Molecular Epidemiology of *Mycobacterium tuberculosis* in Nepal

INAUGURALDISSERTATION

zur

Erlangung der Würde eines Doktors der Philosophie
vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Bijaya Malla
aus Lalitpur, Nepal

Basel, 2013

Genehmigt von der Philosophisch-Naturwissenschaflichen Fakultät auf Antrag von

Prof. Dr. Sébastien Gagneux und Prof. Dr. Douglas B. Young .

Basel, 26 Februar 2013

Prof. Dr. Jörg Schibler

Dekan

Table of Contents

Acknowledgements

I especially thank Prof. Gerd Pluschke for accepting me at first as PhD student in his department and providing me all the necessary administrative support so that I could start my PhD in Prof. Sebastien Gagneux's Laboratory.

My sincerest thanks go to Prof. Sebastien Gagneux for offering me the opportunity to pursue a PhD in the Tuberculosis Research Unit. I thank him for guiding me through the whole PhD process and in writing my thesis. I thank him for financial support of my study as well. I also acknowledge the members of Prof. Gagneux's group (Sonia Borrell, Julia Feldmann, Mireia Coscolla, Daniela Brites, and David Stucki) for helping me during the laboratory work and for the exchange of ideas during the writing process. Importantly, I acknowledge Julia for all her help with the BSL3 laboratory work, without that this work would never have been possible.

Special thanks go to Dr. Lukas Fenner from the Institute of Social and Preventive Medicine, University of Bern for his continuous support and motivation. He is acknowledged for his help in analyzing the results and with writing papers. Most importantly, his cooperation and guidance was productive in getting some papers published in peer reviewed journals.

Dr. Bhawana Shrestha and other staffs from the German Nepal Tuberculosis Project (GENETUP) deserve special thanks. I thank her for overseeing my daily work at GENETUP during patient recruitment and collection of clinical and epidemiology data. My deepest thanks go to the Laboratory In-charge of GENETUP, Bhagwan Maharjan for discussions and suggestions during my laboratory work. I would like to thank Namemitra Shrestha, Sanukaji Tandukar, Sajana Tandukar, Sujit Maharjan, and Chandish Shrestha for their support during my field work at GENETUP.

My sincere gratitude goes to the Amt für Ausbildungbeiträge, Basel-Stadt, Switzerland for kindly providing me with a scholarship for my PhD program.

Occasionally, I thought that pursuing this PhD project is going to be too hard. My family has been supporting, encouraging me through some of my frustrating times. Thanks to my family who reminded me the life beyond work.

Basel

Research Summary

Tuberculosis (TB) is a global health problem. One reason of conducting molecular epidemiology studies is to understand the uneven distribution of the disease in different parts of the world. The global population structure of MTBC can be studied by genotyping strains from different geographic regions, which describes the evolutionary relatedness of MTBC. The choice of appropriate genetic tools is fundamental that can elucidate local as well as global spread of disease. Genetic markers like large sequence polymorphisms (LSPs) and single nucleotide polymorphisms (SNPs) have been used to construct phylogenies of MTBC lineages that are informative for understanding the global distribution of MTBC. Additionally, markers such as direct repeats can differentiate strains within smaller geographical settings or cohort of patients. Studies have shown that the lineages diversity itself could be associated with differences in the pathogenesis and epidemiology of TB. Most importantly, the emergence of drug resistance, which results mostly among treatment failures, is a serious threat to TB control programs.

Our aim was to use those markers to explore the phylogenetic diversity and distribution of MTBC in Nepal and compare it to the global phylogeography of MTBC. Furthermore, to identify the mutational hotspots conferring drug resistance. Understanding the molecular mechanisms of drug resistance will allow us to develop rapid molecular drug resistance detection tools and management of TB cases with improved and more rational drug therapies.

We used SNPs based genotyping tool for 506 *M. tuberculosis* strains from Nepal. This revealed four major lineages of MTBC. This allowed us to map the MTBC structure in Nepal compared to the global diversity. Additionally, the use of spoligotyping and MIRU-VNTR (used for XDR strains only) provided data within particular geographical settings and within human populations. A total of 69 different spoligotypes with unique SIT numbers were

identified. We found Beijing and Central Asian Strain (CAS) family as the predominant genotypes as was expected owing to geo-position of Nepal in Asia.

Molecular analysis of drug resistance for most common anti-TB drugs (i.e. isoniazid, rifampicin) from our sample set confirmed that the polymorphisms were more or less similar as previously documented globally, although we found some additional non-synonymous mutations which need validation. In general, our findings showed that the rapid molecular tools currently developed will detect most of the drug resistance isolates in Nepal. Among drug resistance strains, the *katG* S315T was proportionally more represented by multi-drug resistance strains. However, the patterns of *rpoB* mutation were unrelated to multi-drug resistance or MTBC genotypes. By performing 24 MIRU-VNTR loci plus additional 4 hyper-variable region intended to use for Beijing spoligotypes, we provide evidence of primary transmission of XDR strains.

On the other hand, the aim was to identify risk factors, risk groups, and co-morbidities that may relate to the susceptibility to TB. The number of male patients constituted two-third of the total sample population and most of them were at the age of 15-24 years. However, female TB patients in Nepal seem to be associated with “virulent” strains of TB (Beijing genotype) and drug resistance. We identified four XDR cases; the younger age (median age 21 yrs.) of XDR-TB is a serious matter that requires immediate attention from NTP, Nepal.

