lysis of the bacterial cells followed by a chloroform-isoamyl

alcohol based DNA extraction, necessitating a 2-3 day protocol harsh enough to lyse the bacteria yet sufficiently gentle to prevent DNA shearing (Somerville *et al*.,2005).
The treatment of Lysosyme, proteinase K and sodium dodecyl sulphate (SDS) lyses cell wall and other protein boundaries to nucleic acids making the bacterial lysate. Then, Cetyltrimethylammonium bromide (CTAB) in high concentration of salt removes cell wall debris, denatured protein and polysaccharides by making complex to itself. The isolated bacterial chromosomal DNA is then recovered by precipitation with isopropanol, and further purified with ethanol precipitation and resuspend in TE buffer pH8.0.

CTAB is a cationic detergent that has the useful property of precipitating nucleic acids and acidic polysaccharides from solutions of low ionic strength (<0.5M). Under these conditions, proteins and neutral polysaccharides remains in solution whereas in solutions of high ionic strength (>0.7M), CTAB forms complexes with proteins and most acidic polysaccharides, but will not precipitate nucleic acids (Sambrook and Russell, 2001).

Digestion by PvuII

Restriction endonucleases recognize short DNA sequences and cleave double stranded DNA at specific sites within or adjacent to the recognition sequences. This cleavage results in the formation of discrete DNA fragments (van Soolingen *et al.,* 1995). The physical map of the IS*6110* sequence indicates that various restriction enzymes cleave within the element. BamHI, SstI, BstEII, BssHIII*,* and PvuII have been successfully used to generate RFLP. van Embden *et al.,* 1993 recommend the use of PvuII for the standard protocol, since then it has been used worldwide. PvuII cleaves the IS*6110* sequence only once, so that PvuII digestion of IS*6110*-containing genomic DNA leads to IS*6110*-hybridizing fragment of at least 0.9 or 0.46 kb depending on the probe that is used. PvuII cleaves at the following sequence;

CAG | CTG
GTC | GAC

Gel-electrophoresis

Agarose, which is an extract of seaweed is a linear polymer consisting of D-galactose and 3, 6-anhydro-L-galactose. Agarose gel is made by melting then pouring into a mold and allowing it to solidify at room temp. Upon solidification, the agarose forms a matrix, the density of which depends on the concentration of the agarose. When an electric charge is applied across the gel, DNA, which is negatively charged at neutral pH migrates towards the anode. The rate of migration depends upon a number of parameters, e.g. molecular size of DNA, agarose concentration, voltage applied etc.

The PvuII digested MTB DNA contains DNA fragments ranging from 0.5 to 25 kb in size. So, DNA electrophoresis in large gel (20cm × 20cm) size at low voltage (0.8V/cm) for long time (more than 18 hours) is recommended for good resolution of DNA fragments, which is necessary for Southern blotting. The DNA concentration should be equal in each lane because the higher the concentration of DNA, greater the mobility rate.

Southern blotting

Southern transfer and hybridization (Southern, 1975) are used to study how genes are organized within a genome by mapping restriction sites in and around segments of genomic DNA for which specific probes are available. The process involves the transfer and subsequent detection of electrophoretically separated DNA on a membrane. Analysis of the immobilized DNA is facilitated by hybridization with an appropriately labeled nucleic acid probe.

The transfer of electrophoretically separated DNA from gel to two-dimensional solid supports is a key step in southern hybridization. In upward capillary transfer, the liquid is drawn through the gel by capillary action which is established and maintained by a stack of dry absorbent paper towels. The rate of transfer of DNA depends upon the size of DNA fragments and the concentration of agarose in the gel. Capillary transfer of DNA >15kb in length requires at least 18 hours, and even then the transfer is not

complete. The efficiency of transfer of large DNA fragments is determined by the fraction of molecules that escape from the gel before it becomes dehydrated. As elution proceeds, fluids are drawn not only from the reservoir but also from the interstices of the gel itself. This flow reduces the gel to a rubbery substance through which DNA molecules cannot easily pass. The problem of dehydration due to lengthy transfer can be alleviated by partial acid/base hydrolysis of the DNA before capillary transfer. Acid treatment results in partial depurination, followed by strong base treatment, which hydrolyses the phosphodiester backbone at the site of depurination (Sambrook and Russell, 2001).

DNA Probe

Either DNA or RNA can be used as a probe. DNA probes have the advantages that DNA is inherently a more stable molecule than RNA and that essentially any DNA molecule can be directly labeled by either enzymatic or chemical means. DNA probes have the disadvantages however, that in most labeling protocols both strands of the duplex are labeled. This has two disadvantages. First, in a variety of situations one might wish to have strand specific probes. In addition, when denatured DNA probes are hybridized to an immobilized target the competing reaction of reassociation takes place, which means that during the course of the hybridization the effective probe concentration is continually falling. This probably makes DNA probes somewhat less sensitive than they might otherwise be.

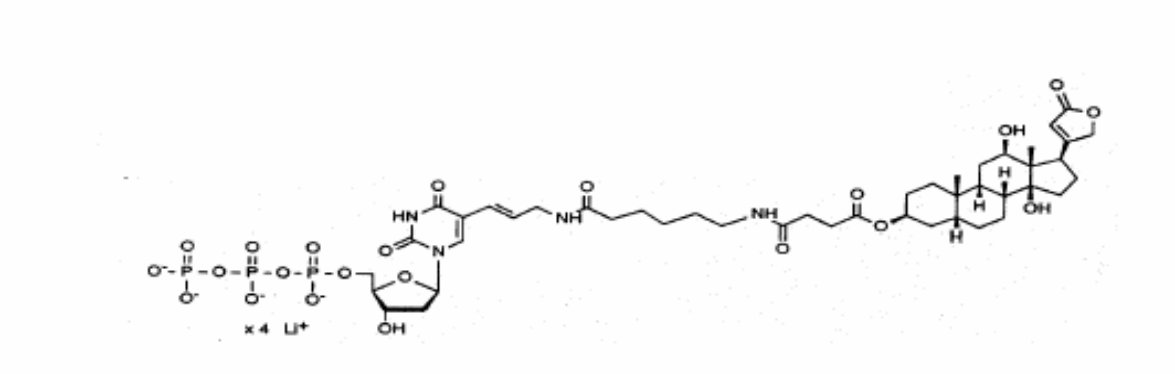
PvuII digestion of IS*6110* containing genomic DNA leads to a IS*6110*-hybridizing fragment of at least 0.90 or 0.46kb depending on the IS*6110* probe used. As high copy number (8-20) of IS*6110* is frequently found in most strains, the use of a DNA probe which overlaps both sides of the PvuII site would result in doubled number (16-40) of bands. Therefore, van Embden et al., (1993) chose a DNA probe to the right of the PvuII site on physical map, reducing the number of bands in the fingerprint. They further recommended to use 3'probe (hybridized to the left of the PvuII site) when the

differentiation of the pattern is not adequate with 5'probe. An oligonucleotide or an insert of a plasmid or a PCR fragment can be used as DNA-IS*6110*- probe.

PCR is a procedure of *in vitro* enzymatic amplification of a specific fragment of DNA. Primers are derived from the sequence to be amplified, which hybridize to opposite strands at either end of the sequence of intrest. Synthesis by DNA polymerase extends across the region of the original DNA template between the primers. In successive cycles of strand denaturations, primer annealing, and enzymatic extension, the amount of DNA from the previous cycle is doubled. The most frequently used polymerase is Taq polymerase.

DNA from any IS*6110* containing *M. tuberculosis* complex strain can be used as a target DNA for PCR. The most frequently used mycobacterial reference strains are *M. tuberculosis* Mt 14323, *M. bovis* BCG P3, *M. avium* R13, *M. paratubrculosis* ATCC19698.Based on the sequence of IS*6110*, primers ISN-1 and INS-2 were designed for preparation of a 245bp PCR fragment probe.

Figure 1 Structure of the Dig-□11□-dUTP molecule (Roche Diagnostics, Germany)



Non radioactive (DIG) labeling of probe

Non-radioactive labeling can be accomplished by direct or indirect labeling methods. The former includes direct incorporation of fluorescent tags, or cross-linking enzyme molecules directly to nucleic acid. Indirect labeling involves the incorporation of

nucleotides tagged with a hapten such as biotin or digoxigenin (DIG) during synthesis of the probe. Used in standard blotting and hybridization procedures, labeled probes that hybridize to a target sequence are detected with streptavidin (biotin) or anti-DIG monoclonal antibody (DIG) conjugated to an enzyme, usually phosphatase or peroxidase. Enzyme activity can be detected either by a chemiluminescent reaction whereby results are captured on X-ray film, or through formation of a colour precipitate deposited directly on the membrane. Nucleic acid probes generated by these means are stable for at least one year, in contrast to the short half-life of 32P-labeled probes (www kpl.com).

In the case of the Digoxigenin system, the hapten is digoxigenin, a steroid exclusively occurring in the plant *Digitalis purpurea*. The molecule is linked to desoxyuracilphosphate by an 11 atoms linear spacer (Dig-[11]-dUTP) (Roche Diagnostics, Germany).

For application in southern blotting, DNA probes can be DIG-labeled by any one of the following methods; random primed DNA labeling, incorporation of DIG-11-dUTP during PCR, 3'-tailing of oligonucleotides or 3'-and/or 5'-end labeling of oligonucleotides. Among them, the first two have been found to be of high sensitivity.

Digoxigenin-11-dUTP (DIG-dUTP) can be incorporated by *Taq* DNA polymerase during PCR. The resulting probes are very sensitive, and the yield from the labeling reaction is quite high. In addition, PCR probe synthesis is a cost-effective use of the DIG-dUTP nucleotide. The amount of DIG-dUTP in the PCR labeling reaction affects both the yield and sensitivity of the probe. When preparing a probe by PCR, it is important to consider the ratio of DIG-dUTP to dTTP in the labeling reaction. Published protocols have described DIG-dUTP: dTTP ratios of 1:2 through 1:20 but ratios of 1:2 through 1:5 are found to be ideal for producing maximally sensitive probes (e.g., for probing genomic Southern blots)(Roche Diagnostics, Germany).

Hybridization

Hybridization is a term used to describe the specific complementary base association due to hydrogen bonding, under experimental conditions, of single-stranded nucleic acids. The process immobilizes "target" nucleic acid previously denatured to single stranded so as to be effectively bound on the membrane, and then anneals to it an appropriately "tagged" or "labeled" single stranded nucleic acid. The mechanism of immobilization of nucleic acids on a membrane is not fully understood. Nylon membrane (Hybond-N^+) binds all nucleic acids under a wide range of salt concentration after baking the membrane to 80^0C in a vacuum for 2 hrs or UV irradiation for 5 min or less, or by treatment with 0.4M NaOH for irreversible or covalent attachment of nucleic acids to membrane (Vernon *et al.,* 1998).

DNA hybrid is dependent on certain factors, such as the melting temperature (T_m) at which the probe is annealed to 50% of its exact complement. The factors influencing the T_m are included in the formula of Meinkoth and Wahl (1984):

$T_m = 81.5°C + 16.6 \log M + 0.41$ (% G + C) - 500/n - 0.61 (% formamide)

Where M is the concentration (mol l-1) of monovalent cations in the hybridization solution/washing solution, (% G + C) the proportion of guanine and cytosine in the probe, and n the length of the probe in base pairs.

The melting temperature T_m together with the selected hybridization and washing temperature T_a determine the conditions for annealing between probe and target DNA. This is called the stringency:

Stringency (%) = 100 - Mf (T_m - T_a)

where Mf is the"mismatch factor" (1 for probes longer than 150 bp).

Under hybridization/washing conditions with a stringency of 100%, all DNA: DNA hybrids with less than 100% homology are resolved. In general, the lower the salt concentration in the washing solution and the higher the hybridization or washing temperature, the higher the stringency (http://www.iaea.org/programmes/nafa/d2).

Chemiluminescent detection

The chemiluminescence substrate CSPD emits light after a two-step reaction. At first, the molecule is de-phosphorylated by the enzyme alkaline phosphatase (AP), and in the second step the molecule decomposes and emits light. The emitted light appears as a continuous glow for more than 24 hours, and it can be documented on X-ray films. The advantages of chemiluminescence include remarkably improved sensitivity, the possibility to test different exposure times, and the facilitation of rehybridization experiments (http://www.iaea.org/programmes/nafa/d2).

Chemiluminescent detection is a three-step process. The membrane is treated with a blocking reagent to prevent nonspecific attraction of antibody to the membrane. Then, it is incubated with a dilution of Anti-Digoxigenin, Fab fragment, which are conjugated to alkaline phosphatase. In the third step, the membrane carrying the hybridized probe and bound antibody antibody conjugate is reacted with a chemiluminescent substrate and exposed to X-ray film to record the chemiluminescent signal (Roche Diagnostics, Germany).

CHAPTER-IV

4. MATERIALS AND METHODS

4.1 Materials

A complete list of bacteriological media, reagents, chemicals, equipment, glassware and miscellaneous materials required for this study is given in appendices I.,II, III and IV.

4.2 Methodology I

4.2.1 Sample collection site

The samples were collected from two different sites at two different time points.

1. National Tuberculosis Centre (NTC)

The samples were collected from March to June 2005. Microscopically AFB positive sputum samples were collected cultured and heat killed at the National Tuberculosis Centre (NTC), Thimi, Bhaktapur. The whole process of Restriction Fragment Length Polymorphism (RFLP) was performed at the Mycobacterial Research Laboratory (MRL), Anandaban Hospital (AH, Lele).

2. German Nepal Tuberculosis Project (GENETUP)

In the second phase of the study, 76 culture positive samples from new cases with more than 5 colonies per culture tube on Ogawa medium were harvested, and heat killed at GENETUP in December 2006.

4.2.2 Patient enrollment

Pulmonary tuberculosis patients with sputum positive for acid fast bacilli (AFB) attending outpatient department (OPD) of NTC from March to June 2005 were selected for the study. These included 117 new cases; patients without case history of treatment

for TB or patients without previous administration of TB drugs before in life. HIV positive new cases were excluded.

4.2.3 Sample collection and processing

The standard procedure recommended by WHO for sample collection, microscopy and culture was followed. All the sputum samples were collected in a leak proof, wide mouth and sterile plastic container. An appropriate portion of the sample was transferred to the slide with the help of the broken end of a wooden stick, smeared over an area of approximately 2.0 by 1.0 cm. The smear had to be thin enough to be able to read through it. Then it was allowed to air dry for 15 minutes, heat fixed and cool to Room Temperature (RT). The fluorochrome staining method was performed. The smeared slides were placed on a staining rack and flooded with Auramine-O, incubated for 15 minutes, rinsed with distilled water followed by 20% sulphuric acid treatment for 5 minutes. The smear was rinsed again with distilled water and counterstained with 1% methylene blue for 2 minutes (Appendix I). The washed and air-dried smear was then examined directly under the UV microscope. The microscopy result was interpreted on the basis of WHO recommended procedure.

On the basis of positive microscopy results, the patients were selected and directly interviewed. The samples from only new cases were selected for culture, excluding samples from relapses, re-infection and HIV co-infection patients.

4.2.4 Culture

Sputum samples after microscopy were mixed with twice of its volume of 4% NaOH in a graduated centrifuge tube of 15 ml capacity and shaken several times to digest, then left to stand for 15 minutes at room temperature with occasional shaking. The specimen was centrifuged at 3000g for 15 minutes and the sediment was suspended with 15 ml distilled water. The tube was again centrifuged at 3000g for 15 minutes. The sediment was harvested after discarding the supernatant and was used for culture. 400µl of concentrated sputum was inoculated into each of two culture tubes containing 3%

modified Ogawa medium (APPENDIX I). The tubes were examined on the 7th day for rapid growers and weekly thereafter for slow growers. If any colonies were seen at any stage, acid-fastness of bacilli was determined by smear examination of the growth. High-grade culture tubes with isolated colonies after growth for 8 weeks were selected for RFLP. Grading of primary culture is given in APPENDIX VII

4.2.5 Transfer and heat killing of cells

After 8 weeks of incubation period, when growth became clearly visible and mature, two loopfuls of cells or 4-5 colonies were transferred in to labeled micro-centrifuge tubes containing 400µl of 1xTE buffer (pH 8.0). The eppendorf tubes containing bacterial cells were incubated at 80ºC for 20 minutes to kill the cells, and then cooled to room temperature. Each of the tubes was then sealed with parafilm and placed in a safety carrier box for transportation at RT to Mycobacterial Research Laboratory.

4.3 Methodology II (MRL)

4.3.1 DNA extraction

The bacterial suspension was made homogenized by vortexing. Fifty microlitres of (10 mg/ml) lysozyme was added to the cells, vortexed and incubated overnight at 37ºC water bath. Next day, 75 µl of 10% SDS/proteinase K (10mg/ml) mix (70µl +5 µl) was added to the lysate, vortexed briefly then incubated at 65ºC for 10 minutes. 100 µl of 5M NaCl, followed by 100 µl of prewarmed CTAB/NaCl (APPENDIX IV) solution was added, vortexed until the liquid content became milky white and incubated at 65ºC for 10 minutes. Then, 750 µl of chloroform/isoamyl alcohol (24:1) was added, mixed (inverted 3-5 times) manually, and centrifuged for 5 minutes at 13000 rpm. All of the aqueous supernatant was transferred in 180 µl amounts to a fresh centrifuge tube using a pipetter (P-200). Four hundred and fifty µl of cold isopropanol (-20 ºC) was added, shaked manually to see the size of precipitated nucleic acids, according to which the final resuspending volume of 1xTE was estimated (ANNEX IV). The content was chilled at -20ºC for 30 min and centrifuged for 15 minutes at 13000 rpm. The

supernatant was discarded leaving 20 µl above the pellet using P-200, 1 ml of cold 70% ethanol was added, mixed manually and centrifuged at 13000 rpm for 5 minutes. The supernatant was discarded as in previous step. The last 20µl supernatant was again centrifuged for 1 min and removed completely using P- 20. The tube was left open for 30 minutes at RT to dry out the pellet. Then finally, the pellet was redissolved in the estimated amount of 1xTE and stored at 4ºC until needed.

4.3.2 DNA quantification

The estimation of the amount of extracted DNA from the samples was carried out using the spectrophotometric method. In this method, the 2 µl sample was taken in eppendorf tube and diluted to 200 (1:100) fold by distilled water. Absorbance was measured in the spectrophotometer at 260 nm and 280 nm by placing the sample in capillary cuvette, against distilled water as blank. OD_{260}/OD_{280} ratio provides an estimate of purity of the nucleic acid, which is derived as follows (Sambrook *etal.,* 1989):

OD_1 =Absorbance at 260nm

OD_2 = Absorbance at 280nm

a = OD_1 x 50 *x 100

b = OD_2 x 50 * x100

Where,

50 = Constant factor for ds DNA (*1OD = 50 µg DNA)

100 = Dilution factor

Pure preparation of DNA had a/b value of 1.8 (Sambrook *et al.,* 1989).

4.3.3 Digestion of chromosomal DNA by PvuII

Before digesting bacterial DNA, the components of digestion were prepared in an eppendorf tube in a safety cabinet, previously cleaned with 20% hypochloride solution and sterilized for 10 minutes by UV light. Assuming X, as the required volume of DNA to make 4.5 µg in final reaction mixture of 20 µl, 15-X µl of was added in each 200 µl tube. Then, 5 µl of the digestion mix was aliquoted in each tube, followed by X µl of

target DNA in respective tubes. Genomic DNA of TBM 14323 was also digested with each set of samples.