List of Tables

List of Figures

Abbreviations

BCG	Bacillus Calmette Guérin
CAS	Central Asian Strain
CDR	Central Development Region
CFP	Culture Filtrate Protein
CRISPR	Clustered Regulatory Short Palindromic Repeats
DOTS	Directly Observed Short Course Therapy
DRs	Direct Repeats
DST	Drug Susceptibility Test
EAI	East African Indian
EPI	Expanded Program on Immunization
ESAT	Early Secreted Antigen Type
FDC	Fixed Dose Combination
FNAC	Fine Needle Aspiration Cytology
GENETUP	German Nepal Tuberculosis Project
GIS	Geographic Information System
GLC	Green Light Committee
ICIMOD	International Centre for Integrated Mountain Development
IQR	Interquartile Range
IS	Insertion Sequences
LSP	large Sequence Polymorphisms
MDGs	Millennium Development Goals
MDR	Multi- Drug Resistance
MIRU	Mycobacterial Interspersed Repetitive Unit
MTBC	*Mycobacterium tuberculosis* complex
MTBDR*sl*	*Mycobacterium tuberculosis* drug resistance second line
NHRC	Nepal Health Research Council
NTC	National Tuberculosis Center
NTMs	Non-tuberculose Mycobacteria
NTP	National Tuberculosis Control Programme
PCR	Polymerase Chain Reaction
RD	Region of Difference
RFLP	Restriction Fragment length Polymorphisms
RRDR	Rifampicin Resistance Determining Region
SIT	Spoligotype International Type
SNPs	Single Nucleotide Polymorphisms
Swiss TPH	Swiss Tropical and Public Health Institute
TDR	Totally Drug Resistance
VNTR	Variable Number of Tandem Repeats
WHO	World Health organization
XDR	Extensively Drug Resistance

Chapter 1: Introduction

1.1 Global epidemiology of human tuberculosis[1]

The most recent global TB report of the World Health Organization (WHO) from 2012 provides the latest TB situation from more than 200 countries, covering both developing and developed countries **(World Health Organization, 2012)**. In the year 2011, 8.7 million new cases of TB and 1.4 million deaths due to TB were recorded (Figure 1). However, the estimates of incidence and death rates varied by country. Geographically, the TB burden is highest in highly populated continents. Together with India and China accounting for almost one quarter of global cases, 40% of the world's TB cases reside in Asia and in Africa,.

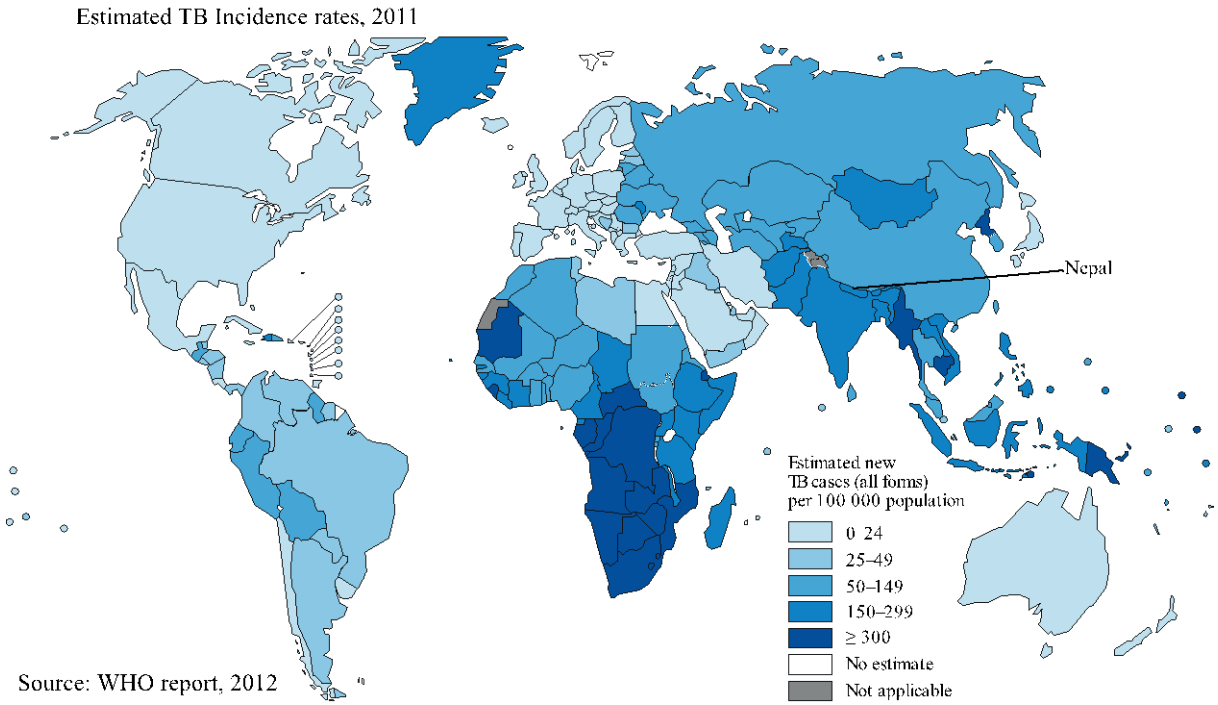

Figure 1: Estimated global incidence rates of tuberculosis (2011).

The new TB cases notified to WHO by the country-specific NTP programs were 5.8 million in 2011. The first similar publication by WHO in the year 1995, which was during the start of the Directly Observed Short Course Therapy (DOTS) strategy, notified 3.4 million new cases. This illustrates how the number of diagnosed TB cases has increased in the past 10-12

[1] Adapted from "Global Tuberculosis Report - 2012"

years. The large disparity between estimated and notified cases for the year 2011 highlights the inability of many NTP programs to correctly diagnose the actual number of TB cases, leaving many undiagnosed and thus untreated. This will lead to adverse consequences, particularly in those settings where drug resistance is on the rise.

With the declaration of TB a "global emergency" by WHO in 1993, the implementation of combined therapy also known as DOTS was successfully implemented in many parts of the world. This had an impact on lowering the prevalence of TB. Currently, the treatment success rate is 80% or more among new TB cases globally. The remaining 20-30% of cases is at increased risk of failure, with increasing drug resistance as being one of the underlying reasons. With the accessibility of drug resistance screening methods such as phenotypic drug susceptibility tests (DST) or molecular tools (despite being only slowly adopted in some countries), data on TB drug resistance have become available from many parts of the world. Worldwide, 3.7% of new cases and 20% of previously treated cases were estimated to have multi-drug resistance (MDR)-TB as recorded in 2011 (Figure 2). However, these drug resistance figures are likely far from being truly representative, as due to technical and logistic constraints, many countries cannot perform routine DST on all patients.

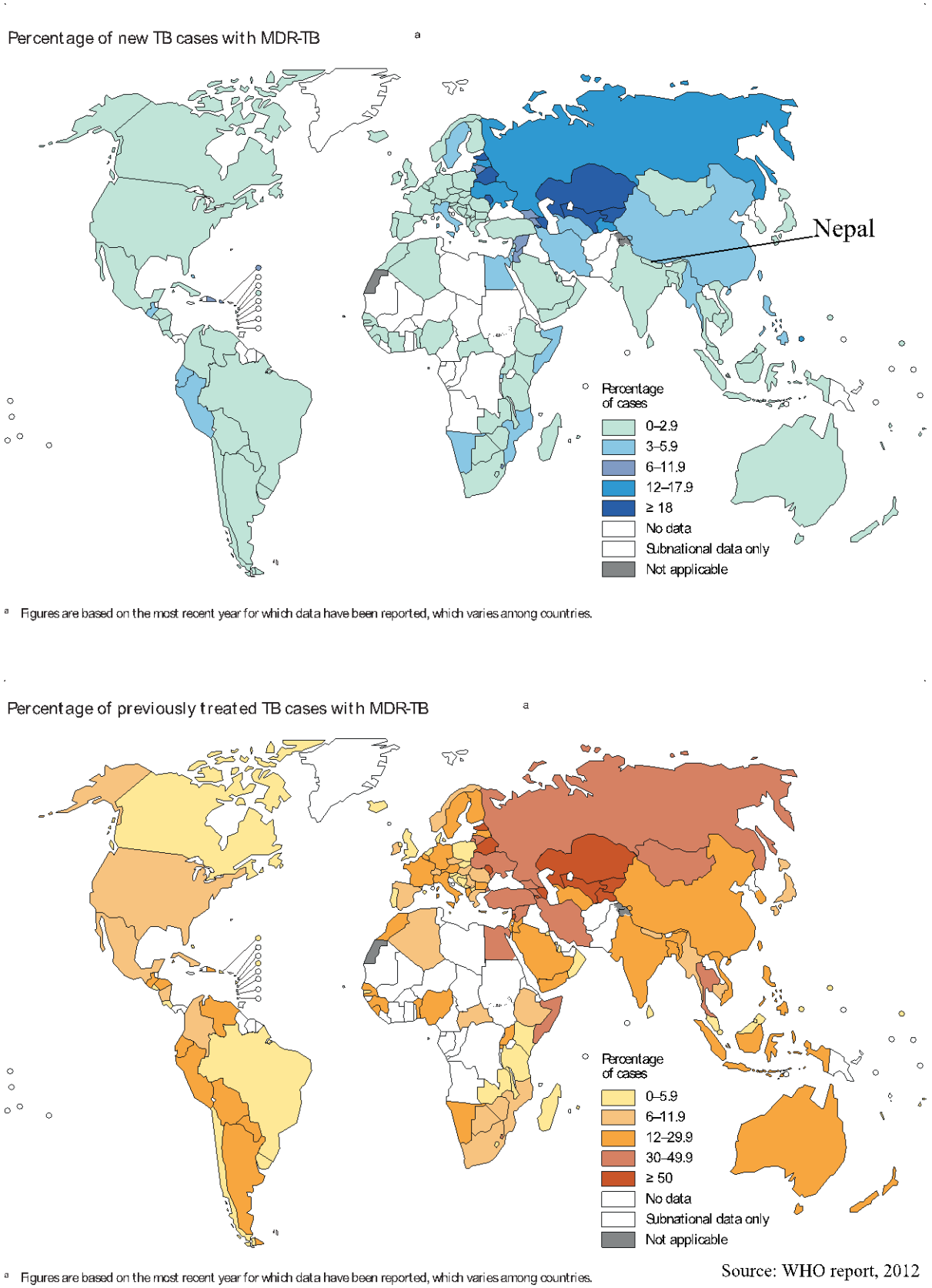

Figure 2: Percentage distribution of global MDR-TB in new and previously treated TB cases.