Table 2 Components of digestion mixture

	Digestion mix components	Stock solution	Required Vol. from stock (µl)	Final concentration
1	Sigma water	-	2	-
2	Digestion Buffer	10x	2	1x
3	PvuII	10U/µl	1	10U/rxn

The content in each tube was mixed thoroughly and touch centrifuged. The tubes then were incubated overnight at 37ºC in a water bath. The DNA loading was carried out on a separate bench. This method of setting reaction minimizes the possibility of pipetting errors and also saves time by reducing the number of reagent transfers.

4.3.4 Separation of digested DNA fragments

4.3.4.1 Gel preparation

Agarose gel (0.8 %) was prepared in 1×Tris-borate-EDTA buffer (TBE) The agarose was solubilized by heating in a microwave It was then cooled to luke warm, poured on the gel casting tray fitted with comb on the horizontal electrophoresis unit and left at RT for 30 minutes to set. Then the electrophoresis unit was filled with 1x TBE so that the buffer just covered the surface of the gel.

4.3.4.2 Electrophoresis

The PvuII digested DNA samples and external marker *MTB14323* were taken out from water bath (37^0C). Each tube was flipped properly to mix the digested fragments in the solution. The tubes were then touch centrifuged and arranged in order to load on the gel. 5 µl of lamda/phi marker with five microlitres BPB loading buffer (APPENDIX IV) was loaded in the first well. The entire digested DNA (20 µl) was mixed with 5 µl of BPB on

a piece of parafilm and loaded in respective wells. The second lane, middle lane, and lane before last were loaded with MTB14323. The last lane was left empty. Firstly, electrophoresis was carried out at 100V for 1 hour. Then the voltage was decreased to 24 V and run for 22 hrs. The gel was checked on UV transilluminator to measure the distance reached by 2kb band of the lamda/phi marker. Electrophoresis was continued till the distance traveled by the band was 7±0.4cm from the slots to obtain an optimal separation of DNA fragments.

4.3.4.3 Staining and visualization

For staining, 0.56 µg/ml of ethidium bromide in 400ml distilled water was used. Gel was immersed in the solution and incubated at RT for an hour. The gel was then washed twice, with 400ml of distilled water each time. For visualization, UV transilluminator was used. In RFLP, visualization was done for two purposes, to optimize the electrophoresis duration and to analyze digested fragments. The lack of DNA in wells and presence of smooth smearing with similar intensity along the lanes indicate the success of digestion and separation of DNA fragments. The top and bottom space of gel without DNA was cut off and the exact size of the gel remaining was measured.

4.3.5 Southern blotting

4.3.5.1 Treatment of gel before blotting

A) Denaturation

For denaturation of DNA, the gel was incubated in 500ml of denaturation solution (APPENDIX - IV) for 1 hour at RT.

B) Neutralization

The gel was incubated in 500 ml of neutralization solution (APPENDEX- III) for 1 hour at RT.

4.3.5.2 Paper ware set up

One Hybond N^+ nylon membrane, 6 pieces of Whatman 3mm filter paper, 12 pieces of thick blotting paper, one and half roll of tissue stack were cut to the same size of the gel. Similarly one piece of Whatman paper with the same width and 38cm length for wick and 4 pieces of parafilm with 4cm × 8cm were cut out.

4.3.5.3 Blotting

Neat and clean blotting tank was taken. Gel holder was placed inverted on the tank. Whatman paper wick was soaked with 20× SSC and placed on the gel holder slowly not forming any air bubbles. Whatman paper 1 soaked with 20× SSC was placed on the wick. The neutralized gel was taken and placed on the Whatman paper1 in face down position. Hybond N+ membrane soaked in 20× SSC was placed on the gel. All four edges were covered with parafilm to prevent short circuit with the buffer to the tissue stack. On the membrane, Whatman paper 2 soaked with 20× SSC followed by paper 3 and 4 soaked with 5× SSC were placed accordingly. Air bubbles were removed in each step by rolling a serological pipette over the membrane and filters. One piece of dry Whatman paper and a tissue stack was placed, above which a glass slab and weight of 1 kg was placed to enhance the capillary blotting. The set up was left untouched overnight (18 hours).

4.3.5.4 Post blotting membrane treatment

After overnight blotting the setup was dismantled, the Hybond membrane was unfolded and briefly washed with 5× SSC for 3 minutes, air dried completely, placed in between two sheets of Whatman paper and baked for 2 hours at 80^oC for covalent linkage of DNA to the membrane.

4.3.6 Preparation and labeling of DNA probe by PCR

4.3.6.1 Optimization of IS*6110*-PCR

The probe was prepared by using a pair of DIG-labelled primers ISN-1 and ISN-2 as recommended by van Soolingen *et al.,* (1995). The probe was simultaneously labeled during PCR by using DIG labeled nucleotide DIG-11-dUTP in ratio of 1:3 with dTTP during PCR.

Sequence of Primers

INS-1: 5'- CGTGAGGGCATCGAGGTGGC
Anneals to bp 631 through 650

INS-2: 5'- GCGTAGGCGTCGGTGACAAA
Anneals to bp 856 through 875

Before commencing PCR, the working area in safety cabinet was cleaned with 20% Sodium hypochlorite solution and sterilized with UV light for 10 minutes. By practicing aseptical techniques, the components of PCR like dNTPS, 100mM (except DIG-dUTP) and primers (100pmol/µl) were diluted to their working concentration 1mM and 10pmol/µl respectively. PCR mix was prepared in 1.5ml eppendorf tube and further aliquoted as required, 40µl to each of 200µl PCR tubes inside safety cabinet. Lastly, 10µl of BCG DNA was added in each tube, outside of UV cabinet.

Table 3 Preparation of working nucleotides solution

Nucleotide	Stock Conc. (mM)	Bench Conc. (mM)	Vol. of stock taken (μl)	Final Vol. (μl)	Working Conc.(mM)
dATP	100	10	10	100	1
dCTP	100	10	10		1
dGTP	100	10	10		1
dTTP	100	10	10	100	1
DIG-dUTP	1	1	1	1	1

As the recommended procedure for IS*6110* –PCR was only for non labeled PCR product; optimization of PCR was carried out by bringing variation among components of PCR mix and cycle parameters. This was brought about by examing the effected variations in concentration of $MgCl_2$, *Taq* polymerase and ratio of dUTP:dTTP. The optimum result was adopted while running PCR for preparation of dig-labeled IS*6110* probe.

Different concentration ratio 1:2, 1:3 and 1:4 of non labeled dUTP:dTTP were tested . The 1:2 ratios of nucleotides was found to give the brightest band and thus employed in PCR with DIG labeled dUTPs. Magnesium chloride was tested at 0, 1.0, 1.5, 2.0, 2.5, and 3.0mM concentrations and the optimized condition was found to be 2.5mM so that concentration was adopted for final PCR. Taq polymerase of 0.9unit (i.e.0.3 μl of 3 units/ μl) per reaction was used as it gave the best result in optimization. Similarly, BCG DNA was diluted to 75ng/ml, 25ng/ml, and 5ng/ml and used in optimization. 5ng/ml DNA gave the best band and was used in probe production.

The parameters of thermocycler were also optimized. 95^{o}C denaturation and 62^{o}C annealing temperature for 40 seconds time periods were found to be optimum so used in final protocol.

Table 4. Optimization of reaction mixture for IS*6110*-PCR using DIG-11-dUTPs

Constituent	**Working Conc. (mM)**	**Volume taken (µl)**	Final Conc./ (50µl) rxn
Sigma Water	-	7.7	-
PCR Buffer	10×	5	1×
dNTPs (dATP,dCTP,dGTP)	1mM	10	200 µM
dTTP	1mM	7.5	150 µM
Dig-11-dUTP/dUTP	1mM	2.5	50 µM
INS-1	10pmol/ µl	1	10 pMol
INS-2	10pmol/ µl	1	10pMol
Magnesium Chloride	25mM	5	2.5 µM
Taq Polymerase	3U/ µl	0.3	1U
BCG DNA	1000×	10	

Table 5. Optimization of thermocycler conditions for IS*6110*-PCR using DIG-11-dUTPs

Parameter	**Condition**		**Number of cycles**
	Temperature (^{o} C)	**Time (min/Sec.)**	
Initial Denaturation	95	5	**30**
Denaturation	95	40 Sec.	
Annealing	62	40 Sec.	
Extention	72	1min	
Final Extention	72	10min	

4.3.6.2 Visualization of amplified DIG-labeled IS*6110* probe

The amplified dig-labeled probe was visualized by gel electrophoresis. Ten microlitres of PCR product probe was mixed with 5 µl of loading buffer. The mixture was then

loaded on a 1% agarose gel. A molecular weight marker was used for determination of required size of band. Electophoresis was carried out at RT for 1 hour at 145V. After completion of electrophoresis, the gel was stained with 0.56 µg/ml solution of ethidium bromide 400ml of distilled water for 1 hour. The PCR product was visualized on UV transilluminator and analyzed for the expected band size by comparing the target band with that of molecular weight marker. The expected band size for nonlabeled IS*6110* control probe is 245bp and slight greater around 250bp for that of dig-labeled one. Finally, the gel was photographed for documentation.

4.3.6.3 Quantification of PCR probe

The concentration of PCR product DNA probe was measured by spectrophotometer using 100 times diluted PCR probe as described in 4.3.2.

4.3.7 Southern hybridization

4.3.7.1 Prehybridization

The baked membrane was placed in the hybridization roller tube and briefly washed with 2× SSC. Prewarmed (42^{o}C) hybridization solution was added to the membrane and incubated for 1 hour at 42^{o}C.

4.3.7.2 Preparation of working probe

The labeling concentration of probe depends upon the size of membrane and the method of labeling the probe. Van Soolingen *et al.,* (1995) recommended 10ng/µl as minimum labeling concentration when probe DNA was labeled with horseradish peroxide (ECL-kit) and used in 0.625ng/cm^2 amount. When probe was labeled with random primed labeling kit, 25ng/ ml and in PCR labeling 2 µl PCR product /ml of hybridization solution was recommended (Roche Diagnostics, Germany). But, 1 µl/ml was found to be good enough and used in final protocol.

Ten microlitres of PCR amplified probe was diluted in 10ml of prehybridization solution. The probe was denatured by boiling for 10 minutes followed instantly by chilling on ice for 10 minutes. The working probe was warmed to 42ºC until used in hybridization.

4.3.7.3 Hybridization

The Prehybridization solution was discarded completely and hybridization solution with probe was added .The membrane was hybridized overnight at 42ºC in hybridizer.

4.3.8 Detection

4.3.8.1 Membrane washing

The membrane was washed with 25ml of 2 ×wash buffers (APPENDIX-IV) for 5 min at 25^{o}C twice. The 2×wash solutions were discarded and stringency wash solution (0.1× wash previously heated to 68^{o}C) was added and washed at 68^{o}C twice for 15 minutes in hybridizer. Finally the wash solution was completely discarded.

4.3.8.2 Membrane DIG detection

The membrane was rinsed with 20ml of detection buffer I (APPENDIX IV) for 20 minutes at RT. It was blocked for 30 minutes in 10 ml detection buffer II (APPENDIX IV) at RT. Then 2µl of anti-DIG AP (2/10,000 dilution) was added and incubated at RT for 1 hour. The excess antibody was washed with 20 ml washing buffer (APENDIX IV) for 30 minutes. The membrane was equilibrated with 10 ml detection buffer III (APPENDIX IV) for 5 minutes, drained well and incubated with 10 ml CSPD solution (1:100 dilution in Buffer III) for 5 minutes.

4.3.8.3 Autoradiography of membrane

The membrane was wrapped in cling film (saran wrap) and incubated at 37^{o}C for 15 minutes to activate the CSPD. It was then attached on white board of Kodak catridge. In

dark room, an X-ray film was placed above the membrane with marker cut at one side, locked the catridge box and exposed for 2 hours to 48 hours. For development, the film was immersed in developer solution (APPENDIX III) for 1 to 5 minutes, washed in water for 10 seconds, immersed in fixer for 2 minutes and washed in water for 10 seconds.

4.3.8.4 Stripping and storage of membrane

The membrane was stripped by alkali probe stripping method. At first the membrane was rinsed in distilled water for 1 minute and then incubated in alkali probe stripping solution twice (APPENDIX IV) for 10 min. at 37^{o}C .Then it was rinsed thoroughly in 2× SSC and packed in saran wrap.

4.3.9 Result analysis

The results of hybridization were inspected visually. The sizes of the bands of the isolates with similar patterns were calculated based on the assumption that the migration distance has linear relationship with the logarithm of the size of the DNA bands. The similarity coefficient for visually suspected cluster cases was calculated by the formula.

$$\underline{S=2m \times 100/(a+b)}$$

Where, S is the similarity coefficient, m is the number of matched bands and a and b are the total number of bands in the two isolates being compared. The numbers of matched bands were confirmed by applying grid lines over the scanned autoradiograph (Palittapongarnpim *et al.,* 1997).

Clustering was defined as the occurrence of two or more isolates from different patients that contained more than one copy of IS*6110* and had similarity coefficient of more than 78%. For single copy isolates molecular weight was estimated by comparing the band with that of external marker *M. tuberculosis* 14323.

Fig-2 Flow chart of methodology I (NTC)

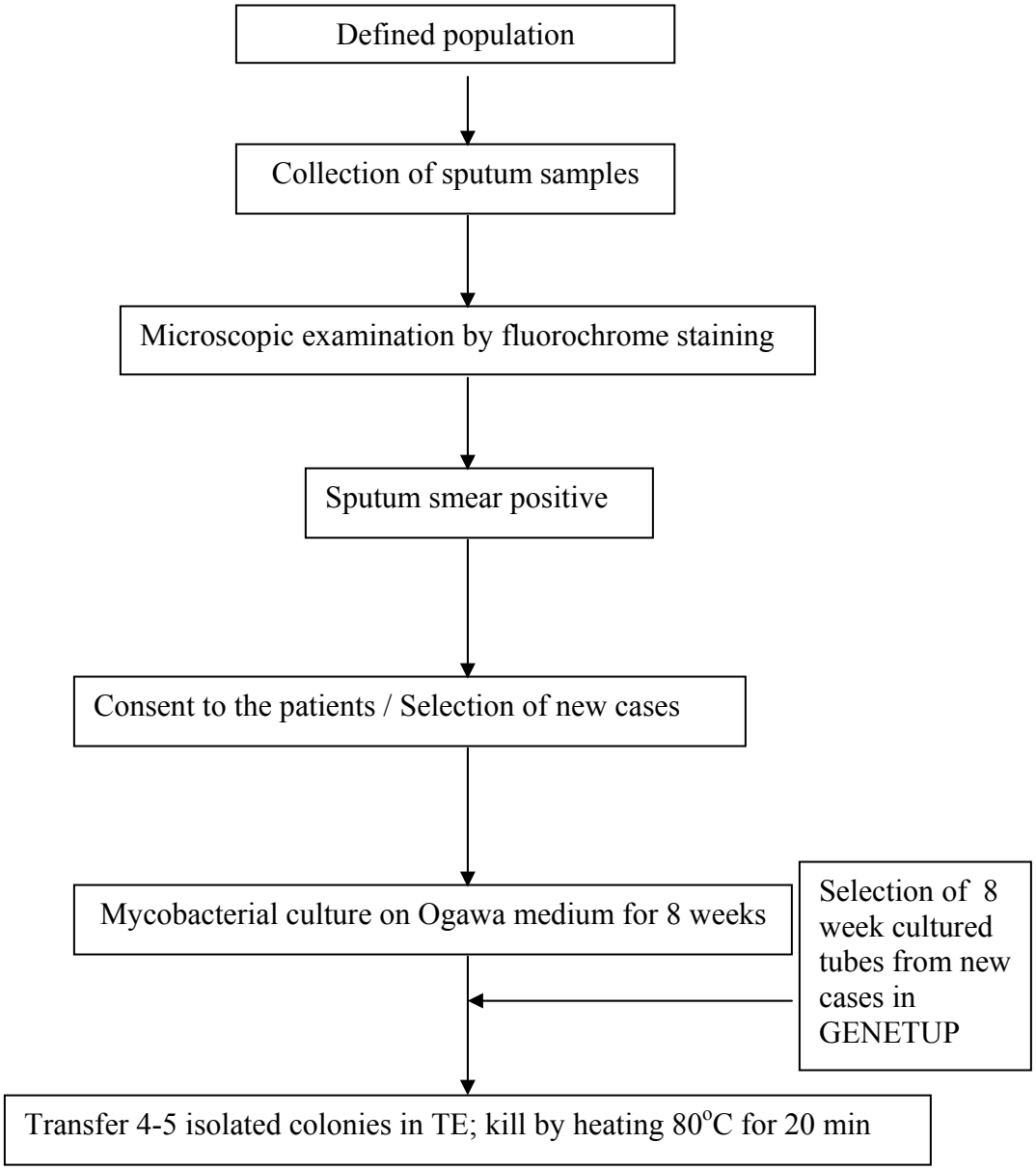
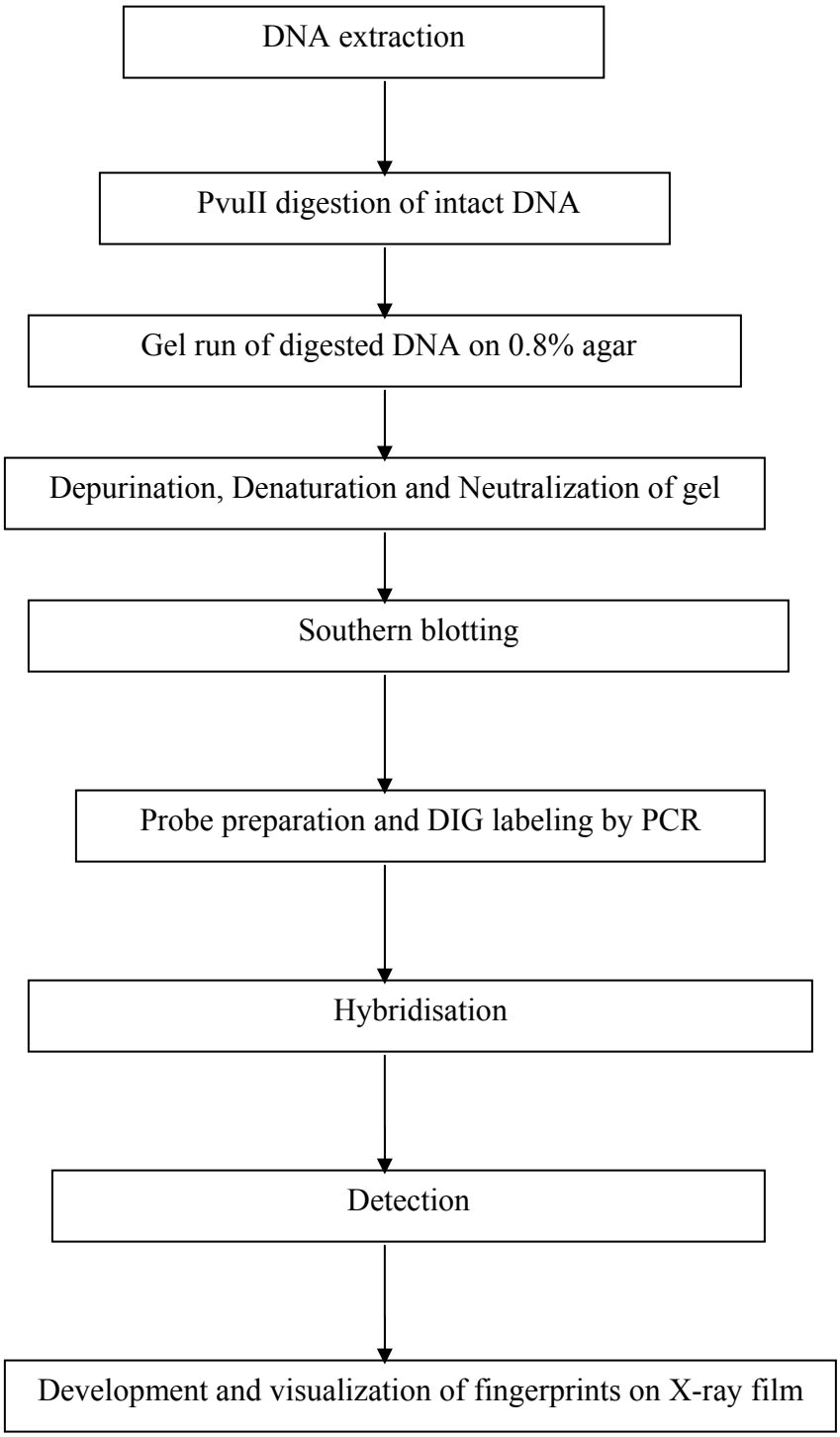


Fig 3 Flow Chart of Methodology (MRL)



CHAPTER-V

5 RESULTS

5.1 Sample analysis

In this study, a total of 280 culture positive samples were collected from two different tuberculosis centers. On that basis the study was categorized in two parts.