The global plan to Stop TB 2011-2015 includes targets that by 2015, all cases of TB should be considered as at high risk of MDR-TB, and hence should have access to standard DST. Similarly, all patients with MDR-TB should undergo DST for second-line drugs to detect potential extensively drug resistance (XDR) TB, as this form of TB is particularly difficult to treat. By 2011, XDR-TB had been reported from 84 countries (Figure 3). However, because of the limited technical and financial resources in many countries, only a small proportion of TB cases are currently tested for drug resistance.

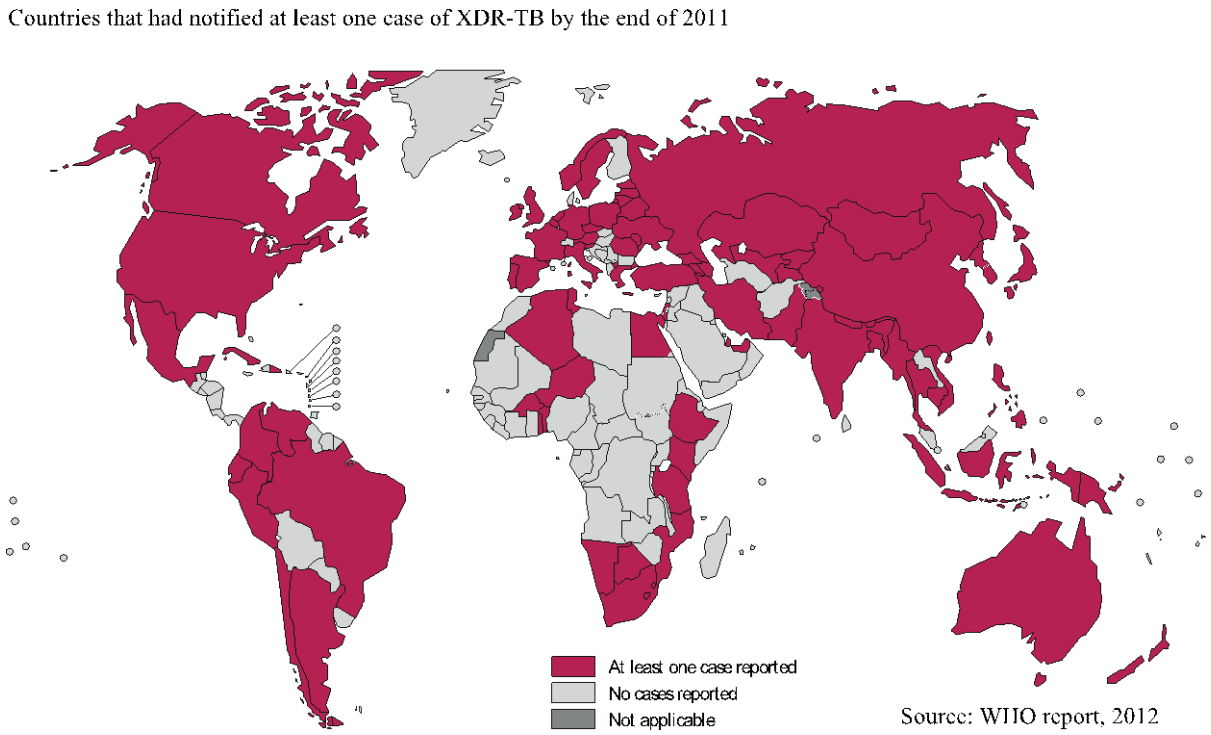

Figure 3: Countries that had notified at least one case of XDR-TB by the end of 2011

In addition to being threatened by the emergence of drug resistance, global TB control is further complicated by co-morbidities such as HIV and diabetes. These diseases lead to immunosuppression in the host and make the patient particularly susceptible to TB. In persons co-infected with HIV, the risk of developing active TB exceeds 10% per year, as opposed to 10% per a life-time in HIV-uninfected individuals. HIV/TB co-infection is one of

the most common causes of death in least developed countries **(Aaron *et al.*, 2004)**. The WHO data from the year 2011 estimated 430,000 HIV-associated TB deaths globally. In an effort to better control HIV/TB, screening programmes are being scaled up. Globally, 79% of people living with HIV were provided with co-trimoxazole preventive therapy in 2011.

To address the concerns of growing number of TB cases and drug resistance, WHO has developed the "Stop TB Strategy" that was launched in 2006. With a vision of "A TB-free world", Stop TB Strategy has set an objective linked to Target 6.c of the Millennium Development Goals (MDGs), which plans to "Halt and begin to reverse the incidence of TB by 2015" and to "reduce prevalence of and deaths due to TB by 50% compared to their levels in 1990". Another MDGs target associated to the Stop TB Strategy is to "eliminate TB as a public health problem by 2050".

Starting in 2009, several publications have led to controversies regarding the classification of a new form of TB that is resistant to all anti-TB drugs and which has been coined "totally drug resistant TB (TDR-TB)". These reports came first from Italy **(Migliori *et al.*, 2007)**, Iran **(Velayati *et al.*, 2009)**, and later also from India **(Udwadia *et al.*, 2012)**. WHO argues that such a definition of "TDR-TB" cannot be used unless proper verification and standardization of DST guidelines have been established that i) cover all anti-TB drugs, and ii) can be applied in all TB diagnostic laboratories **(World Health Organization, 2008)**. Moreover, the reproducibility and reliability of second-line DST is limited, and critical concentrations to define resistance have been found to differ in different laboratory settings. Finally, new anti-TB drugs are still being evaluated in clinical trials, and the "TDR-TB" has not yet been tested against those drugs **(World Health Organization, 2008)**.

1.2 The origin of pathogenic mycobacteria

The mycobacteria exhibit great diversity in growth and live in diverse ecological niches. Most of the ~120 mycobacterial species are saprophytes that grow and are able to replicate in soil or water. It is hypothesized that the pathogenic species of mycobacteria diverged from early ancestors that evolved in different environments, eventually developing the capacity to survive intracellularly (i.e. in free-living amoebas). Some of these mycobacteria eventually evolved to become true pathogens, depending on the host environment to survive and multiply. The three major mycobacteria species that are pathogenic to humans are known as the *Mycobacterium tuberculosis* complex (MTBC), *M. leprae* and *M. ulcerans*, and cause TB, leprosy, and Buruli ulcer, respectively. In addition, several "non-pathogenic" mycobacteria (also known as the "non-tuberculous mycobacteria (NTMs)") that are commonly found in the environment can cause opportunistic infections in humans; examples include *M. intracellulare* or *M. kansasii,* which are often associated with opportunistic infections in HIV co-infected individuals. Many of these NTM infections are difficult to treat as many environmental mycobacteria are naturally resistant to many of the drugs used to treat TB. The growth rate of these environmental mycobacteria is much faster and similar to other organisms like *Escherichia coli.* In contrast, generation time of parasitic mycobacteria ranges from ~24h in MTBC to >72h in *M. ulcerans* and 14 days in *M. leprae*.

TB is primarily caused by the different members of the MTBC. Based on 16S rRNA sequencing, MTBC forms a single and compact clonal group together with *M. canettii* and other so-called "smooth strains" (Figure 4) **(Gutierrez *et al.*, 2005)**. *M. canettii* and the other smooth strains have distinct phenotypes, forming smooth colonies with a shorter generation time compared to the other members of the MTBC which form rough colonies. These smooth strains have been proposed to represent the population of mycobacteria from which the MTBC evolved. Hence collectively, *M. canettii* and the other smooth strains have been referred to as *M. prototuberculosis* **(Gutierrez *et al.*, 2005)**. Little is known on the epidemiology of *M. prototuberculosis*, but the fact that only about 60 patient isolates have been reported to date, and almost all of them were associated with the Horn of Africa, suggests that an animal or environmental reservoir might exist **(Fabre *et al.*, 2004)**.

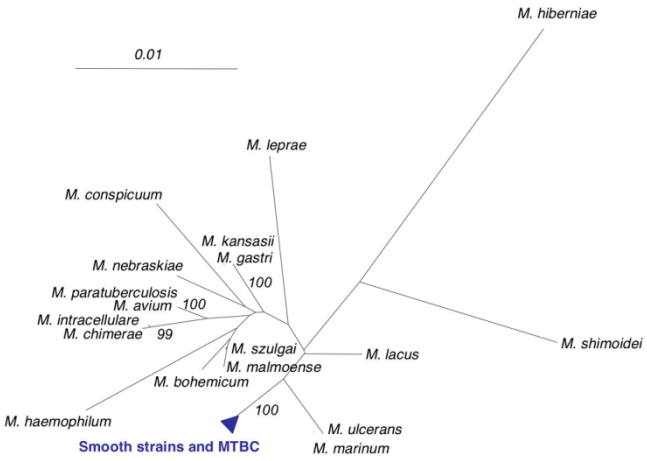

Figure 4: Phylogenetic position of tubercle bacilli within the genus mycobacterium

By contrast, the other members of the MTBC are obligate pathogens with no known environmental reservoir. The MTBC *sensu stricto* (i.e. excluding *M. canettii*) comprise several human- and animal-associated species and sub-species. Human TB is primarily caused by *M. tuberculosis sensu stricto* and *M. africanum*. In addition, several MTBC lineages are thought to be specially adapted to particular domestic and wild animal species. These include *M. bovis* (a cattle pathogen) *M. caprae* (sheep and goat), *M. pinnipedii* (seals and sea lions), *M. mungi* (mangoose), *M. orygis* (antilope) and the "dassy bacillus" (rock hyrax). One of first molecular markers used to define these different MTBC lineages were genomic deletions also known as Regions of Difference (RDs) **(Brosch *et al.*, 2002)** .

Analyses based on these RDs provided insights into the evolution of MTBC. For example, the *M. tuberculosis*-specific deleted region 1 (TbD1) is of importance as it differentiates the modern and ancient lineages based on the absence or presence of it (Figure 5).