Table 6. Distribution of samples in different parts of the study

Specimen collection centre	**Study Design**	**No. of cases selected for study**	**No. of cultures used in DNA extraction**	**No. of isolates subjected to RFLP**	**Isolates with Interpretable results in RFLP**
NTC	First Part	117	204*	53 (45%)	*25 (47%)*
GENETUP	Second part	76	76	63 (83%)	*34 (53%)*
Total		*193*	*280*	*116 (60%)*	*59 (51%)*

*** Two or more than two samples were collected from each case**

5.1.1 Sample analysis for part I study

Two hundred and four sputum samples were collected from 117 smear positive PTB patients without the previous history of anti-TB treatment. The samples were cultured on Ogawa medium, 8 week old colonies were harvested in TE buffer, heat killed and DNA was extracted. Of 204 DNA extracted, 54 (26%) gave insufficient amount of DNA, 12 (6%) were of low quality and 33 (16%) DNA samples were found to be sheared and thus rejected. Of the remaining 105 DNA samples from 96 patients, 52 were used in finalizing standard operating procedure and 53 (45%) isolates from total

117 selected cases, were subjected to the final study. Of them, 25 (47%) of isolates were analyzed finally for their IS*6110* RFLP patterns.

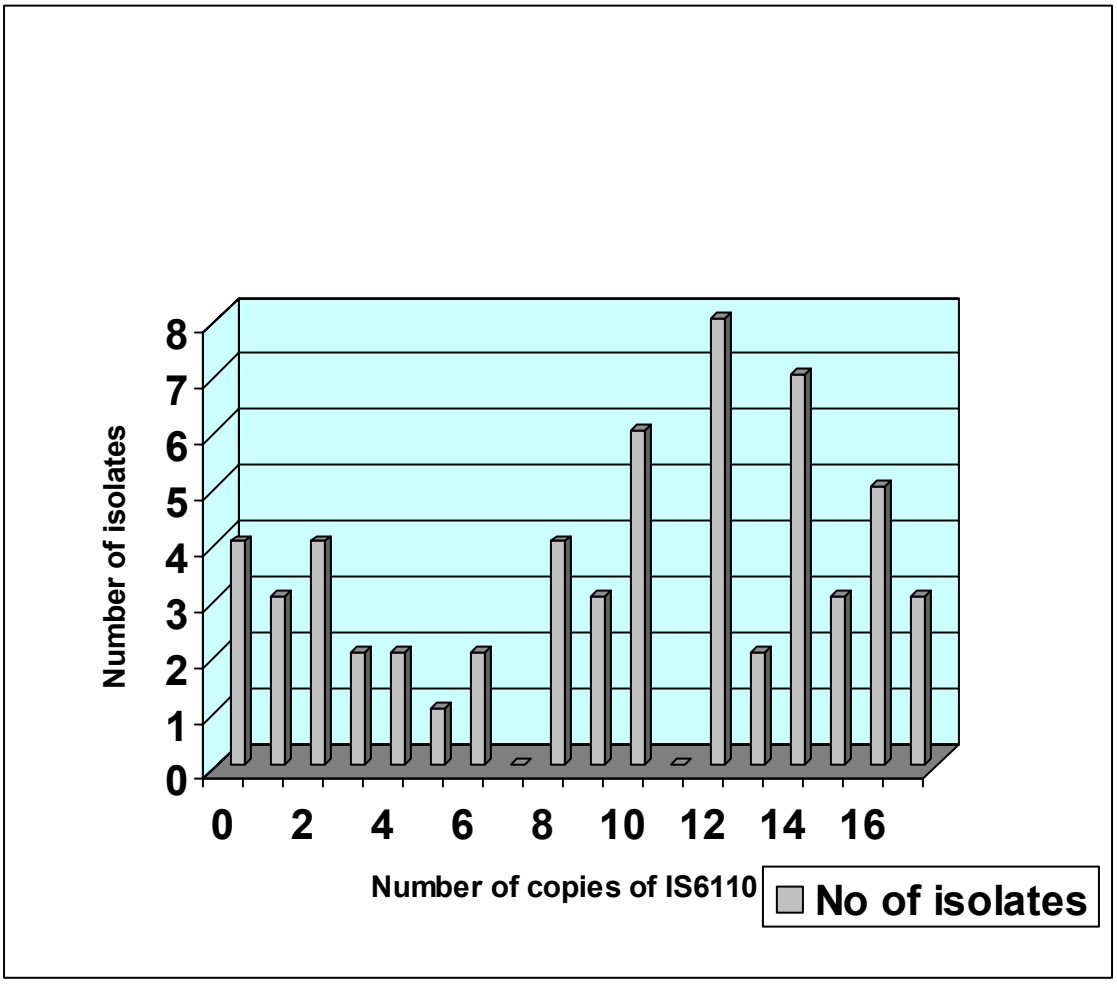
5.1.2 Sample analysis for part II study

Seventy six culture tubes bearing at least 5 colonies from 76 new PTB patients were selected for the study; their colonies harvested in TE buffer, the cells killed and DNA extracted. Of the 76 samples, 13 (17%) gave insufficient amount of DNA and 63 (83%) were subjected to RFLP, of which 34 (54%) isolates were analyzed for their fingerprinting patterns.

5.2 Result analysis

Out of 116 isolates subjected to RFLP, fingerprints from 59 isolates were found to be interpretable and used for final result analysis. At first, isolates were grouped on the basis of number of copies of IS*6110* and analyzed with respect to gender, age geographical regions and races. The fingerprints with similar copies of IS*6110* and banding pattern were further analyzed for clustered cases and separated into five different clusters.

Figure 4 Genetic variation of *M. tuberculosis* on the basis of number of copies of IS*6110*



In this study, DNA of 7% (4 isolates) did not hybridize to the probe suggesting them to be isolates with no IS*6110* copies. The number of IS*6110* copies per isolate studied varied from 1 to 17; however isolates with 7 and 11 copies were not found. Five percent (3 isolates) hybridized at only one position. Of these, the hybridized fragment of two isolates was of same length (1.4kb) and third with different length around 0.8kb. Similarly banding pattern in two of 4 double banded isolates was identical. The least and most frequent copy number of IS*6110* were found to be 5 (1 isolate; 2%) and 12 (8 isolates; 14 %).

Figure 5 Gender wise distributions of IS*6110* copies

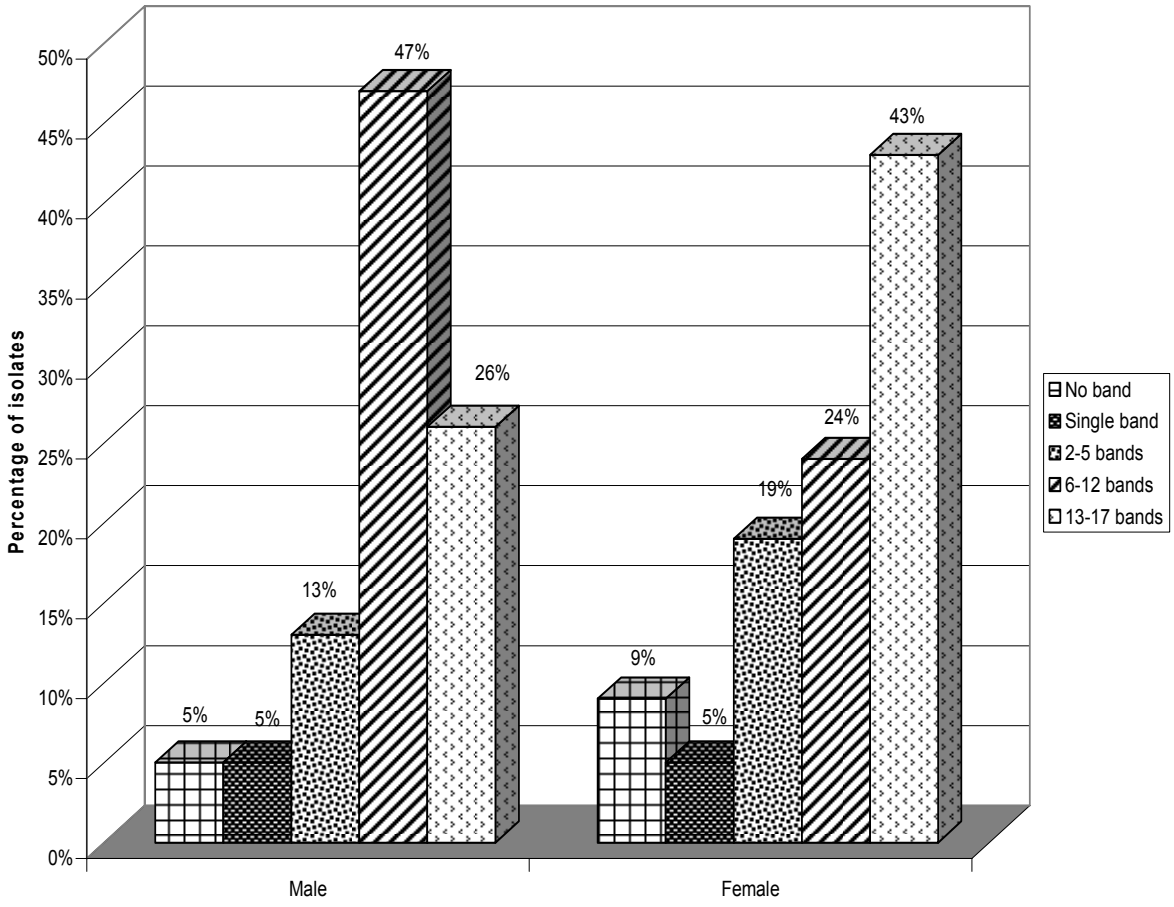


Figure 6 Distribution of IS*6110* copy numbers in various age groups

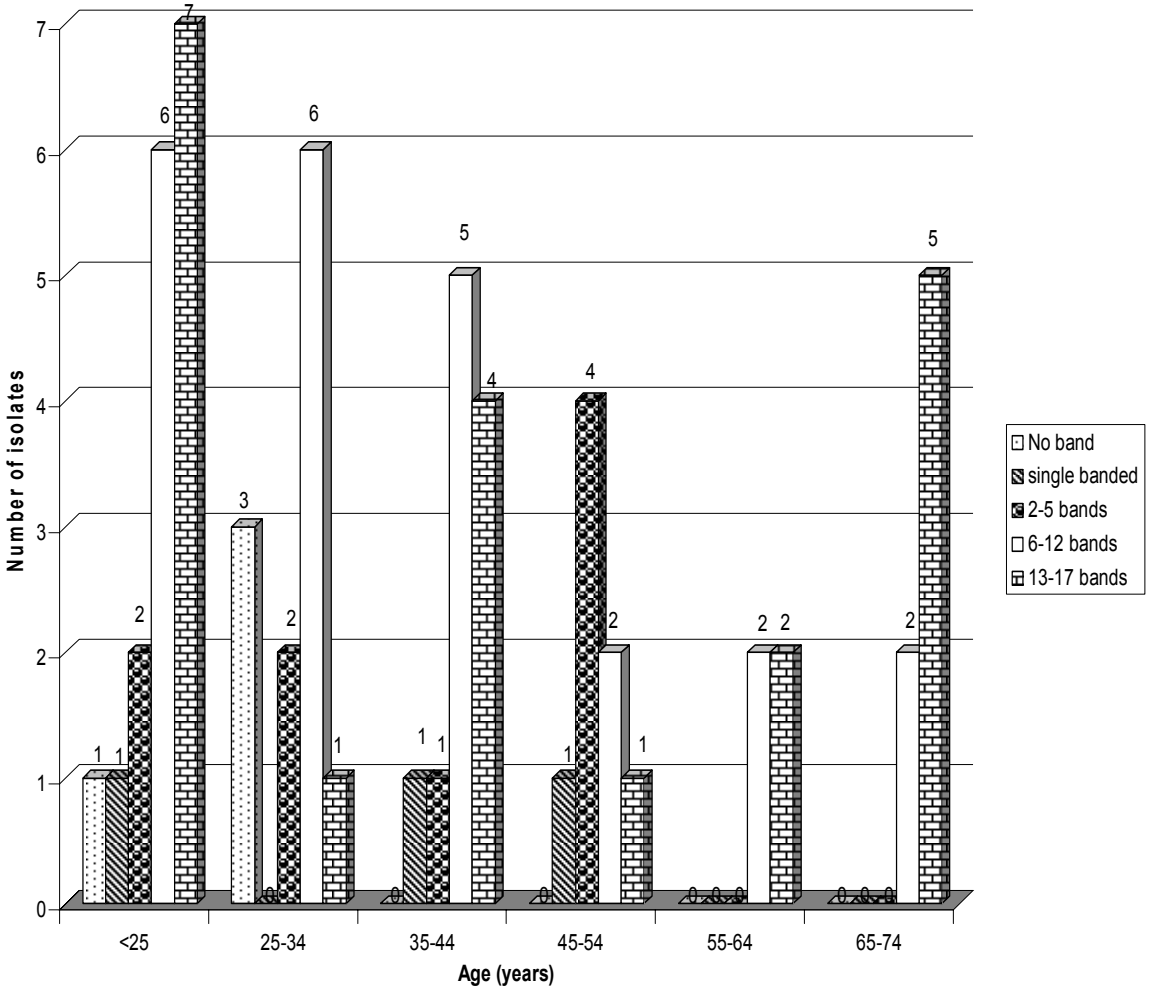


Figure 7 Distribution of different number of copies of IS*6110*-RFLP in various geographical regions

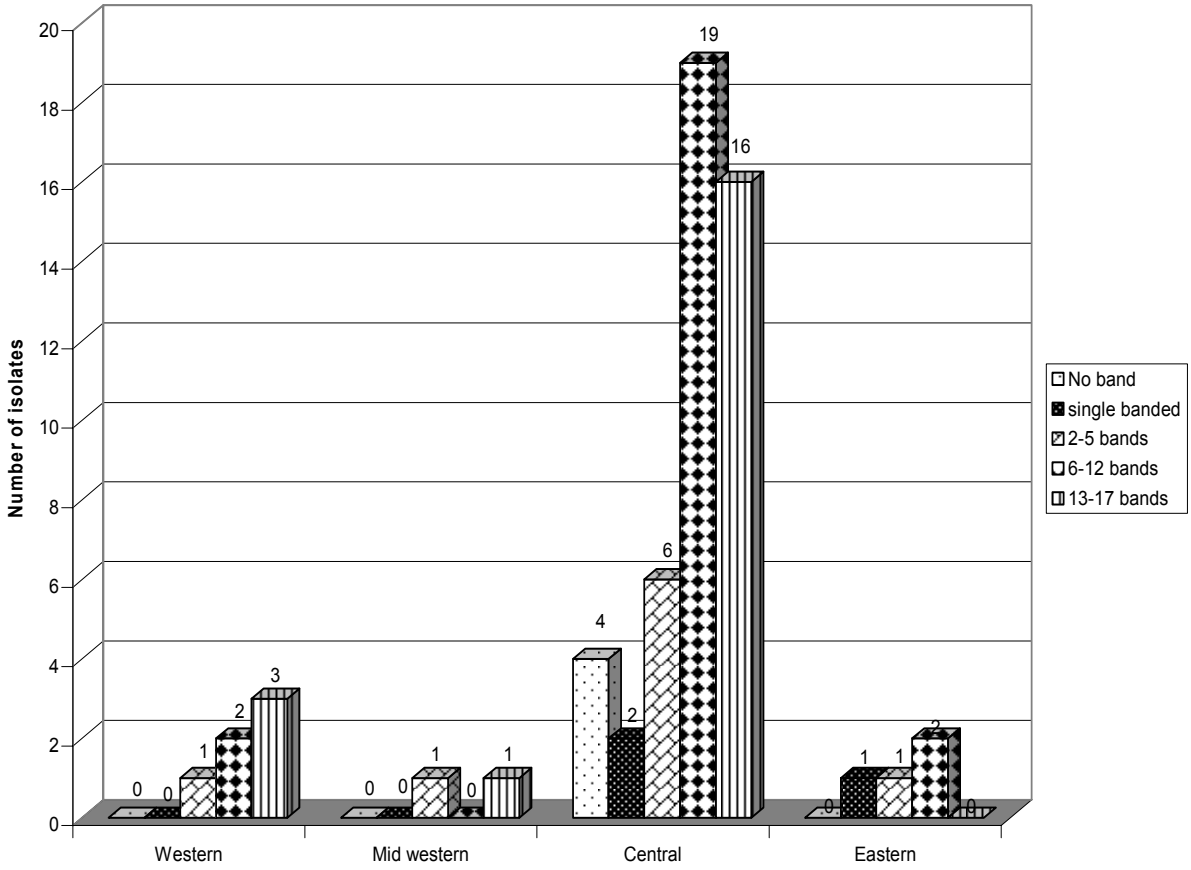


Table 7. Gender wise distribution of IS*6110* RFLP of *M. tuberculosis*

Gender	No band	Single banded	2-5 bands	6-12 bands	13-17 bands	Total
Male	2 (5%)	2 (5%)	5 (13%)	18 (47%)	11 (29%)	**38 (64%)**
Female	2 (9%)	1 (5%)	4 (19%)	5 (24%)	9 (43%)	**21 (36%)**
Total	**4(6.78%)**	**3(5.08%)**	**9(15.25%)**	**23(40.67%)**	**20(32.2%)**	**59**

The ratio of male: female in total isolates was found to be 1.8:1. Eighteen (47%) of isolates from males contained 13-17 copies where as 9 (43%) of isolates from females contained 6-12 copies of IS*6110*.

Table 8. Distribution of different IS*6110*-RFLP pattern among various age groups

Age, years	No band	Single band	2-5 bands	6-12 bands	13-17 bands	Total
17-24	1 (6%)	1 (6%)	2 (12%)	6 (35%)	7 (41%)	**17 (29%)**
25-34	3 (25%)	0	2 (17%)	6 (50%)	1 (8%)	**12 (20%)**
35-44	0	1 (9%)	1 (9%)	5 (45%)	4 (36%)	**11 (19%)**
45-54	0	1 (12%)	4 (50%)	2 (25%)	1 (12%)	**8 (13%)**
55-64	0	0	0	2 (50%)	2 (50%)	**4 (7%)**
>64	0	0	0	2 (29%)	5 (71%)	**7 (12%)**
Total	**4**	**3**	**9**	**23**	**20**	**59**

Of the total isolates, the highest number of patients 17 (29%) were from age group below 24 years, and the lowest 4 (7%) from the age group 55-64. All types of banding patterns were observed in very young age group below 25 years, however 13-17 IS*6110* copy numbered strains were responsible for 41% of disease cases in this age groups. No banded isolates were only seen in age below 35 years. Similarly, single banded and 2-5 bands isolates were observed in the patients below 55 years. Fifty percent (4/8) of isolates from age group 45-54 contained 2-5 copies of IS*6110*. Old aged patients were

seem to be infected with high copy numbered (>6) strains. Seventy one percent of disease cases above the age 65 years were due to 13-17 IS*6110* copy numbered strains. However, no significant association was observed in terms of IS*6110* copy number of isolates in various age groups of patients ($\chi 2$, $P<0.05$).