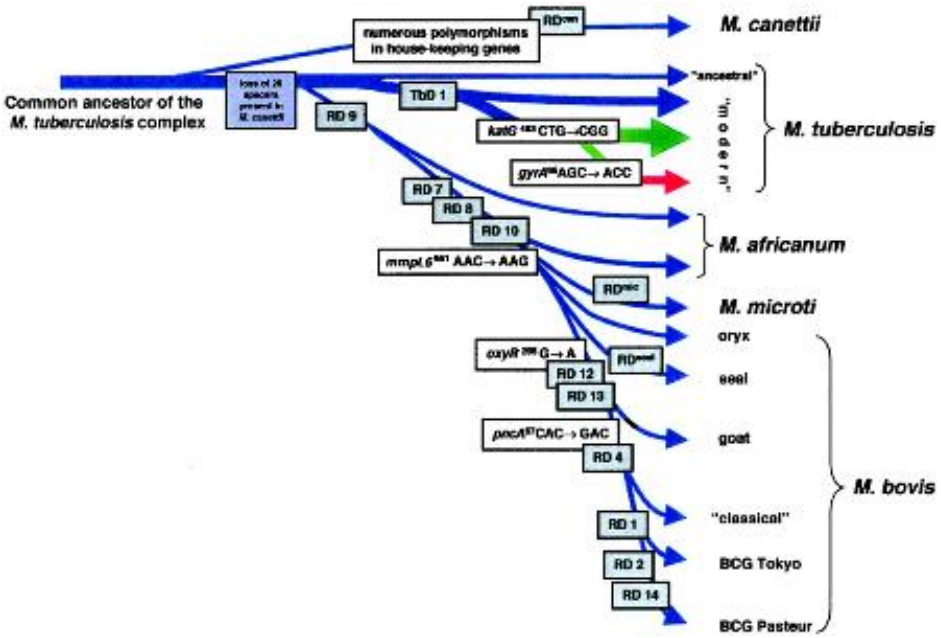

Figure 5: Proposed evolutionary pathway of tubercle bacilli based on deletions and sequence polymorphisms in five selected genes *katG, gyrA, oxyR, pncA, mmpl6* and TbD1

Among MTBC adapted to humans, the genetic variability of strains is evident both in terms of phylogenetic relationship and geography.

1.3 The global diversity of human-associated MTBC

The phylogenetic structure of human-associated MTBC has been extensively studied using LSPs in 875 globally representative strains from 80 different countries **(Gagneux *et al.*, 2006)**. This study found that human-associated MTBC consists of six major lineages which show biogeographic specificities in that the individual lineages are associated with particular geographic locations (Figure 6). Lineages 1, 5 and 6, which are referred to as the "ancient" lineages are predominant in Africa (Lineage 5 and 6) and around the Indian Ocean (Lineage 1), while the "modern lineages are more widespread but still strongly associated with

particular geographic settings; Lineage 4 in Europe, the Americas and Africa, Lineage 2 in East Asia, and Lineage 3 in South-Asia.

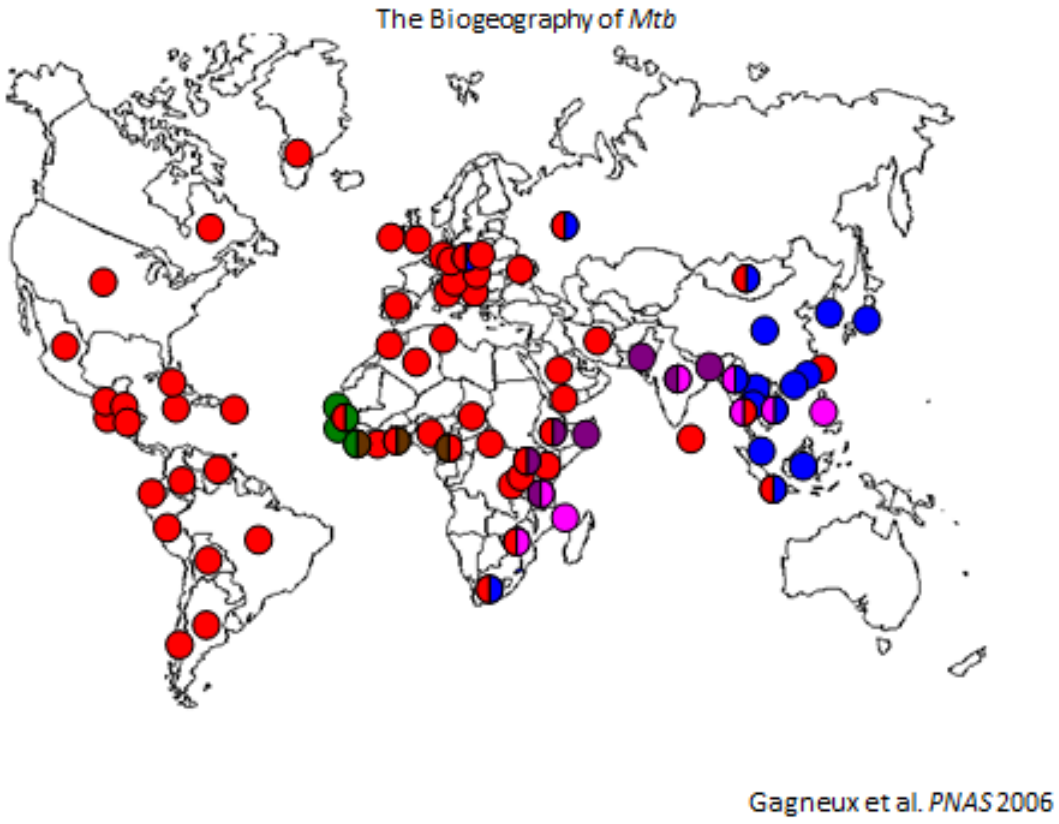

Figure 6: The biogeography of MTBC of six lineages

The same authors used another approach using maximum parsimony analysis of 89 genes in a global collection of 108 human and animal strains **(Hershberg *et al.*, 2008)**. This analysis yielded a single comprehensive phylogenetic tree (Figure 7), which showed analogy to ancient and modern lineages defined based on the presence/absence of TbD1. Analysis of genetic distances revealed that human MTBC strains are genetically diverse as represented by different phylogenetic lineages. Lineages from Africa and animal hosts are represented mostly in ancient lineage while others are represented in modern lineages

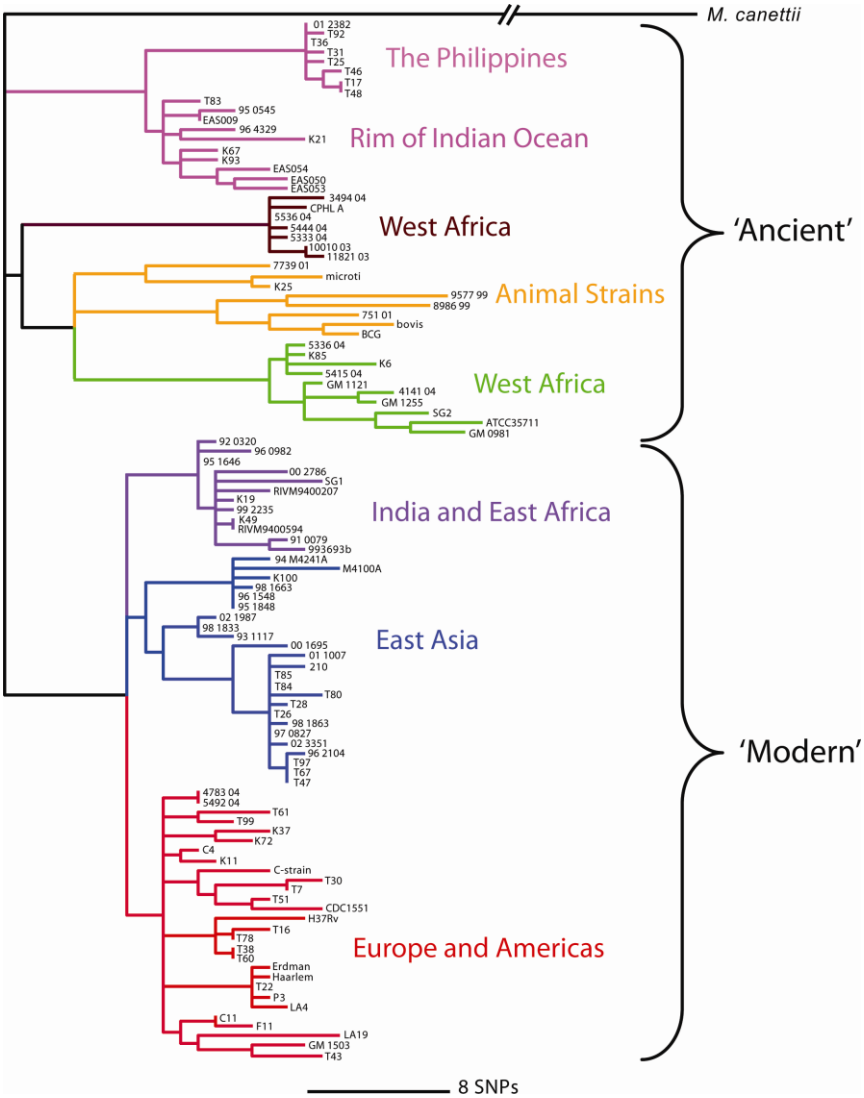

Figure 7: Phylogeny of *M. tuberculosis* showing six major lineages (Source: Hershberg *et al*, 2006)

More recently, the global phylogeny of human-associated MTBC was defined into six major lineages based on whole genome sequences **(Coscolla and Gagneux, 2010)**. This lineage classification corresponds to genotypes as detected and defined by other techniques.

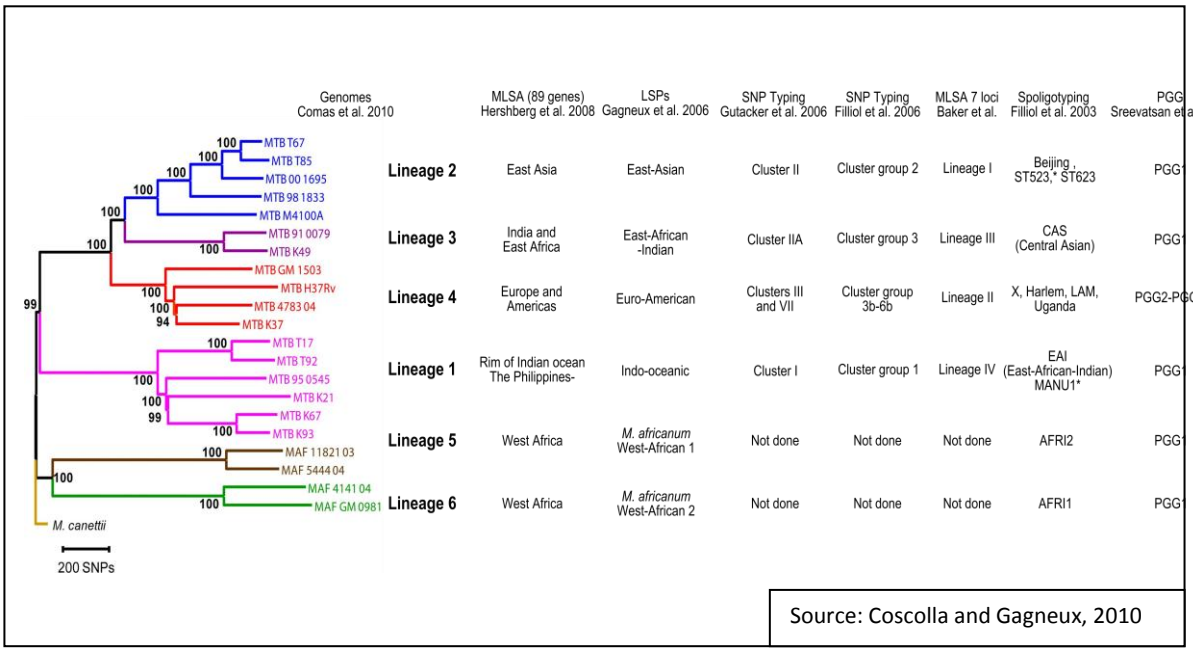

Figure 8: Phylogeny of *M. tuberculosis* based on 9037 variable common nucleotide positions

Although the explanation on genetic diversity of global MTBC continues using different genes and molecular biological tools, it is equally important to understand whether this variability has biological consequences in terms of host immune recognition, pathogenesis, and the outcome of infection and disease in clinical settings **(Portevin *et al.*, 2011,Brites and Gagneux, 2012)**.