Table 9. Distribution of IS*6110*-RFLP patterns in various regions of the country

	No band	**Single banded**	**2-5 bands**	**6-12bands**	**12-17bands**	**Total N (%)**
Western	0	0	1 (17%)	2 (33%)	3 (50%)	**6 (10%)**
Mid west	0	0	1 (50)	0	1 (50)	**2 (3%)**
Central	4 (8%)	2 (4%)	6 (13%)	19 (40%)	16 (34%)	**47 (80%)**
Eastern	0	1 (25%)	1 (25%)	2 (50%)	0	**4 (7%)**
Total	**4**	**3**	**9**	**23**	**20**	**59**

In this study, the variation in IS*6110*-RFLP banding patterns with geographical variation was not observed apparently as the most of the patient 47 (80%) were from central region. The least number of patients 2 (3%) were from mid western region. All types of banding pattern with 0-17 copy numbered strains were observed in the isolates from Central region however the most predominant strains (40%) had 6-12 copies of IS*6110* which is also the most frequent type of isolate observed in the present study but not found in isolates from Mid-Western region. Non hybridized isolates were observed only in central region. Excluding Central region, single banded isolate was only found in Eastern region of the country, where high (>13) copy numbered strains were not observed. Fifty percent (3/6) from Western region contained 13-17 copies of IS*6110*. Only two patients were from Mid Western region and their isolates contained 3 and 16 copies of IS*6110*.

Figure 8 Distribution of number of copies of IS*6110* between racial groups

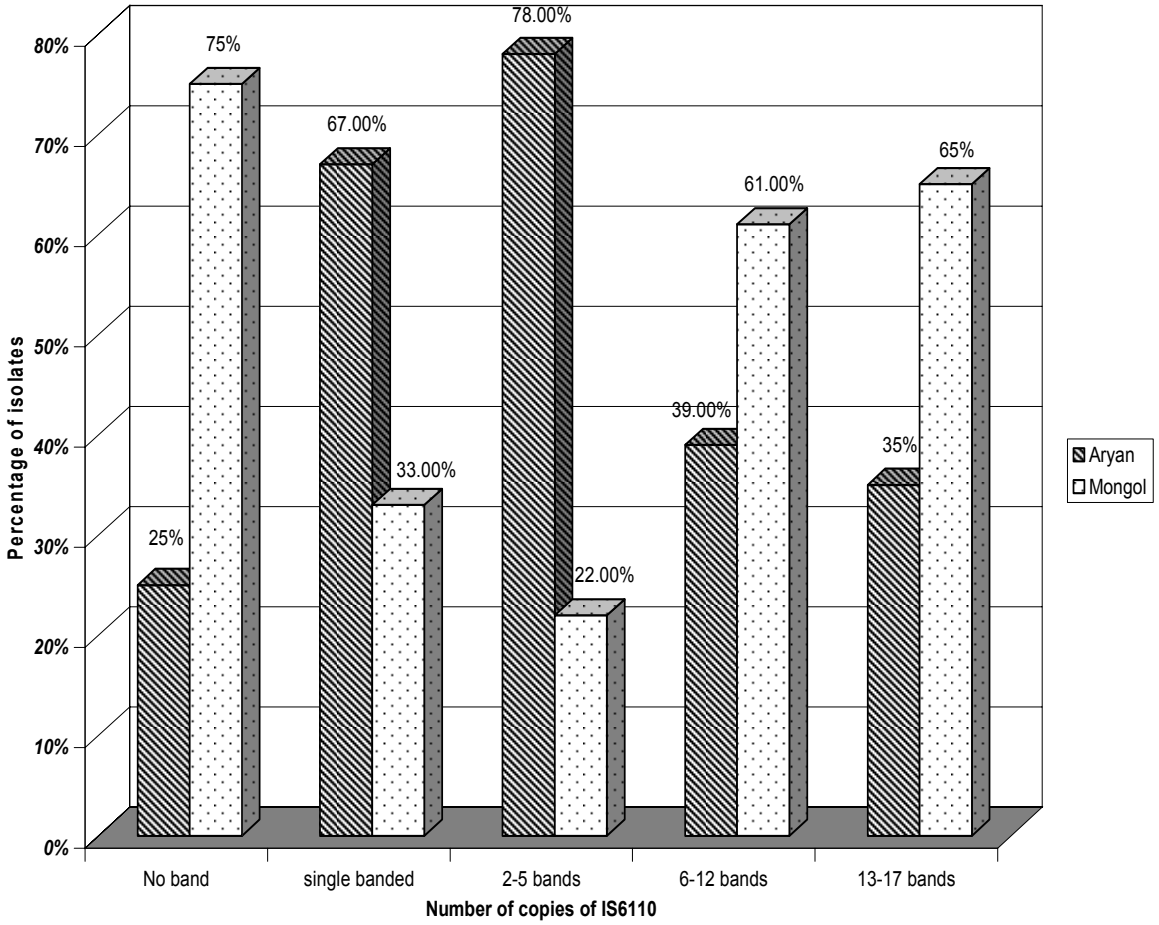


Table 10. Distribution of IS*6110* copy number between racial groups

	No band	Single banded	2-5 bands	6-12 bands	13-17 bands	Total
Aryan	1 (25%)	2 (67%)	7 (78%)	9 (39%)	7 (35%)	26 (44%)
Mongol	3 (75%)	1 (33%)	2 (22%)	14 (61%)	13 (65%)	33 (56%)
Total	4	3	9	23	20	59

The percentage of Aryans in LCS i.e. single banded and 2-5 copy numbered isolates were very high i.e. 67% and 78% in comparison with 33% and 22% of Mongols. The reverse condition was observed for the isolates without IS*6110* and high copy numbered strains (HCS). Only 39% of 6-12 copy numbered isolates and 35% of 13-17 copy numbered isolates were from Aryan racial group in comparison with 61% and 65% of the respective isolates from Mongols. On statistical analysis, the distribution of isolates having different copy number of IS*6110* was found to be significantly heterogeneous between two racial groups tested ($\chi 2$, $P<0.05$).

Table 11. Distribution of caste and ethnicity of patients in two racial groups studied

Race	Ethnicity (No.)	Caste	Number
Aryan	Brahamin (9)	Bhandari	1
		Chohan	1
		Dhakal	1
		Gautam	1
		Koirala	1
		Neupane	2
		Pandey	2
	Chettri (5)	Basnet	1
		Bujhel	1
		Gharti	1
		K.C	1
		Rana	1
	Sudra (5)	BK	2
		Dadal	1
		Dadhraj	1
		Danuwar	1
	Others (8)	Ansari	1
		Balki	1
		Dharti	1
		Majhi	1
		Musalman	1
		Pasi	1
		Rishidev	1
Mongol	Magar (8)	Budha	1
		Thapa Magar	7
	Newar (8)	Dangol	1
		Maharjan	2
		Shrestha	5
	Gurung (1)	Gurung	1
	Lama (8)	Lama	8
	Tamang (8)	Tamang	8

Table 12. The profile of clustered and non-clustered cases of *M. tuberculosis isolates*. The single banded and non hybridized isolates were excluded from this table.

	Clustered cases	Isolated cases
Total isolates in each group	12	40
Percentage of isolates	23%	77%
Male :Female	1.4 :1	2.1 :1
Median age of patients (years)	39.5	35.7

Of the total cases studied, 52 (88%) of isolates had more than one copy of IS*6110*. Analysing these strains only, 23% clustered isolates (12/52) and 77% (40/52) isolated cases were observed, when similarity coefficient of 78% was taken as marginal value to include two isolates in the same cluster. Of the clustered isolates, 10 (83%) had 12 or more copies of IS*6110*.The median age of the patients in clustered, non-clustered isolates and the total cases were found to be 39.5, 35.7 and 34.9 years, which were in the same age group (35-44) years Similarly, the sex ratio of both clustered and non clustered cases were similar. Predominance of male patients was observed in both groups as in total cases.

Table 13. The profile of clustered and non-clustered cases of *M. tuberculosis* isolates. The single banded and non hybridized isolates were excluded from this table.

Characteristics	Cluster A		Cluster B				Cluster C		Cluster D		Cluster E	
No. of isolates	2		4				2		2		2	
Sample code	G2	G27	G14	G25	T16	T36	G43	G54	T86	T124	G52	G67
Similarity coefficent	100%		96.81%				88.23%		84.6%		80.1%	
No. of bands	2	2	16	15	15	16	17	17	12	14	12	13
Male:Female	1:1		1:3				1:1		M		M	
Age, Years	24	53	66	28	23	20	67	17	35	40	70	66
Geo. Region	W	C	W	C	C	C	C	C	E	C	C	C
Aryan :Mongol	Aryan		1:1				Mongol		Mongol		1:1	

Analyzing the banding pattern of the clustered isolates, 5 different clusters with 2 to 4 isolates were found and named as cluster A to Cluster E.

Cluster A contained 2 isolates harbouring two copies of IS*6110*, with identical banding pattern. The mean age of patients was 38.5 with equal sex ratio. Both of patients were Aryans but no geographical relationship was observed.

Cluster B was the biggest cluster observed in this study, containing 4 isolates with 15 to 16 copies of IS*6110* in their genome. Among the 4 isolates, the two isolates with 15 copies of IS*6110*, were identical to each other and most similar (similarity coefficient 96.8%) to other two isolates of the cluster having 16 copies of IS*6110* and vice versa. Three out of four patients in the cluster were females. Fifty percent (2/4) patients were from the age group <24 years, 25% (1/4) of 28 years and 25 % (1/4) of 66years, however the mean age of the cluster was 34.2 years and not significantly different from

the mean age of total patients. Seventy five percent (3/4) patients of this cluster was localized to the central region including 50% (2/4) patients from the same ward of Kavre The distribution of these isolates in both the races studied were found to be equal. Comparing the banding pattern of these isolates with photographs of identified Beijing strains from published papers, the patterns were more or less similar. Finally they were confirmed as Beijing strains (Palittapongarnpim, personal communication)

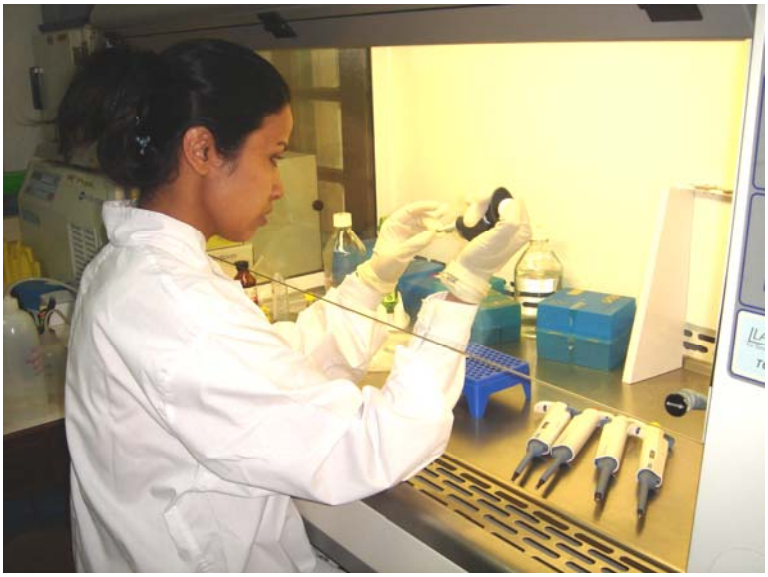
Cluster C contains two isolates with 17 copies of IS*6110*. They had 15 matched hybridized fragments or 88.23% similar to each other. No gender discrimination was observed. Both of the patients were Mongols and from Central region with mean age 42 years, however one of them was >65 and other 17 years.

Cluster D contains two male, Mongol patients with mean age of 37.5 years. Their strains shared 84.6% similarity in IS*6110*-RFLP patterns, however no geographical correlation was observed from the available data.

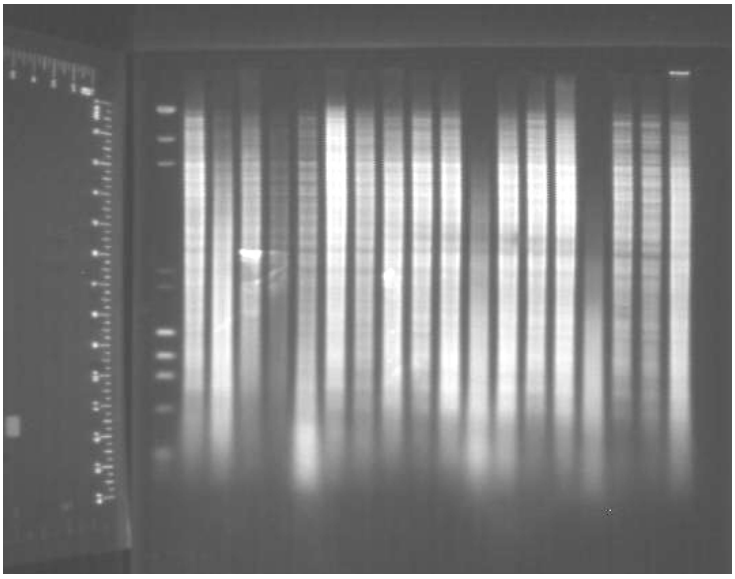
Finally, **the cluster E** also contains two old aged male patients with mean age of 68 years from Central region of the country. Their strains harboured 12 to 13 copies of IS*6110* with 80.1% similarity in their hybridization patterns. The predominance of one race was not observed in this strain like cluster B

Besides the clusters with most similar patterns described above, some more strains were seem to be similar to each other and some to the isolates of defined clusters .However they were discounted from the clusters and grouped in isolated cases as their similarity coefficient were only 66-75% i.e. less than the marginal value we followed.

Photograph 1. Colonies of *M. tuberculosis* on L-J medium

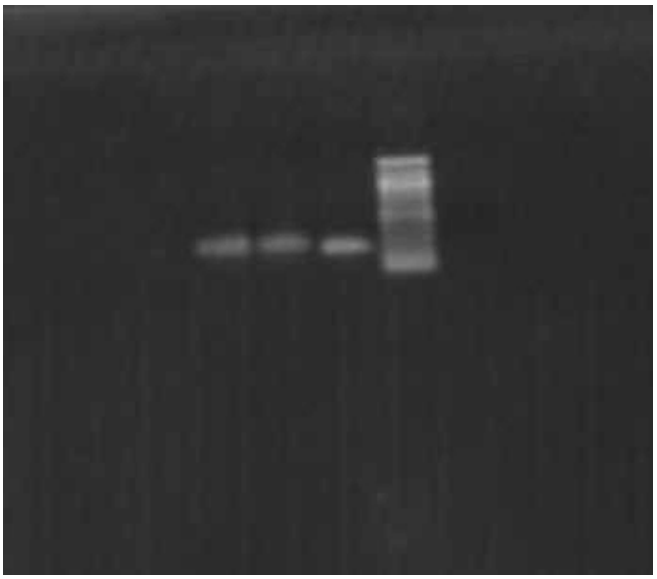


Photograph 2. Processing of PCR in the laboratory of MRL

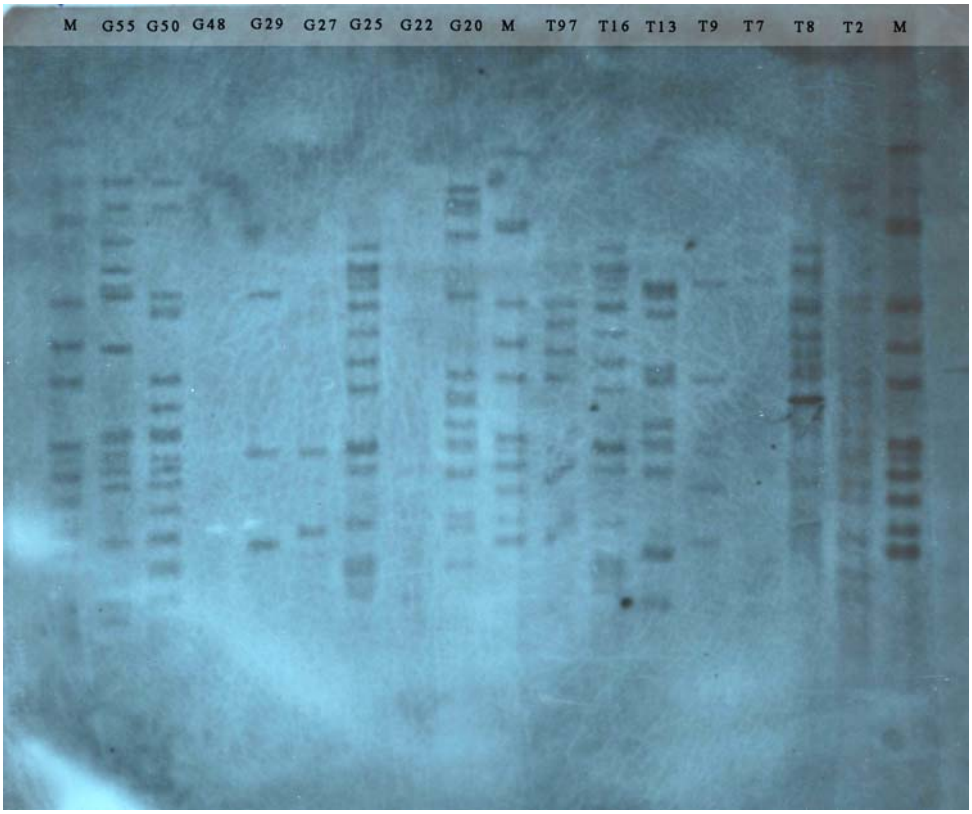


Photograph 3. Overnight run PvuII digested DNA samples of *M. tuberculosis* isolates on 0.8% agarose. The first lane contains molecular weight marker and the last lane is empty.

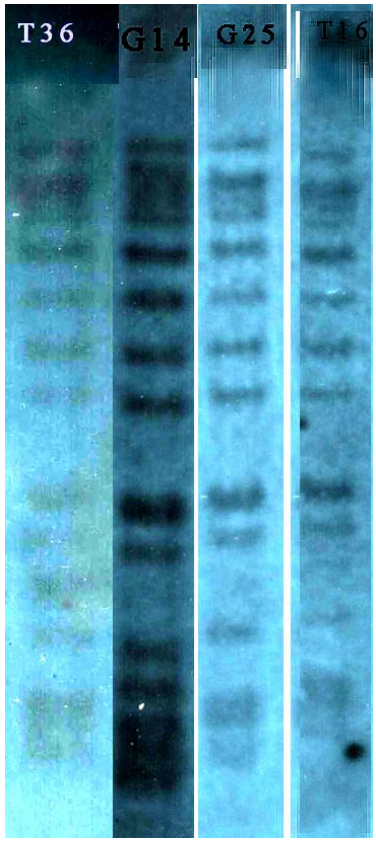
Photograph 4. Capillary transfer set-up for southern blotting of PvuII digested DNA to Hybond N^+ membrane



Photograph 5: PCR amplified DIG-11-UTP labeled IS*6110* probe. D1 and D2 are DIG-11-UTP



Photograph 6: IS*6110*-RFLP banding pattern of *M. tuberculosis* isolates



Photograph 7: Matched IS6110- RFLP pattern of clustered isolates of *M. tuberculosis*

CHAPTER–VI

6. DISCUSSION

The general objective of this study was to determine genomic polymorphisms of *M. tuberculosis* isolates in Nepal by restriction fragment length polymorphism. DNA fingerprinting using the IS*6110* insertion sequence as the basis of the analysis of RFLP is regarded as the standard method for the molecular typing of *M. tuberculosis*. *M. tuberculosis* contains an insertion sequence IS*6110* in its chromosome which belongs to the IS*3* family of enterobacterial insertion sequence (IS) elements. This insert is present in other species of the *M. tuberculosis* complex but not in the more distantly related mycobacteria (Cave *et al.,* 1991). The presence of multiple copies of IS*6110* within the chromosome of *M. tuberculosis* allows the easy detection of *M. tuberculosis* in tuberculous patients by PCR (Sapkota, 2003). Also, irregular insertion of IS*6110* within the genome permits the discrimination of strains of *M. tuberculosis* species (Cave *et al.,* 1991) by RFLP analysis. IS*6110* provides the most variable patterns yet discovered when it is used as a probe for RFLP analysis for the subdivision of *M. tuberculosis* strains

IS*6110*-RFLP provides a versatile tool for the identification of transmission (Stephen *et al*., 2002), the investigation of TB outbreaks (Daley *et al.,* 1992), the distinction between reactivation and reinfection (Small *et al.,* 1993), as well as proof of laboratory cross-contamination (Bauer *et al.,* 1997) in most of developed countries with low incidence cases. Similarly the technique was already found to be successful in tracking the RFLP types of a number of developing countries with high incidence of cases and revealed that, although *M. tuberculosis* isolates from each country had comparatively similar characteristics depending on the classification factor, each country's isolates showed characteristic DNA fingerprints and differed slightly from the isolates of the other countries in the mode number of IS*6110* copies or analyzing more distinctly, isolates from each country had their own peculiar DNA fingerprinting traits (Park *et al.,* 2000). However the situation is different in some developing countries like Nepal, as

one of endemic countries for the disease. There had not been any study on molecular epidemiology of human tuberculosis of Nepalese strains of *M. tuberculosis* until Rijal, (2006) recently used the tool to study MDR tuberculosis up to the molecular level. His study revealed the presence of high genetic polymorphism among Nepalese isolates with the endemic strains from neighboring countries, predominance of Beijing strains and also the presence of local types referred as Nepalese strains. However further investigation with large number of isolates representing the country is necessary for proper identification and characterizations of strain types and to initiate the control measure accordingly. Thus the secondary aim of this study was to optimize the standard operating procedure of IS*6110*-RFLP in the lab of our own country.

6.1 Optimization of standard operating procedure (SOP)

IS*6110*-RFLP has been considered as the gold standard technology in the molecular epidemiology of tuberculosis till date. The standard methodology for this matter has been recommended by van Embden *et al.,* (1993).However it presents some difficulties. First, it requires large quantities (4.5µg) of high molecular weight chromosomal DNA with minimal protein contamination, obtainable only from prolonged cultures or sometimes subculture with considerable delay (20-40 days). The method is expensive and labour-intensive. Before going through the whole protocol, every step was optimized in part one of the study. Of 204 culture positive samples collected from 117 smear positive PTB patients from NTC, 54 (26 %) gave insufficient amount of DNA, 12 (6 %) were of low quality and 33 (16 %) DNA samples were found to be sheared and thus rejected. The quality and quantity of extracted DNA partially depends upon the type of colonies one has chosen for the DNA preparation and partially on the practical exposure of the person in handling chromosomal DNA. For this study, mycobacterial cultures bearing more than 5 colonies were selected, however many samples had small colonies or smooth growth with pin pointed colonies resulting in low quantity or poor quality DNA respectively. On the other hand very large colonies were tough enough to prepare homogeneous solution. Vortexing with micro-glass beads was applied as the alternative for that problem but found to produce sheared DNA. So, in the final

protocol, glass beads were not used, vortexing and pipetting steps were minimized, lyzozyme treatment time was prolonged and final DNA resuspending temperature was minimized to 4°C. Once the protocol for DNA extraction was optimized, 76 samples from GENETUP were extracted in part II of the study and it was found that only 13 (17 %) gave insufficient amount of DNA samples. The remaining 63 (83 %) were good enough in terms of quality and quantity and thus used in RFLP.

Electrophoresis of PvuII digested DNA was performed at different voltages and time periods. For the available electrophoretic unit, 100V for 1 hour followed by 24 V (0.8V/cm) for 22 hours was found to produce interpretable banding patterns i.e. 2kb fragment of lamda/phi molecular weight marker reached 8 cm at this point. Increasing the time of electrophoresis may results in high resolution of banding pattern but could not be applied in this work due to the limitation of working hours per day.

Denaturization and neutralization of gel was performed as indicated by Palittapongarnpim *et al.,* (1997). Depurination of DNA fragments in gel by UV exposure was done as in the protocol of van Soolingen *et al.,* (1993) but further acid depurination followed by strong base treatment step was not performed. Acid depurination is necessary for DNA fragments of >15kb lengths. Shorter fragments may break into pieces so small that they may escape from the membrane during southern blotting (Sambrook *et al.,* 2005). Southern blotting hybridization and detection was done as indicated in the publication of Palittapongarnpim, *et al.,* (1997) except the final documentation step i.e. unlike the colorimetric method with NBT/BCIP they used, auto radiological development was done which has the further advantages of easy handling, is non toxic and can be stored for long term.

6.1.1 Optimization of PCR and DIG labeling of probe

Van Embden *et al.,* (1993) recommended use of a DNA probe 5' to the PvuII digestion site on the physical map of IS*6110*, producing the number of bands in the fingerprint equilibrate to the number of copies of IS*6110* in genome of the bacteria. Most

researchers prepare 245bp IS*6110* DNA probe by PCR followed by horse radish peroxide enzymatic labeling of probe with ready to use kit by Amersham (ECL-kit, direct nucleic acid labeling and detection system) used by van Soolingen *et al.,* (1993). Palittaponggarnpim *et al.,* (1997) used a B*am* HI-*Sal* I fragment of plasmid pDC73 as probe and labeled by non radioactive digoxygenin random primed DNA labeling kit (Roche Diagnostics, Germany). Chauhan *et al.,* (2004) used PCR probe and labeled with Dig labeling kit described above. In this study, the DNA probe was prepared and simultaneously labeled by incorporation of digoxygenin-11-dUTP during PCR as previously described by Dolzani *et al.,* (2004).Unlike his work, the primers used in this study, INS 1 and INS2 were initially 3' end labeled with Dig molecules on either end of primers in addition to incorporation of dig-11-dUTPs during PCR. The detection sensitivity of the resulting probe was found to be enhanced even in use of the half volume of the PCR product recommended by the company.

The PCR conditions were optimized by changing different parameters such as Taq and Mg concentration, the ratio of dig-11-dUTP: dTTP in nucleotide preparations, concentration of template DNA and conditions of thermocycler. After optimization, 1U of Taq, 2.5μM of MgCl, 1:3 nucleotide ratio (50 μM dig-dUTP: 150 μM dTTP) and 1000x dilution of template DNA was found to produce best band. For the thermocycler conditions, van Soolingen *et al.,* (1993) used 96ºC and 65ºC as denaturation and annealing temperature for 1 min period each. However, incorporation of DIG in primer will change the melting temperature and annealing temperature of PCR. Melting temperature for oligonucleotide primers can be calculated from the single formula Tm=2(A+T) + (G+C). Generally a rule of thumb is to use a temperature approximately 5ºC lower than Tm of the primer (www.jax.org/imr/optimize pcr.html). For the primer pair available, 95ºC and 62ºC of denaturation and annealing temperature were found to produce best result at 40 seconds each.

For the matter of labeling, Roche Diagnostics, Germany recommended to use 1:2 ratios of nucleotides for detection of single gene present in the membrane. But for this study, even 1:3 was found to be optimal and produced sensitive probe similar as those labeled

by random primed labeling kit. Labeling methods by kit system either enzymatic (ECL, Amersham) or chemiluminescence's (DIG, Roche) are very expensive and hard to afford by most of research laboratories in developing countries such as Nepal. Incorporation of dig-11-dUTPs during PCR was found to be quite economic in comparison with the methods described above. One further advantage of this method over the former ones is that no purification of template DNA or PCR products was needed before labeling procedure (Roche Diagnostics, Germany).

Thus finally total 116 isolates, 53 (45 %) isolates from 117 selected cases from NTC, and 63 (83 %) isolates from GENETUP were subjected to the final study. Of them, 25 (47 %) of isolates from NTC and 34 (54 %) from GENETUP were analyzed finally for their IS*6110* RFLP patterns.

6.2 Result analysis

Out of total 116 isolates subjected to RFLP, fingerprints from 59 (51 %) isolates were analyzed. Among the isolates with un-interpretable results, some had incomplete banding pattern, some with very faint bands, few had problem with DNA. The IS*6110* hybridization patterns were inspected visually as done previously by palittapongarnpim *et al.,* (1997) and Chauhan *et al.,* (2004). The numbers of bands were counted for each isolate. Among the similar patterns, numbers of matched bands were determined and similarity coefficient was calculated to group the isolates in different clusters.

6.2.1 Analysis of fingerprints by number of copies of IS*6110*

The number of IS*6110* copies per isolate in population varied from 0-17 bands. The result was identical with North Indian isolates (Chauhan *et al.,* 2004) and similar with 0-19 copy numbered Thai-isolates (Rienthong *et al.,* 2005). Further classifying the result in three groups namely A, B and C (having either no copies, 1-4 copies or 5-17 copies of IS*6110*, respectively), 7 % belongs to A, 19 % isolates to group B and 75 % isolates to group C. This division agrees closely with the observations of north Indian isolates (Chauhan *et al.,* 2004).

Most predominant strains have their peculiar banding pattern with specified number of copies of IS*6110*. Although the number of copies of IS*6110* in the genome of *M. tuberculosis* can range from 0–25, population-based molecular epidemiological studies report that most strains contain 8–18 copies, a number sufficient to enable discrimination between the majorities of strains. The numbers of copies were found to be reduced in some regions of Asia and Africa (Burgos and Pym, 2003). On the other hand, other predominant strains from East Asia such as Beijing strain harboured multiple copies with minimum 13 to 20 copies of IS*6110* (Toungoussova *et al.,*2002; van Soolingen *et al.,*1995).

The strain types observed in this study were categorized in 5 groups on the basis of number of copies of IS*6110* per isolate for further analysis of the strain types in respect to different parameters studied. They were no band, single banded, 2-5 bands, 6-12 bands and 13-17 bands. The classification helped to compare the strains with the previously identified strain types on the basis of number of copies of IS*6110*.

DNA of 4 (7 %) isolates of *M. tuberculosis* did not hybridize to the probe, suggesting that they contained no IS*6110*. These strains were not observed in the only previous study from Nepalese isolates (Rijal, 2006). However this observation was in concordance with the results (8 %) from North Indian isolates but much high than those (1%) from South Indian isolates (Chauhan *et al.,* 2004 and Narayan *et al.,* 2002).These strains were observed only in certain regions of Asia and are the limitations for the IS*6110* based DNA fingerprinting techniques. Methods based on other epidemiological markers could be applied for these strains.

Three (5 %) isolates hybridized at only one position. Of these, the hybridized fragment of two isolates was of same length (1.4kb) and third with different length around 0.8kb. The percentage of isolates containing a single copy of IS*6110* was low in this study, unlike that previously reported in patients from Nepal (26 %), and other countries in Asia, such as India, Vietnam and Thailand but the fragment length in two isolates were of similar size as the isolates from Thailand (Rijal *et al.,* 2006; Das *et al.,* 1995; Yuen *et*

al., 1993; Palittapongarnpim *et al.,* 1997). The high frequency of these single-copy strains, at least in some countries, indicates a high degree of stability of this pattern, presumably due to the virtual absence of transposition of the element in such strains. Except South East Asian countries the percentage of single banded isolates in this study was similar to the most studies that have reported less than 5% single-copy strains from not only Europe but also China and Mongolia (Hermans *et al.,* 1995; van Soolingen *et al.,* 1995). Although, the single hybridized bands appeared at the same position, an epidemiological relationship between the isolates could not be implied (van Soolingen *et al.,* 1993) since the single copy of IS*6110* usually locates at the site common to all strains which has been referred to as a 'hot spot' in the bacterial chromosome (Hermans *et al.,* 1991). This can be further confirmed by the southern hybridization with the probes specific to DR and PGRS or other PCR based fingerprinting techniques which can differentiate these strains.

Nine (15 %) isolates in this study harbored 2-5 copies of IS*6110* which was similar with 17% of isolates from African city Bostuwana however very high in comparing with isolates from Hong Kong (2 %) and low in comparison with 26% reported from South Indian isolates, 20% from Thailand and 37% from Alabama. (Lockman *et al.,* 2001; Chan-Yeun *et al.,* 2003; Narayan *et al.,* 2002; Rienthong *et al.,* 2005; Kempf *et al.,*2005). These LCS observed in Thailand were grouped as Heterogeneous (H) type because they had heterogeneous banding pattern, quite different from the major characterized groups they studied. Similarly, in the previous study from Nepal, 13 % of the isolates studied were of this (H) type with <6 bands. The result was in favor with the observation of the present study. However as gel compare software analysis was not performed in this study, the strains could not be confirmed as Heterogeneous type (H). Low copy number strains were observed in most of countries studied but their frequency is common in some regions of Asian and African countries and uncommon in other parts of them. Association of LCS with drug resistance was found to be low comparing with Beijing and heterogeneous with more than 5 bands (Rienthong *et al.,* 2005). No resistance was observed in LCS types from Nepal (Rijal, 2006). Differentiation of *M.*

tuberculosis strains carrying fewer than six IS*6110* copies is poor when IS*6110*-RFLP alone is used. It has been shown that matching IS*6110* RFLP patterns with at least 6 bands is more predictive for clonality than patterns with fewer bands. So, secondary molecular typing techniques with additional DNA probes are required to differentiate these strains (Kempf *et al.,* 2005).

IS*6110*-RFLP was found to be quite discriminating for the strains harboring more than 6 copies of IS*6110*. Most of countries categorized the isolates in to low copy numbers and high copy number considering the 6 copies as marginal one. By following this principle, the present study was enabled to type 73 % of the clinical isolates analyzed which is in agreement with most of Asian countries, lower than that found in European countries but high than south Indian isolates where low copy number strains was found to be prevalent (Narayan *et al.,* 2002).

The fourth group of strains analyzed contained 6-12 copies of IS*6110* in their genome. It was the largest group observed in this study in terms of number of copies of IS*6110* and represented by 39 % of the isolates analyzed. The mode number of copies was found to be 12. The majority copy number range and mode numbers of copies can be considered as country specific. Most Chinese, Mongolian, and Hong Kong isolates were found to have 20 copies of IS*6110*. Mongolian and Hong Kong isolates had similar distributions in terms of the number of IS*6110* copies and the presence of 21 copies with single mode. Similarly Korean isolates and Fillipino isolates had 10 and 13 copies as mode numbers respectively (Park *et al.,* 2000). In the previous study of the Nepalese isolates, Rijal (2006) observed a group of isolates with 8-12 copies of IS*6110* and heterogeneous banding pattern that was not reported from other countries till date and so considered as local type of Nepal. In present study also, 8-12 copy numbered strains were found to be most predominant type and was represented by 21 (36 %) of isolates analyzed. Thus they may be heterogeneous (h) Nepal type, however further confirmation should be done by comparing with international database.

The second largest group of isolates (34 %) observed in this study harbored 13-17 copies of IS*6110* element. These HCS were found to be predominant in most of countries. The most successful strain (Beijing family) harbored 13-17 copies IS*6110* in Russia however 15 to 20 copies was the predominant number for the family (Toungoussova *et al.,* 2002; van Soolingen *et al.,* 1995). The Beijing strains were prevalent in China (85%), Mongolia (50%), South Korea (43%), Thailand (37%), Vietnam (50%), and Indonesia (34%) suggests that this clone spreads in Southeast Asia where TB is endemic, whereas they were rarer in other countries in Middle East Asia, Africa, South America, and Europe. Beijing family strains were defined as containing only 9 of the 43 spacer sequences by spoligotyping and a 3.6-kb PvuII fragment by IS*1081* fingerprinting (van Crevel *et al.,* 2001; van Soolingen *et al.,* 1995). Other groups of strains such as Nonthaburi (11-14) from Thailand or the F11 family (11-19) from South Africa were found to contain high copy numbers. Some of the strains in the present study seem to be Beijing type and some Nonthaburi group by visual comparision with published papers (Palittapongarnpim, personal communication). However, in depth analysis either by standard software or other secondary typing should be done for proper identification and characterization of typed strains.

6.2.2 Analysis of fingerprints with respect to gender

Males were found to be predominant with the ratio of 1.8: 1 for male: female in the total population analyzed. Further analyzing the fingerprints on the basis of number of copies of IS*6110* in each gender specifically, great variation was not observed for low copy numbered strains in comparison with the percentage of isolates representing each gender for high copy numbered strains. Fourty seven percent of male patients were found to be infected with *M. tuberculosis* isolates harboring 6-12 copies of IS*6110*. In contrast, 43 % of the female patients harbored 13-17 copied strains. As stated before, 6-12 copy number strains were the predominant type observed in this study and may be heterogeneous type of Nepal. Similarly, 13-17 copied strains may be Beijing type as these strains were endemic to most of South East Asian countries; sharing of north border of the country with China and also by the previous study in which predominance

of these strains was already revealed (Rijal, 2006). The latter study revealed 42 % of Beijing type and 29 % of heterogeneous Nepal type isolates were MDR. This might be interpreted as the evidence that both strains were transmitted in the country with dominance of single gender for each strain type and there seem to be increasing risk for transmission of MDR strains for the respective gender. Thus, further studies on this matter should be done with caution.

6.2.3 Analysis of fingerprints with in different age groups

The isolates analyzed were from the patients aged between 17-79 years. Of these, the highest percentage was aged below 25 years and the number of isolates decreased with increase of patient's age. High genetic polymorphism was observed in young age groups in terms of number of copies of IS*6110* where as old-aged patients above 55 years were found to be infected with only HCS. Tuberculosis in old age is generally due to reactivation of previously infected strains rather than reinfection with new ones. Thus, the involvement of HCS in old patients indicates that these strains have been circulating in the country for a long time. On the other hand, predominance of these strains in every age group studied suggests that they were endemic to the country and therefore available for the transmission over a considerable period of time among the residents with in that population. Similarly, LCS were observed only in the patients below 55 years which suggests that these strains have not been circulating in the country for a long time so that they could be distributed fairly evenly in all age groups as the endemic strains but again their presence in youngsters indicates them as the spreading types. Beijing strains were found to be associated with younger age in agreement with the hypothesis that these strains were spreading. In contrast the reverse condition was observed for the single band isolates suggesting them as the common isolates of the past in many countries, e.g. Vietnam Thailand and Malaysia (Dale *et al.,* 1999; Rienthong *et al.,* 2005; Anh *et al.,* 2000). The unusual observation in the present study might be due to bias in proportion of age groups. However, median age for LCS and HCS were found to be 32.5 and 36.16 years and no significant association was observed in various age groups in terms of number of copies of IS*6110*.

6.2.4 Analysis of fingerprints in different geographical regions

In the present study, the selections of samples in terms of geographical distribution of patients seemed to be biased as the study relied on the hospitals located in the central region. However genotypic variation observed among the regions seems to be useful for portraying the distribution of genotypes in the country. Eighty percent of the patients studied were from the central region and only 3 % from mid western region. Analyzing the banding pattern among the regions, all groups of strains with 0-17 copies of IS*6110* were observed in central region. Besides there, single bands were observed only in Eastern region where high copy numbered strains were absent. In contrast, the reverse condition was observed in Western region. Nepal is a land locked country with various geographical conditions between China and India. It was suggested that ancestors of the Beijing family might have originated in China and had clonally expanded for less than one hundred years. Consequently, the family had previously been reported in many countries of South East Asia where Ethnic Chinese migration from the same area comprise a significant proportion of the total population (Palittapongarnpim *et al.,* 1997). Such a significant migration was unknown in Nepal. However, two way business migrations have always been observed for a long time through Tibet. The prevalence of high copy numbered strains in Western and Central region and absence of them from Eastern region indicate that the Western region might be the portal of entry of 13-17 copy numbered strains (may be Beijing strains) for the country.

The predominance of LCS, especially single banded, were previously reported in India (Narayan *et al.,* 2002). The Southern border of Nepal is shared by India and there has always been two way open migration of people. Presence of single copy numbered isolates in Eastern region even in comparatively small fraction of isolates might indicate the predominance of these strains in that region.

Besides the two major strains assumed to be transported from the territories, the country should have its own genotypes. The assumption was supported by the observation that 8-12 copy numbered strains (might be heterogeneous Nepal type) were distributed

apparently in all regions studied except in Mid Western (MW). Their absence from MW might be due to presence of comparatively low proportion of isolates representing that region. The complementary explanation for geographical distribution of strains is that 6-12 copy numbered strains have been circulating in Nepal for such a long time that they are distributed fairly evenly in all regions. On the other hand predominance of 13-17 copy number over these local types in Western region may indicate their possible replacement with high copy numbered strains (may be Beijing family). Presence of all strain types in Central region might be the reflection of its diverse as well as crowded population structure. However, a long term population based study should be done to confirm this.

6.2.5 Analysis of distribution of fingerprints between racial groups

The distribution of RFLP banding pattern on the basis of number of copies of IS*6110* between two racial groups tested i.e. Aryans and Mongols was found to be significantly different ($\chi 2$, $P<0.05$). The association of Black race with LCS has previously been reported (Kempf *et al.,* 2005).

The distribution of both racial groups in the study population and in the total population of the county was comparable. Although their migration around the world and to Nepal may be different,. Aryans were assumed to come to Nepal from India. Further, it is assumed that they invaded India through Iran, whereas the Mongols came to Nepal through China. From the literature, it seems that LCS types were reported from southern part of India, whereas the HCS were reported from China which is situated close to the northern part of Nepal. LCS (<6 copies) and HCS (>5 copies) were predominant in Aryans and Mongolians population respectively. The distribution of Aryan and Mongol is not homogenous itself in Nepal. Mostly the Mongols are predominant in Northern part of Nepal which is nearby the border of China. This suggests that either the HCS were coming from China or the Mongols are more susceptible to them. The same arguments can be made for the Aryans as they are more predominant to the southern parts of Nepal which is near India. This further suggests that there is the chance of

getting the mixed or heterogeneous local strains in Nepal. The observation of the present study was in agreement with the observations reported from these countries with predominance of respective racial groups (van Soolingen *et al.,* 1995; Narayan *et al.,* 2002). Similarly, the data of E.U, 2006 was also in favour with the observation and shows that the proportion of TB attributable to the Beijing genotype is variable: high in Asia, apart from the Indian subcontinent, increasing further east; low in parts of Africa, Latin America, and Western Europe; intermediate in the United States and Cuba; low in Eastern Europe (other than the former Soviet Union); low in the Middle East including <1% in a recent study from Tehran. In Western Europe Beijing genotype is more common among immigrants. The expansion of the W-Beijing family strains throughout the Asian continent may be as a result of their resistance to the BCG vaccine (Bifani *et al.,* 2002). However, the LCS were found to be most predominant comparing with all reported cases till date, at the study site in India where the world's largest *Mycobacterium bovis* BCG trial was conducted. As the present study was only concerning about number of copies of IS*6110*, further studies with characterization of strain types should be done to reflect a clear association of strain with races and also to their genetic susceptibility towards the respective strains.

6.2.6 Analysis of clustered samples

Among the 59 isolates analyzed, 48 (81 %) different patterns were observed which is in agreement with the 80 % observed in Indian isolates (Vijaya Bhanu *et al.,* 2004). Excluding the non hybridized isolates and single banded ones, 12 (23 %) isolates were grouped in 5 clusters and 40 (77 %) were observed to be isolated ones. The percentage of clustering observed in present study was in agreement with the results from Thailand 26 %, where as very high in comparison with 9% of South Indian isolates excluding the cluster of single banded ones, 8 % of Japan and low in comparison with 52 % observed in Spain (Rienthong *et al.,* 2005; Narayan *et al.,* 2002; Fujikane *et al.,* 2004; Ruiz Garcia *et al.*, 2002).

In contrast to the widely held belief that IS*6110* fingerprinting shows poor discrimination with LCS, the degree of clustering was not higher for those patterns with two to five copies (17 % clustering) than for those with more than five copies (23 % clustering). So, for subsequent analyses the data were therefore used to divide the isolates into two groups: single-copy strains and isolates with more than one copy.
Analyzing the population structure of clustered patients, no age and gender variation was observed. However 3 out of 4 patients in cluster B, which was confirmed as Beijing strains (Palittapongarnpim in personal communication) were females and of age below 30 years. The observation might indicate the recent transmission of these strains. The Beijing strain was also found to be more common in younger patients in Vietnam (Dale *et al.,* 1999) which was interpreted as the expansion of the strain.

The patients whose strains exhibit an identical (or highly similar) fingerprint pattern during a defined period of time are included in a so-called cluster. However, considerable caution should be exercised in conducting and interpreting these studies. Groups of strains may be identical for reasons other than recent transmission, depending, for example, on the stability of the marker and the number of strains in the population over time. Cases actually due to recent transmission may not be seen as clustered if they are new immigrants to the population or if not all cases in the population are included in the study. The amount of clustering seen will depend on the duration of the study (Glynn *et al.,* 1999). The present study was not a time dependent. Also, patient contact tracing and their epidemeological corelation was not studied. So, presence of the clustering could not be interpreted as recent transmission. The strains might be endemic ones to the population. However presence of the large cluster of Beijing strains even in small study population and also the predominance of them within youngsters should be taken with caution.

Despite the most similar banding patterns discribed above, many strains were seen to be similar with similarity coefficent below 75%, however were excluded from the cluster analysis as the 78% of similarity coefficient was the minimum requirement for the grouping system we followed. Groups of *M. tuberculosis* isolates with very similar

hybridization patterns are commnly found in developing countries. The most important factor accounting for the grouping was reported to be the geographical area, implying that transmission of the bacteria was high in these countries (Palittapongarnpim *et al.,* 1997).

Non-inclusion of large population representing all geographical regions and lack of identification and characterization of strains has made the present study limited to analyze the clusters in depth. However this study has succeeded in pointing out the insights of molecular epidemiology in the context of Nepal and their indespensable necessity for clinical management of tuberculosis and their control programmes in future.

CHAPTER-VII

7. SUMMARY AND RECOMMENDATIONS

7.1 SUMMARY

Restriction Fragment Length Polymorphism was carried out using protocol described by van Embden *et al.,* (1993) with some modifications. PCR product IS*6110* probe was used as described in van Embden's protocol and subsequently labeled with DIG-dUTPs during PCR as indicated by Dolzani *et al.,* (2004) and Roche, (2006).Analysis of autoradiographs was done visually. Strains were typed on the basis of number of copies of IS*6110* per isolate. Then based on the number of copies of IS*6110* in visual inspection, strains were further analyzed for their banding pattern and grouped in different clusters by calculating their similarity coefficient as described by Palittapongarnpim *et al.,* (1997).

Of 59 *M. tuberculosis* isolates analyzed, the number of copies of IS*6110* varied from 0 to 17 with the mode number of 12 copies. When strains were further classified in 5 groups harbouring 0 copy, single copy, 2-5 copies, 6-12 copies and 13-17 copies of IS*6110*, 6-12 copied strains were found to occur most frequently in all age groups, geopolitical regions and races studied. Patients of ≥55 years were found to be infected with strains harbouring ≥6 copies of IS*6110*. Distribution of IS*6110* copy number between two races tested (Aryans and Mongols) was found to be significantly different. Low and high copy numbered strains were found to be predominant in Aryans and Mongol patients respectively.

Further analyzing the banding pattern of individual strain types, 48 different patterns were observed showing 81 % of genetic variations. Twelve isolates were found to be in 5 clusters. The largest cluster was represented by four 15-16 copy numbered strains predominantly found in younger age groups and females.

From the present study it can be concluded that a high genomic variations exists among the *M. tuberculosis* isolates circulating in the country. On the other hand, some strains are more frequently transmitting in the populations than others.

7.2 RECOMMENDATIONS

The present study succeeds in resolving some questions regarding molecular epidemiology of *M. tuberculosis* in Nepal. But some questions were generated which can be taken as future research projects. The following recommendations have been made from the findings of this study.

1. Optimized RFLP protocol used in this study can be utilized to conduct population based studies to understand molecular epidemiology of *M. tuberculosis* in Nepal
2. It is recommended to use secondary typing methods for strains with single or no copies of IS*6110* like DR-RFLP, PGRS-RFLP.
3. The distribution of copy number of IS*6110* was found to be heterogeneous between two races in this study. Further research with significantly larger population should be done to confirm their genetic susceptibility towards respective strains.
4. Strains from different geopolitical regions were seemed to be different in this study. Future, researches from each region should be recommended to understand the transmission dynamics of the bacteria with in and between the regions and to compare them with the strains across the borders of the country
5. One of molecular clusters was observed predominantly in youngsters, female patients from central region. This should be taken with caution and thus time dependent researches on this matter should be strongly recommended.
6. In addition, the present study might pave the way for any molecular epidemiological studies of *M. tuberculosis* in Nepal concerning on different parameters of the disease such as HIV status, MDR patterns, relapses reinfections and treatment failure cases.

CHAPTER VIII

8. REFERENCES

Abed Y, Bolted C and Mica P (1995a) Identification and strain differentiation of *Mycobacterium* species on the basis of DNA 16S-23S spacer region polymorphism IS*6110*, an IS-like element of *M. tuberculosis* complex. Nucleic Acids Res. 18:188

Abed Y, Davin-Reglia A, Bollet C and De Micco P (1995b) Efficient discrimination of *M. tuberculosis* strains by 16S-23S spacer region-based random amplified polymorphic DNA analysis. J Clin Microbiol 33:1418-1420

Alland D, Kalkul GE, Moss AR, Hahn JA, Bosworth W, Drucker E and Bloom BR.(1994) Transmission of Tuberculosis in New York city and analysis by DNA fingerprinting and conventional epidemiological method. New England Journal of Medicine 350:1710-1716

American Thoracic Society and Center for Disease Control (1981) The tuberculin skin test. Am Rev Respir Dis 124:356-363

American Thoracic Society and Center for Disease Control (2000) Diagnostic standards and classification of tuberculosis in adults and children. Am J Respire Crit Care Med 161:1376-95

Ananthanarayan R. and Panikar CKJ (1996) Textbook of microbiology. Fifth edition, Orient Longman Limited, Hyderabad, India

Anderson P (1997) Host responses and antigens involved in protective Immunity to *Mycobacterium tuberculosis*, scand. J Immunol, 45:115-131

Anh DD, Borgdorff MW, van LN, Lan NTN, van Gorkom T, Kremer K, and van Soolingen D (2000) *Mycobacterium tuberculosis* Beijing Genotype Emerging in Vietnam Emerg Infect Dis 2000;6:302–305

Bauer J, Thomsen VO, Poulsen S, Andersen B (1997) False-positive results from cultures of *Mycobacterium tuberculosis* due to laboratory cross-contamination confirmed by restriction fragment length polymorphism. J Clin Microbiol 35:988–991

Bauer J, Yang Z, Poulsen S and Andersen AB (1998) Results from five years of nationwide DNA fingerprinting of *Mycobacterium tuberculosis* complex isolates in a country with a low incidence of *M. tuberculosis* infection. J Clin Microbiol 36:305-308

Betts JC, Dodson P, Selwyn Q, Lewis AP, Thomas PJ, Duncan K and McAdam RA (2000) Comparison of the proteome of *Mycobacterium tuberculosis* strain H37Rv with clinical isolate CDC 1551. Microbiology 146:3205-3216

Bifani PJ, Mathema B, Kurepina NE and Kreiswirth BN (2002) Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. Trends Microbiol.10 (1):45-52

Blackwood KS, Assaad AA, Elliott LJ, Hershfield ES and Kaban AM (2003) Conventional and molecular epidemiology of Tuberculosis in Manitoba. BMC Infect Dis. 2003; 3: 18.

Bloom BR (1994) Tuberculosis-Pathogenesis, Protection and Control. American Society for Microbiology, Washington, DC, USA

Blot M (1994) Transposable elements and adaptation of host bacteria. Genetica 93:5–12

Brosch RWJ, Stavropoulos PE, Colston MJ, Cole ST and Gorden SV (1999) Genomic analysis reveals variation between *Mycobacterium tuberculosis* H37Rv and the attenuated *M. tuberculosis* H37Ra strain. Infec Immun 67:57668-57674

Burgos MV and Pym AS (2002) Molecular epidemiology of tuberculosis. Eur Respir J 20:54S-65S

Burgos M, DeRiemer K, Small PM, Hopewell PC, and Daley CL (2003) Effect of drug resistance on the generation of secondary cases of tuberculosis. J Infect Dis 188:1878–1884

Camus JC, Pryor MJ, Medigue C and Cole ST (2002) Reannotation of genome sequence of *Mycobacterium tuberculosis* H37Rv. Microbiology 148:2967-2973

Casper C, Singh SP, Rane S, Daley CL, Schecter GS and Riley LW (1996) The transcontinental transmission of tuberculosis: a molecular epidemiological assessment. Am J Public Health 86: 551-553

Cave MD, Eisenach KD, McDermott PF, Bates JH and Crawford JT (1991) IS*6110*: conservation of sequence in the *Mycobacterium tuberculosis* complex and its utilization in DNA fingerprinting. Mol Cell Probes 5:73–80

Centers for Disease Control (1991) Purified protein derivatives (PPD)-tuberculin anergy and HIV infection: guidelines for anergy testing and management of anergic persons at risk of tuberculosis. Morbid Mortal Weekly Rep 40:27-33

Chakraborty P (2003) A textbook of microbiology. Second edition, New Central Book Agency (P) LTD 8/1, Chintamoni Das Lane, Kolkota 700 009, India

Chambers MAD, Gavier-Widen and Hewinson RG (2004) Antibody bound to the surface antigen mp8 of *Mycobacterium bovis* enhances survival against high dose and low dose challenges, FEMS. Immunol Med Microbial 41:93-100

Chandler M and Mahillon J (2002) Insertion sequences revisited. In: Craig NL, Craigie R, Gellert M and. Lambowitz A M (ed.) Mobile DNA II. American Society for Microbiology, Washington, D.C.

Chan-Yeung M, Tam CM, Wong H, Leung C, Wang J, Yew WW, Lam CH and Kam KM (2003) Molecular and Conventional Epidemiology of Tuberculosis in Hong Kong:

a Population-Based Prospective Study. Journal of Clinical Microbiology 41(6):2706-2708

Charlesworth B, Sniegowski P and Stephan W (1994) The evolutionary dynamics of repetitive DNA in eukaryotes. Nature 371:215–220

Chauhan A, Chauhan DS, Parashar D, Gupta P, Sharma VD, Sachan AS, Gupta R, Agarawal BM and Katoch VM (2004) DNA fingerprinting of *Mycobacterium tuberculosis* isolates from Agra region by IS*6110* probe. Indian Journal of Medical Microbiology; 22(4):238-240

Cheesbrough M (1989) Medical Laboratory Manual for Tropical Countries.Volume II, Microbiology. ELBS.

Cole ST,Brosch R, Parkhill J, Garnier T, Chursher C, Harries D, Gordon SV,Eiglmeier K, Gas S, Barry CE, Tekaia F, Badcock, Basham D, Brown D, Chilling-Worth T, Connor R, Davies D, Devlin K, Feltwell T, Gentles S, Hamlin N, Holroyd S, Hornsby T, Jagels K, Krogh A, McLean J, Moule S, Murphy L, Oliver K, Osborne J, Quail MA, Rajandream MA, Rogers J, Rutter S, Seager K, Skelton J, Squares R, Squares S, Sulston JE, Taylor K, Whitehead S and Barrell BG (1998) Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. Nature 393:537-544

Cole ST, Eisenach KD, Mc Murry DN and Jacobs WR (2005) Tuberculosis and Tubercle Bacillus, American Society for Microbiology, Washington, DC, USA: 191-199

Crawford JT and Bates JH (1984) Phase typing of Mycobacteria. In: Kubica GP Wayne LW The Mycobacteria, a source book, part A , Marcel Dekker, Inc., New York pp 123-132

Dale JW, Mat Nor R, Ramayah S, Tang TH, and. Zainuddin JF (1999) Molecular Epidemiology of Tuberculosis in Malaysia. J Clin Microbiol; 37(5):1265–1268

Daley CL, Small PM, Schecter GF, Schoolnik GK, McAdam RA, Jacobs Jr WR and Hopewell PC (1992) An outbreak of tuberculosis with accelerated progression among persons infected with the human immunodeficiency virus. An analysis using restriction-fragment-length polymorphisms. N Engl J Med 326:231–235

Daniel TD (1994) Tuberculosis. In: Isselbacher KJ, Braunwald E, Wilson JD, Martin JB, Fauci A and Kasper DL (eds) Harrison's principles of internal medicine, 13th edition, Vol. 1, Mc Graw-Hill Inc. Health Profession Division

Das S, Chan SL, Allen BW, Mitchison DA, Lowrie DB (1993) Application of DNA fingerprinting with IS986 tosequential mycobacterial isolates obtained from pulmonary tuberculosis patients in Hong Kong before, during and after short-course chemotherapy. Tuberc Lung Dis 74:47-51

Das S, Paramasivan CN, Lowrie DB, Prabhakar R and Narayanan PR (1995) IS*6110* restriction fragment length polymorphism typing of clinical isolates of *Mycobacterium tuberculosis* from patients with pulmonary tuberculosis in Madras, South India. Tuber Lung Dis 76:550-554

De Boer AS, Borgdorff MW, de Haas PE, Nagelkerke NJ, van Embden JD and van Soolingen D (1999) Analysis of rate of change of IS*6110* RFLP patterns of *Mycobacterium tuberculosis* based on serial patient isolates. J Infect Dis 180:1238–1244

Dolzani L, Rosato M, Sartori B, Banfi E, Lagatolla C, Predominato M, Fabris C, Tonin E, Gombac F and Monti-Bragadin C (2004) Mycobacterium tuberculosis isolates belonging to katG gyrA group 2 are associated with clustered cases of tuberculosis in Italian patients.J Med Microbiol 53:155-159

Doran TJ, Hodson AL, Davies JK and Radford AJ (1993) Characterization of a highly repeated DNA sequence from *Mycobacterium bovis*. FEMS Microbiology Letters 111: 147–152

Douglas JT, Qian L, Montoya JC, Musser JM,. van Embden JDA, Soolingen DV and Kremer K (2003) Characterization of the Manila Family of *Mycobacterium tuberculosis.* J Clin Microbiol 41(6):2723–2726

Edlin BR, Tokaras JL, Grego MH, Crawford JT, Williams J, Sordillo EMK, Kilburn JO, Dooley SW, Castro KG, Jasvis WR and Holmberg SD (1992) An outbreak of multidrug resistant tuberculosis among hospitalized patient with acquired immunodeficiency syndrome. New England Journal of medicine 326:1514-1521

Eisenach KD, Cave MD, Bates JH and Crawford JT (1990) Polymerase chain reaction amplification of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. J Infect Dis 161:977–981

Enarson DA, Rieder HL, Arnadottir T and Trebucq A (2000) Management of tuberculosis: a guide for low-income countries. 5th edition, International Union against Tuberculosis and Lung Disease (IUATLD), 68 boulevards Saint-Michel, 75006 Paris, France

Enarson DA, Rieder HL, Arnadottir T and Trébucq A (1996). Management of Tuberculosis: A Guide For Low Income Countries. 4th edition, International Union Against Tuberculosis and Lung Disease (IUATLD), 68 boulevard Saint-Michel, 75006 Paris, France

European Concerted Action on New Generation Genetic Markers and Techniques for the Epidemiology and Control of Tuberculosis (2006) Beijing/W genotype *Mycobacterium tuberculosis* and drug resistance. Emerging Infectious Diseases; 5 (1).

Forbes BA, Sahm DF and Weissfeld AS (1998) Bailey and Scott's Diagnostic Microbiology. 10th edition, Mosby, Inc. USA

Foulds J and O'Brien R (1998) New tools for the diagnosis of tuberculosis: the perspective of developing countries. Int J Tuberc Lung Dis 2:778-783

Friedman CR, Stoeckle MY, Johnson JR and Riley LW (1995) Double-repetitive-element PCR method for sub-typing *M. tuberculosis* clinical isolates. Journal of Clinical Microbiology 33:1383–1384

Frothingham R (1995) Discrimination of *M. tuberculosis* strains by PCR. J Clin Microbiol 33:2801

Fujikane T, Fujiuchi S, Yamazaki Y, Matsumoto H; Takahashi M., Fujita Y, Shimizu T and Kikuchi K (2004) Molecular epidemiology of tuberculosis in the north Hokkaido district of Japan. The International Journal of Tuberculosis and Lung Diseasee 8:39-44

Glennon M and Smith T (1995) Can random amplified polymorphic DNA analysis of the 16S-23S spacer region of *M. tuberculosis* differentiate between isolates? J Clin Microbiol 33:3359

Glynn JR, Bauer J, de Boer AS, Borgdorff MW, Fine PEM, Godfrey-Faussett and Vynnycky P (1999) European Concerted Action on Molecular Epidemiology and Control of Tuberculosis.Interpreting DNA fingerprint clusters of *Mycobacterium tuberculosis*. The International Journal of Tuberculosis and Lung Disease 3(12):1055-1060

Grange JM (1990) Tuberculosis In: Parker MT and Duerden BI (eds.) Topley and Wilson's Principle of Bacteriology, Virology and Immunity. Vol. II, 8th edition, Butler and Tanner Ltd, London: 93-118

Greenwood D, Slack RCB and Peutherer JF (2002) Medical microbiology. 16th edition, Churchill Livingstone, UK: page number

Groenen PM, van Bunschoten AE, van Soolingen D, van Embden JD (1993) Nature of DNA polymorphism in the direct repeat cluster of *M. tuberculosis*, application for strain differentiation by a novel method. Mol Microbiol 10:1057-1065

Groothius DC and Yates MD (eds.) (1991) Diagnostic and Public Health Mycobacteriology. Bureau of Hygiene and Tropical Diseases, England

Harries A, Maher D, Upleker M., Chaulet P and Nunn P (1998) National tuberculosis Programme. A clinical Mannual for Nepal. 1st edition, National Tuberculosi Centre, Thimi

Hermans PWM, Messadj F, Guebrexabher I.I., van Soolingen D, de Haas PWE, Heersam H, de Neeling H, Ayoub A, Portaels F, Frommel D, Zribi M, and van Embden JDA (1995) Analysis of the population structure of *Mycobacterium tuberculosis* in Ethiopia, Tunisia and the Netherlands: usefulness of DNA typing for global tuberculosis epidemiology. J Infect Dis 171:1504-1513

Hermans PW, van Soolingen D, Dale JW, Schuitema AR, McAdam RA, Catty D and van Embden JDA (1990) Insertion element IS986 from *Mycobacterium tuberculosis*: a useful tool for diagnosis and epidemiology of tuberculosis. J Clin Microbiol 28:2051-2058

Hermans PW, van Soolingen D, Bik EM, de Haas PE, Dale JW and van Embden JD (1991) Insertion element IS987 from *Mycobacterium bovis* BCG is located in a hot-spot integration region for insertion elements in *Mycobacterium tuberculosis* complex strains. Infect Immun 59:2695-2705

Hermans PW, van Soolingen D and van Embden JD (1992) Characterization of a major polymorphic tandem repeat in *M. tuberculosis* and its potential use in the epidemiology of *Mycobacterium kansasii* and *Mycobacterium gordonae*. Journal of Bacteriology 174, 4157–4165.

Hornick D B (1993) Tuberculosis:Public Health Preventive Medicine:208-219

Huebner RE, Schein MF and Bass JB (1993) The tuberculin skin test. Clin. Infect. Dis. 17:968-975

Iwamoto T, Sonobe T and Hayashi K (2003) Loop-mediated isothermal amplification for direct detection of *Mycobacterium tuberculosis* complex, *M. avium*, and *M. intracellulare* in sputum samples. J Clin Microbial 41:2616–2622

Kanduma E, Mc Hugh TD and Gillespie SH (2003) Molecular Methods for *M. tuberculosis* strain typing: A Users Guide. J Applied Microbiology 94:781-791

Kato-Maeda M, Bifani PJ, Kreiswirth BN and Small PM (2001) The nature and consequence of genetic variability within *Mycobacterium tuberculosis*. Clin Invest 107(5):533-537

Kempf MC, Dunlap NE, Lok KH, Benjamin Jr. WH, Keenan NB, and Kimerling ME (2005) Long-Term Molecular Analysis of Tuberculosis Strains in Alabama, a State Characterized by a Largely Indigenous, Low-Risk Population. Journal of Clinical Microbiology: 870-878

Kox LFF (1996) Diagnosis of tuberculosis and other mycobacterioses: development and clinical evaluation of PCR assay. A thesis of department of biomedical research, Royal Tropical Institute, Amsterdam and Division of Pulmonary Disease, Academic Medical Center, Amsterdam

Kremer K, Glynn JR, Lillebaek T, Niemann S, Kurepina NF, Kreiswirth BN, Bifani PJ and van Soolingen (2004) Recognition of the Mycobacterium tuberculosis Beijing family on the basis of three genetic markers. J Clin Microbiol.; 42:4040-4049

Kremer K, van Soolingen D, Frothingham R., Haas WH, Hermans PWM, Martin C, Palittapongarnpim P, Plikaytis BB, Riley LW, Yakrus MA, Musser JM and van Embden JDA (1999) Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: inter-laboratory study of discriminatory power and reproducibility. Journal of Clinical Microbiology 37:2607–2618

Lalvani A, Pathan AA and Durkan H (2000) Rapid detection of *Mycobacterium tuberculosis* infection by enumeration of antigen-specific T cells. Am J Respir Crit Care Med 164(4):824-8

Lauzoro M and Ashken D (2000) Phthisiology at the drawn of new century. A review of tuberculosis and the prospects for its elimination. Chest 2:288-295

Lillebæk T (2005) The molecular epidemiology of tuberculosis Recent trends in a low burden country. Danish Medical Bulletin 52:143-59

Lillebaek TA, Dirksen I, Baess B and Strunge VO, Thomsen and Andersen AB (2002) Molecular evidence of endogenous reactivation of *Mycobacterium tuberculosis* after 33 years of latent infection. J Infect Dis 185:401–404

Lockman S, Sheppard JD, Braden CR, Mwasekaga MJ, Woodley CL, Kenyon TA, Binkin NJ, Steinman M, Montsho F, Kesupile-Reed M, Hirschfeldt C, Notha M, Moeti T and Tappero JW (2001) Molecular and Conventional Epidemiology of *Mycobacterium tuberculosis* in Botswana: a Population-Based Prospective Study of 301 Pulmonary Tuberculosis Patients. Journal of Clinical Microbiology:1042-1047

Madigan MT, Martinko JM and Parker J (1997) Brock Biology of Microorganisms. 8th edition, Prentice Hall International, Inc, USA: 787-788

Maguire H, Dale JW, Mc Hugh TD, Butcher PD, Gillespie SH, Costetsos A, Al-Ghulsein H, Holland T, Dickens A, Marston L, Wilson P, Pitman R, Strachan DA, Drobniewisky FA and Manerjee DK (2002) Molecular Epidemiology of Tuberculosis in London demonstrating low rate of active transmission. Thorax 57:617-622

Maher D, Chaulet P, Spinaci S and Harries A (1997) Treatment Of Tuberculosis: Guidelines For National Programmes for the Global Tuberculosis Programmes. 2nd edition, World Health Organization (WHO), Geneva, Switzerland

Mahillon J and M. Chandler (1998) The insertion sequences. Microbiol Mol Biol Rev 62:725-774

Mathema B, Kurepina NE, Bifani PJ and Kreiswirth BN (2006) MolecularEpidemiology of Tuberculosis: Current Insights. Clinical microbiology: 658–685

McAdam RA, Hermans PW, van Soolingen D, Zainuddin ZF, Catty D, van Embden JD and Dale W (1990) Characterization of a *M. tuberculosis* sequence belonging to the IS3 family. Molecular Microbiology 4:1607–1613

McHugh TD, Dickens A and Gillespie SH (2000) False molecular clusters due to non-random association of IS*6110* with *Mycobacterium tuberculosis*. Journal of Clinical Microbiology 38:2081–2086

McMurray DN (2001) Baron's Medical Microbiology Textbook. Online textbook, http://gsbs.utmb.edu/microbook/toc.htm

Moström P, Gordon M, Sola C, Ridelland M and Rastogi N (2002) Methods used in the molecular epidemiology of tuberculosis. Clinical Microbiology and Infection 8:694

Meinkoth J and Wahl G (1984) Hybridization of nucleic acids immobilized on solid supports. Anal Biochem 138(2):267-84

Narayanan S, Das S, Garg R, Hari L, Vijay Rao B, Frieden TR, and Narayanan PR (2002) Molecular Epidemiology of Tuberculosis in a Rural Area of High Prevalence in South India: Implications for Disease Control and Prevention. J Clin Microbiol 40(12):4785–4788

National Tuberculosis Control Programme, Annual Report (2004/2005) NTC, Thimi, Bhaktapur,Nepal

Nguyen D, Proulx JF, Westley J, Thibert L, Dery S and Marcel AB (2003)Tuberculosis in the Inuit Community of Quebec. Canada American Journal of Respiratory and Critical Care Medicine 168:1353-1357

Niemann S, Harmsen D, Rüsch-Gerdes S and Richter E (2000) Differentiation of clinical *Mycobacterium tuberculosis* complex isolates by gyrB DNA sequence polymorphism analysis. J. Clin. Microbiol. 38: 3231-3234

Niemann S, Rusch-Gerdes S and Richter E (2000) Stability of IS*6110* restriction fragment length polymorphism patterns of *Mycobacterium tuberculosis* strains in actual chains of transmission. J Clin Microbiol 38:2563–2567

Otal I, Martín C, Vincent-Lévy-Frebault V, Thierry D and Gicquel B (1991) Restriction fragment length polymorphism analysis using IS*6110* as an epidemiological marker in tuberculosis. J Clin Microbiol 29(6):1252–1254

Palittapongarnpim P, Luangsook P, Tansuphaswadikul S, Chuchottaworn C, Prachaktam R, Sathapatayavongs B (1997) Restriction fragment length polymorphism study of *Mycobacterium tuberculosis* in Thailand using IS*6110* as probe. Int J Tuberc Lung Dis1:370–376

Palmero D, Ritacco V, Ambroggi M, Natiello M, Barrera L, Capone L, Dambrosi A, Di Lonardo M, Isola N, Poggi P, Vescovo M and Abbate E (2003) Multidrug-Resistant Tuberculosis in HIV-Negative Patients, Buenos Aires, Argentina. Emerg Infect Dis 9(8):965-69

Park YK, Bai GH and Kim SJ (2000) Restriction Fragment Length Polymorphism Analysis of *Mycobacterium tuberculosis* isolated from Countries in the Western Pacific Region. J Clin Microbiol 38(1):191–197

Poulet S and Cole ST (1995) Characterization of the highly abundant polymorphic GC-rich-repetitive sequence (PGRS) present in *M. tuberculosis*. Archives of Microbiology 16, 87–95.

Ravins M, Bercovier H, Chemtob D, Fishman Y and Rahav G, (2001) Molecular Epidemiology of *Mycobacterium tuberculosis* Infection in Israel Journal of Clinical Microbiology, Vol. 39, No. 3, p. 1175-1177

Rastogi N, Legrand E and Sola C (2001) The mycobacteria: an introduction to nomenclature and pathogenesis. Rev Sci Tech 20(1):21-54.Res Microbiol 146: 405-413

Rienthong D, Ajawatanawong P, Rienthong S, Smithitikarn S, Akarasewi P, Chaiprasert A and Palittapongarnpim P (2005) Restriction fragment length polymorphism study of nation wide samples of *Mycobacterium tuberculosis* in Thialand, 1997-1998. Int J Tuberc Lung Dis 9(5):1-6

Rijal BP (2006) Study of multi-drug resistance tuberculosis in Nepal A PhD dissertation submitted to Central Department of Microbiology, Tribhuvan University

Roche Diagnostic, Germany (www.roche-applied-science.com)

Rosenberg NA, Tsolaki AG and Tanaka MM (2003) Estimating change rates of genetic markers using serial samples: applications to the transposon IS*6110* in *Mycobacterium tuberculosis*. Theor Popul Biol 63:347–363

Ruiz Garcia M, Rodriguez JC, Navarro JF, Samper S, Martin C and Royo G (2002) Molecular epidemiology of tuberculosis in Elche, Spain, A 7 year study. J Med Microbiol 51:273-277

SAARC (2004) Tuberculosis in SAARC region, an update 2004. SAARC Tuberculosis Centre, Thimi, Bhaktapur

Salfinger M, Demchik BS and Kafader FM (1990) Comparison between the MB check system, radiometric and conventional methods of recovery of mycobacteria. J Microbiol Method 12:97-100

Sambrook J and Russell DW (2001) Molecular Cloning, A laboratory manual, 3rd edition, Cold spring Harbor Laboratory press, Cold spring Harbor New York

Sambrook J, Fritsch EF, and Maniatis T (1989). Molecular Cloning: A Laboratory Manual. Vol. 2. Second edition. Cold Spring Harbor Laboratory Press, USA.

Sapkota BR (2003) Optimization and use of polymerase chain reaction (PCR) for the diagnosis of Tuberculosis and Leprosy. An academic dissertation presented to the Central Department of Microbiology, Tribhuvan University, Kirtipur, Kathmandu, Nepal

Small PM and Moss A (1993) Molecular epidemiology and the new Tuberculosis. Infect Agents Dis 2:132-138

Small PM, Schafer RW, Hopewell PC, Singh SP, Murphy MJ, Desmond E, Sierra MF and Schoolnik GK (1993b) Exogenous reinfection with advanced HIV infection. N Engl J med 328:1137-1144

Snider DE Jr, Jones WD and Good RC (1984) The usefulness of phase typing *Mycobacterium tuberculosis* isolates. American review of respiratory diseases 130:1099

Weis SE, Pogoda JM, Zhenhua Y M, Cave D, Wallace C, Kelley M and Barnes PF (2002) Transmission Dynamics of Tuberculosis in Tarrant Country, Texas. American Journal of Respiratory and Critical Care Medicine 166:36-42

Struelens MJ (1998) Molecular Epidemiologic Typing Systems of Bacterial Pathogens: Current Issues and Perpectives. Mem Inst Oswaldo Cruz. 93(5):581-586

Strump FIJ, Tsang AY and Sayre JW (1979) Reevaluation of sputum staining for the diagnosis of pulmonary tuberculosis. Am Rev Respir Dis 199:599-602

Suffy PN, Araujo MEI and Degrave WM (1997) The Changing Face of the Epidemiology of Tuberculosis due to Molecular Strain Typing - A Review. Mem Inst Oswaldo Cruz, Rio de Janeiro 92(3):297-316

Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B and Locht C (2000) Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. Molecular Microbiology 36:762–771

Tekaia F, Gordon SV, GarnierT, Brosch R, Barrell BG and Cole ST (1999) Analysis of the proteome of *Mycobacterium tuberculosis* in silica. Tubercle Lung Dis 79:329-342

Theirry D, Cave MD, Eisennach KD, Crawford JT, Bates JH, Gicquel B and Guesdon JD(1990) IS*6110*, an IS-like element of *Mycobacterium tuberculosis* complex. Nucleic Acids Res. 1990; 18:188

Thomson N, Sebaihia M, Holden M, Cerdeno-Tarraga A and Park Hill J (2001) Size matters? Trends Microbiol 9:359

Toungoussova OS, Per S, Mariandyshev AO, Nizovtseva NI, Bjune G and Caugant DA (2002) Spread of Drug-Resistant *Mycobacterium tuberculosis* strains of the Beijing Genotype in the Archangel Oblast, Russia. J Clin Microbiol 40(6):1930–1937

van Crevel R., Nelwan RHH, de Lenne W, Veeraragu Y, van der Zanden AG, Amin Z, van der Meer JWM, and van Soolingen D (2001) *Mycobacterium tuberculosis* Beijing genotype strains associated with febrile response to treatment. Emerg Infect Dis 7:880-883

van Deutekom H, Hoijng SP, de Haas PEW, Langendam MW, Horsman A, van Soolingen D and Coutinho RA (2004) Clustered Tuberculosis Cases Do They Represent

Recent Transmission and Can They Be Detected Earlier? American Journal of Respiratory and Critical Care Medicine 169:806-810

van Embden JDA, van soolingen D, Small PM and Hermans PWM (1992) Genetic markers for epedimiology of tuberculosis. Res microbial 143:385-391

van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, Hermans P, Martin C, Mc Adam R and Shinnick TM (1993) Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. J Clin Microbiol 31:406–409

van Embden JD, van Gorkom T, Kremer K, Jansen R, van Der Zeijstqq BA and Schouls LM (2000) Genetic variation and evolutionary origin of the direct repeat locus of *Mycobacterium tuberculosis* complex bacteria. J Bacteriol 182:2393–2401

van Soolingen D, Qian L, De Haas PE, Douglas JT, Traore H, Portaels F, Qing HZ, Enkhsaikan D, Nymadawa P and van Embden JD (1995) Predominance of a single genotype of *M. tuberculosis* in countries of East Asia. J Clin Microbiol 33:3234-3238

van Soolingen D, Borgdorff MW, de Haas PE, Sebek MM., Veen J, Dessens M, Kremer K and van Embden JD (1999) Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. J Infect Dis 180:726–736

van Soolingen D, De Haas PE, Hermans PW, Groenen PM and van Embden JD (1993) Comparison of various repetitive DNA elements as genetic markers for strain differentiation and epidemiology of *M. tuberculosis*. Journal of Clinical Microbiology 31:1987–1995

van Soolingen D, Hermans PW, de Haas PE, Soll DR and van Embden JD (1991) Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J Clin Microbiol 29:2578–2586

van soolingen D, Hermans PWM, De Haas PEW and van Embden JEA (1992) Insertion element IS1081-associated restriction fragment length polymorphism in *Mycobacterium tuberculosis* complex species: A reliable tool for recognizing *Mycobacterium Bovis* BCG. J Clin Microbiol 30:1772-1777

Vernon EC, Diane MJ, Reid SJ, and Rybicki EP (1998) Detection of Nucleic acids by Hybridization.

Vijaya bhanu N, Banavalikar JN, Kapoor SK and Seth P (2004) Suspected small-scale interpersonal transmission of mycobacterium tuberculosis in wards of an urban hospital in Delhi, India. Am J Trop Med Hyg 70(5):527- 531

Warren RM, van der Spuy GD, Richardson M, Beyers N, Borgdorff MW, Behr MA and van Helden PD (2002) Calculation of the stability of the IS*6110* banding pattern in patients with persistent *Mycobacterium tuberculosis* disease. J Clin Microbiol 40:1705–1708

Watt B, Rayner A and Harris G (1996) Mycobacteria. In: Collee JG, Fraser AG, Marmion BP and Simmons A (eds.) Mackie and McCartney Practical Medical Microbiology. 14th edition, International Student Edition, Churchill Livingstone:

Wayne LG and Kubica GP (1986) The mycobacteria, p.1435-1457. *In* P.H.A. Sneath, and *et al.,* (ed.), Bergey's Manual of Systematic Bacteriology. Vol. 2. Section 16. Williams and Wilkins Publications

WHO (1997) TB: A clinical manual for South East Asia. World Health Organization, Geneva, Switzerland

WHO (1998 a) Laboratory services in tuberculosis control. Part I: microscopy. World Health Organization, Geneva, Switzerland

WHO (1998 b) Laboratory services in tuberculosis control. Part II: microscopy. World Health Organization, Geneva, Switzerland

WHO (2003) Treatment of Tuberculosis: Guidelines for national programmes, 3rd edition, Geneva:11

WHO (2004) TB/HIV a clinical manual. World Health Organization, Geneva, Switzerland

Yang ZH, Chaves F, Barnes PF, Burman JW, KoehlerJ, Eisenach KD, Bates JH and Cave DM (1996) Evaluation of a method for secondary DNA typing of *M. tuberculosis* with pTBN12 in epidemiological study of TB. Journal of Clinical Microbiology 34:3044–3048

Yuen KY, Chan KS, Yam WC, Ho PL and Chau PY (1995) IS*6110* based amplityping assay and RFLP fingerprinting of clinical isolates of *M. tuberculosis*. Journal of Clinical Pathology 48:924-928

Zainuddin Z F and Dale J W (1989) Polymorphic repetitive DNA sequences in *Mycobacterium tuberculosis* detected with a gene probe from a *Mycobacterium fortuitum* plasmid. J Gen Microbiol.135:2347–2355.

APPENDIX-I

Equiptments and reagents used in NTC

1. Equipments

Biological safety cabinet, Class IIA	:	Kubota, Japan
Autoclave	:	Kubota, Japan
Coagulator	:	Hirasawa, Japan
Centrifuge	:	Kubota, Japan
Incubator	:	IKemoto, Japan
Microscope	:	Olympus, Japan
Refrigerator	:	Sanyo, Japan

1. Reagents for fluorochrome staining

A. For auramine solution

Auramine	1 gm/liter
Phenol	30ml/litre
Distilled water	1000ml
Ethanol (95%)	100ml

B. 20% Sulphuric acid

H_2SO_4	200ml
Distilled water	800ml

C. 0.1% Methylene blue

Methylene blue powder	1 gm
Distilled water	1000ml

2. Solution for sample decontamination

A. NaOH solution

Stock solution: 40% NaOH (10 M) (100 ml)

NaOH pellet	40 gm
Distilled water	100 ml

Working solution: 4%NaOH (500ml)

40%NaOH	50 ml
Distilled water	450ml

Glassware's

Miscellaneous

Inoculating loop

Bunsen burner

Cotton

Forceps

Gloves

Staining rack

Spirit lamp Soaps

Tube holder

Tissue paper

Trays

Bacteriological media

(A) Modified Ogawa Medium (2%)

1. Preparation of salt solution 500ml/flask

Potassium dihydrogen phosphate	2.0 gm
Magnesium citrate	0.1 gm
Sodium glutamate	0.5 gm
D/W	100ml

- Mix well and heat at 100°C for 30 minutes in a water bath (or autoclave at 121°C for 15 minutes).
- Add glycerol 4 ml into the salt solution (while it is hot).
- Add 4ml of 2% malachite green solution.

2. Preparation of whole egg homogenate

- Wipe off eggshell with spirit cotton.
- Break down the egg into a plate to check the decomposition.
- Transfer the egg into the beaker (500ml).
- Homogenize the egg with a pair of chopsticks until the egg become watery.
- Place the two layers of sterile gauze pieces on the funnel.
- Filter the egg homogenate until to get 200ml.

3. Mix 1 with 2 (Raw Modified Ogawa Medium)

4. Distribution of raw medium

   Dispense the medium 6 ml into each tube (avoid bubble formation).

5. Inspissations

   Arrange the tubes in slant position and coagulate them at 90°C for 1 hr with caps closed loosely.

6. Store at 4°C -6°C with caps closed tightly.

(B) Lowenstein-Jensen Medium (L-J Medium used in MRL)

This medium is used for the culture of mycobacterial species (except *M. leprae*).

Ingredients

i) Mineral salt solution

Potassium dihydrogen phosphate, anhydrous [KH_2PO_4]	2.40 gm
Magnesium sulphate [$MgSO_4$]	0.24 gm
Magnesium citrate, quadrihydrate [$Mg_3 (C_6H_5O_7)_2.4H_2O$]	0.60 gm
Asparagine	3.60 gm
Glycerol	12 ml
Distilled water	600 ml

Dissolve these ingredients in the water by heating and through mixing. Boil the solution for 2 hours in a steamer. Sterilize in an autoclave at 121^{o}C and 15-lbs/sq. in. pressure for 20 minutes.

ii) Malachite green solution

Prepare malachite green solution (2 gm/dl) in sterile condition. Placing the container in 37^{o}C incubator for 2 hours will help to dissolve the dye.

iii) Egg solution

Prepare egg solution, using fresh eggs (not more than 4 days old). Depending on the size of the eggs, 20-24 eggs are needed. Wash the eggs carefully with soap and warm water, using the brush. Rinse in running water for 30 minutes. Drain the water. Place eggs in sterile tray. Cover with sterile paper and leave to dry until next day. (This can be hastened by cleaning shells of eggs with methylated spirit and then burning off the excess spirit.)

Then thoroughly wash hands with soap and warm water. Rinse in running water, then with spirit and dry. Crack the eggs with a sterile knife and empty contents into a sterile beaker and beat with a sterile eggbeater.

Combine these three solutions as follows:

Mineral salt solution	600 ml
Malachite green solution	20 ml
Beaten egg solution	1000 ml

Mix, then distribute in 5 ml aliquots in sterile screw-capped tubes or bottles.

Lay tubes or bottles on their sides and inspissate at 80-85^{o}C for 50 minutes. Medium can be leave in the inspissator overnight and then inspissate again at 80-85^{o}C for 30 minutes the next day.

APPENDIX- II

Equiptments used in MRL

Incubator 37 ºC,

Heat block 65 ºC, for Eppendorf tubes 1.5 ml

Water bath 37 ºC

Microcentrifuge centrifuge

Vortex

Freezer -20 ºC

Refrigerator

Electrophoresis unit for small gel

Power Supply

Polaroid camera for Photodocumentation system

Hybridizer

Oven

PCR machine- T3 thermocycler Whatman Germany

Horizontal electrophoresis unit (large gel 20x20 cm), Whatman, Germany

Southern Blotting set –Gel tank, glass plate, weight (1kg)

UV transilluminator, Ultra Violate products (UVP)

Microwave o**ven-Lifes Good (LG)**

pH meter

Analytic Balance

Plastic boxes for washing of membrane

Box for large gel incubation

Development and fixing of film

Hypercassette Kodak

Dark room, red lamp

Gel doc and Polaroid camera

APPENDIX-III

Reagents and Disposables used for IS*6110*-RFLP

(A) Reagents

Lysozyme (Himedia India)

SDS (sodium dodecyl/lauryl sulphate

Cetyl Trimethyl Amonium Bromide (CTAB)-Central Drug House (CDH), India

Sodium Chloride- Central Drug House (CDH) India

ProteinaseK _Sigma, U.S.A

Chloroform-Merck,India

Isoamyl alcohol-Sigma, U.S.A

Isopropanol-Sigma U.S.A

Ethanol

Trisma base-Sigma, U.S.A

EDTA Sigma, U.S.A

Boric acid-Sigma, U.S.A

Agarose-Sigma, U.S.A

Bromophenol blue-Sigma, U.S.A

Molecular weights marker (100bp)-New England Biolabs

Ethidium bromide

PvuII –New Enland Biolabs(NEB)

Molecular weights markers: lambda-HindIII/ PhiX174/HaeIII (NEB)

Nucleotides dNTPs(Roache diagnostic,Germany)

Dig-11-dUTPs for labeling of probe (Roches Diagnostics, Germany)

Taq-polymerase Banglore Genei

PCR buffer (10x)-Banglore Genei

Magnesium chloride Banglore Geinie

Oligonucleotides primers-3' end labeled with DIG used for IS*6110*-PCR (MWG)

INS1: 5'- CGT GAG GGC ATC GAG GTG GC

INS2: 5'- GCG TAG GCG TCG GTG ACA AA

Sodium Hydroxide-Qualigens,India

Tri-Sodium citrate-Merck,India

Hydrochloric acid, Merck- India

Whatman 3MM and blotting papers –Whatman,Germany

Membrane Amersham Hybond N+ RPN303B

Tissue Stalk,Picaso

Saran Wrap

Blocking Reagent- Boehringer Mannheim,Germany

Maleic Acid- Sigma, U.S.A

Sodium-N-Lauryl Sulphate

Ice cubes

Tween-20

Fixer Solution

Developer

(B) Disposables

Plastic tubes 0.5 and 1.5 ml (e.g., Eppendorf, Treff)

Plastic tubes 30 ml (e.g., Falcon)

PCR tubes 0.2ml Molecular bio products

Microfuge tubes 1.5ml-Griener Germany

Falcon tubes 50ml Griener Germany

Propyline tubes 20 ml Interpath

Blottig tissue 24X21 Picaso

Micropettes- 20µl, 50 µl. 200 µl, 1000 µl-Biohit,

pipettes-5 and 10 ml

Gloves-Sterile and Non sterile -Surgicare, India

APPENDIX–IV

Solutions used in IS*6110* -RFLP

(1) Tris EDTA **1×** buffer pH 8.0 (**10mM Tris-HCl, 1mM EDTA**)

Tris	0.3gm
EDTA	0.093gm
D/W	250ml
pH 8.0 with HCl	

(2) 10% SDS (sodium dodecyl/lauryl sulphate)

10 g SDS (Sigma) /100 ml of sterile distilled water. Dissolve by leaving at hot water bath (65^{o}C). Do NOT autoclave. Store at room temperature (>20°C); at cool temperature it becomes crystallized and should be heated in water bath to be ready for use.Note: SDS is harmful when inhaled. Wear gloves and mask when weighing.

(3) CTAB/NaCl (10% CTAB, 4.1%NaCl)

Dissolve 4.1 g NaCl in 80 ml preheated (70-80°C) distilled water. Weigh separately 10 g of CTAB, add it (carefully stirring) to the just prepared NaCl solution. Place at water bath (70°C) to completely dissolve. Adjust with distilled water to the final volume 100 ml. Do NOT autoclave. Store at room temperature (>20°C). This is very viscous solution. You need to preheat it at 70°C before use. CTAB is harmful when inhaled. Wear gloves and mask.

(4) ProteinaseK _ (10 mg/ml) Sigma Dissolve at room temperature. Aliquot (200µl per tube) and store at +4°C

(5) NaCl (5M)

NaCl powder	73.05 g
dH_2O	up to 250 ml

(6) Ethanol 70%- 30ml of Ethanol+ 70ml of D/W (store at -20°C)

(7) chloform-isoamylic alcohol –24:1.
Add 4 ml of Isoamyl alcohol to 96ml of Chloroform. Store at RT in dark bottle.

(8) Tris Boric EDTA (5×)

Trizma base	27.25g
Boric acid	13.9g
EDTA	1.85g
dH_2O	up to 500ml

(9) Bromphenol blue (60% W/V sucrose, 0.25% W/V BPB in TE 1× pH 8)
Add 6 gm of Sucrose to 10 ml of 1x TE. Then add 0.025gm of BPB to it .Store at RT.

(10) Ethidium bromide 0.5 mg/ml (1000 ×): preferably purchase a solution, not a powder. Store in dark at +4 ºC. Ethidium bromide is mutagenic!

(11) Denaturation solution (NaCl 1.5M, NaOH 0.4 M)

NaCl	43.83 g
NaOH	10g
dH_2O	upto 500 ml

(12) Neutralization solution (NaCl 1.5M, Tris 1M pH8.0)

NaCl	87.6645 g
Trisma base	30.275 ml
HCl	2.3085
dH_2O	upto ml

(13) SSC 20× (NaCl 3M, Tri-sodium citrate dihydrate 0.3M, pH 7.0)

NaCl	175.3
Sodium Citrate	88.2
Distilled water	upto 1000ml

(14) Maleic acid (1M) (RT)

Maleic acid	116.08g
NaOH	until pH 7.5
Distilled water	upto 1000ml

(15) Pre-hybridization / hybridization solution for membrane

20X SSC	12.5ml
10% Blocking reagent	10ml
10% N-laurylsarcosine	0.5ml
10% SDS	0.1ml
Formamide	25 ml
dH_2O	1.9ml
	50ml

(16) SSC Wash 2× for membrane:

20× SSC	5ml
10% SDS	0.5ml
dH_2O	44.5ml
	25ml

(17) SSC Wash 0.1× for membrane:

20× SSC	0.25ml
10% SDS	0.5ml
dH_2O	49.25ml
	50ml

(18) Detection Buffer 1 for membrane:

10X Maleic Acid Buffer	3ml
dH_2O	27ml
	30ml

(19) Detection Buffer 2 for membrane:

10× Maleic Acid Buffer	1ml
10% Blocking reagent	1ml
dH_2O	8ml
	10ml

(20) Detection Buffer 3 for membrane:

0.1M Tris pH 9.5	2ml
5M NaCl	0.4ml
dH_2O	17.6ml
	20ml

(21) Washing Buffer for membrane:

1× Maleic Acid Buffer	20ml
Tween 20	2
	20ml

(22) Alkaline probe striping solution (0.2N NaOH, 0.1%SDS)

1N NaOH	8ml
10% SDS	0.4ml
dH_2O	31.6ml

(23) Sodium hydroxide NaOH (1N)

NaOH	0.8gm
dH_2O	10ml

(23) Developer for X-ray films (Kodak)

(24)Fixer (Kodak)

APPENDIX V

Classification of Mycobacteria

Mycobacterial species - Potential Pathogens in Humans

Complex name of closely related species	Species	Clinical significance
Tuberculosis complex	*M. tuberculosis* *M. bovis*	Both species are pathogenic and cause tuberculosis; they regularly exhibit susceptibility to anti-TB drugs; they are communicable
Leprosy group	*M. leprae*	Causes leprosy; can't be cultivated in the laboratory media; can be grown in mouse footpad or in Armadillos, where the temperature is favorable for growth (2-5^{o}C below that of most mammals)
Runyon group		
I. Photochromogens	*M. kansasii* *M. marinum* *M. simiae*	Usually pathogenic; susceptibility to isoniazid and rifampicin. Usually pathogenic causes "swimming pool granuloma"; only in skin lesions (never in sputum); susceptibility to streptomycin. Nonpathogenic; resistant to anti-TB drugs.
II Scrofulaceum complex (Scotochromogens)	*M. scrofulaceum* *M. szulgai*	Commonly nonpathogenic; resistant to anti-tuberculosis drugs. Newly recognized species; susceptible to ethambutol and rifampicin.
III Battey-avium complex (nonphotochromogens)	*M. avium and/or* *M. intracellulare* *M. xenopi* *M. ulcerans*	Both species are usually pathogenic; most strains are resistant to antituberculosis drugs. Commonly nonpathogenic; susceptible to isoniazid and streptomycin. Usually pathogenic; found in superficial lesions.
IV Rapid growers	*M. fortuitum* *M. chelonei*	Usually nonpathogenic; resistant to antituberculosis drugs. May be agents of abscesses in puncture wounds; resistant to antituberculosis drugs.

Source: Pelczar, 1993.

APPENDIX VI

Treatment regimens for tuberculosis

The essential anti-TB drugs, their mode of action and recommended dose (range in parenthesis).

Essential anti-TB drug (Abbreviation)	Mode of action	Recommended dose (mg/kg) intermittent		
		Daily	3x/week	2x/week [a]
Isoniazid (H)	Bactericidal	5 (4-6)	10 (8-12)	15 (13-17)
Rifampicin (R)	Bactericidal	10 (8-12)	10 (8-12)	10 (8-12)
Pyrazinamide (Z)	Bactericidal	25 (20-30)	35 (30-40)	50 (40-60)
Streptomycin (S)	Bactericidal	15 (12-18)	15 (12-18)	15 (12-18)
Ethambutol (E)	Bacteriostatic	15 (15-20)	30 (25-35)	45 (40-50)
Thioacetazone (T)	Bacteriostatic	2.5	Not applicable	

Source: Maher *et al.*, 1997.

[a] *WHO does not generally recommend twice weekly regimens.*

Possible alternative treatment regimens for each treatment category

TB treatment category	TB patients	Alternative TB treatment regimens	
		Initial phase (daily or 3 times per week)	Continuation phase
I	New smear-positive PTB; new smear-negative PTB with extensive parenchymal involvement; new cases of severe forms of EPTB	2 EHRZ (SHRZ) 2 EHRZ (SHRZ) 2 EHRZ (SHRZ)	6 HE 4 HR 4 H_3 R_3
II	Sputum smear-positive: relapse; treatment failure; treatment after interruption.	2 SHRZE/1 HRZE 2 SHRZE/1 HRZE	5 H_3 R_3 E_3 5 HRE
III	New smear-negative PTB (other than in category I); new less severe forms of EPTB	2 HRZ 2 HRZ 2 HRZ	6 HE 4 HR 4 H_3 R_3
IV	Chronic case (still sputum- positive after supervised re-treatment)	Not applicable (Refer to WHO guidelines for use of second-line drugs in specialized centres)	

Source: Source: Maher *et al.*, 1997.

There is a standard code for TB treatment regimens. Each anti-TB drug has an abbreviation (shown in above table). A regimen consists of 2 phases. The number before a phase is the duration of that phase in months. A number in subscript (e.g. 3) after a letter is the number of doses of that drug per week. If there is no number in subscript after a letter, then treatment with that drug is daily. An alternative drug (or drugs) appears as a letter (or letters) in brackets.

APPENDIX VII

Reading and reporting of laboratory results of culture

Reading	Report
No growth	Negative
1-19 colonies	Positive (Number of colonies)
20-100 colonies	Positive (1+)
100-200 colonies	Positive (2+)
200-500 colonies (almost confluent growth)	Positive (3+)
>500 colonies (confluent growth)	Positive (4+)
Contamination	Contaminated

Source: (WHO, 1998 b)