1.4 Clinical and epidemiological consequences of MTBC diversity:

In animal models, it has been found that virulence can dependent on the genotype of the infecting MTBC strain **(Lopez *et al.*, 2003)**. With recent development in genotyping tools and use of them for exploration of local, regional, and global distribution of MTBC diversity, much attention has been given to strain variation and its association with drug resistance and disease outcome. One of the approaches is to find associations between inter-genotype characteristics within pathogen population and corresponding clinical features among TB patients. The major motivation is to provide evidence to explain if the MTBC diversity matters for global TB control **(Coscolla and Gagneux, 2010)**.

Studies in Vietnam have shown that lung cavitations were found in higher proportion in TB patients infected by Lineage 4 (i.e. the Euro-American lineage) (**Thwaites *et al.*, 2008)** and that Lineage 4 was also less likely to cause TB meningitis **(Caws *et al.*, 2008)**. The same study showed that the shorter duration of illness among TB meningitis was related to infection by Lineage 2 (East Asian/Beijing lineages) **(Thwaites *et al.*, 2008)**. However, a study in Netherlands where TB cases and controls were stratified by age, previous episode of TB and ethnicity, showed that the bacterial genotypes were not associated with chest radiological presentation **(Borgdorff *et al.*, 2004)**. A study among HIV negative TB patients and contacts in Gambia showed that the progression of disease was less likely among patients infected with *M. africanum* compared to *M. tuberculosis* **(de Jong *et al.*, 2008)**. These differing findings on links between bacterial genotypes and disease phenotypes could also be due to factor like sample size, stage of disease, geographical differences, and patient ethnicity. On the other hand, clinical presentation and disease outcomes could also be due to different treatment strategies, immunization, patient's predisposition such as HIV, diabetes, ethnicity, age which may be coupled with immigration history, past TB outbreaks, substance abuse, place of patient's origin, homelessness, and year of first episode of TB **(Dye and Williams, 2010,Dalton *et al.*, 2012)**. A prospective study from South Africa showed an association of late sputum smear conversion among TB patients who smoked and who were infected with W-Beijing genotype **(Visser *et al.*, 2012)**. If the MTBC genotype affects the formation of lung cavities, then the degree of lung cavitation will have an effect on the grading of the sputum smear. Patients with larger lung cavitations tend to be positive for maximum smear grade due to a higher bacterial load. Hence, severity of disease could also be correlated with MTBC lineages.

In another study conducted in India, old patients were more frequently associated with the East African Indian (EAI) spoligotype as compared the Central Asian Strains (CAS)

spoligotype (which correspond to Lineage 1 and Lineage 3, respectively) **(Arora *et al.*, 2009)**. This may suggest that the predominance of one MTBC genotype may have biological advantages in specific host populations, and such findings may help our understanding of the host and environmental risk factors as well as the pathogen characteristics determining the outcome of TB.

1.5 Overview of genotyping tools used for MTBC

The whole genome sequencing of *M. tuberculosis* H37Rv in 1998 paved the way for a better understanding of the biology of the bacilli **(Cole *et al.*, 1998)**. MTBC is a genetically monomorphic bacterium that has little sequence diversity compared to most other bacteria **(Achtman, 2008)**. After sequencing of MTBC strains collected worldwide, progress has been made in finding other discriminatory markers that can be used to trace the evolutionary patterns of spread of the MTBC worldwide, country wise, and locally **(Mathema *et al.*, 2006)**.

Genome sequencing has revealed polymorphic regions at nucleotide and gene level. The former also referred as single nucleotide polymorphisms (SNPs) are ideal for use as genetic markers for inferring deep phylogenies of *M. tuberculosis* **(Filliol *et al.*, 2006)**. Moreover, compared to repetitive elements, SNP-based analysis is less prone to distortion due to homoplasy (i.e. the emergence of convergent DNA fingerprints in unrelated strains) **(Schork *et al.*, 2000)**. Furthermore, horizontal gene transfer is presumed rare in MTBC, so deletions of large sequences and polymorphisms at nucleotide level (that are used as SNP based markers) are unidirectional, and evolutionary history can be determined. These changes can be purposively used to construct phylogenetic linkages/trees that represent the evolution and global spread of MTBC. However, LSP- and SNP-based typing methods have a relatively low discriminatory power and cannot be used to infer ongoing transmission or identifying outbreaks **(Mathema *et al.*, 2008)**. Hence, polymorphic markers which have a faster

molecular clock are generally used to trace the transmission of disease at the community level **(van Embden *et al.*, 1993)**. The insertion sequence IS*6110* shows polymorphic patterns in MTBC as it is present in multiple copies in the genome. This has made it possible to use this insertion sequence as a molecular marker to trace the ongoing transmission of MTBC during molecular epidemiological investigations. The classical method of studying TB epidemiology is restriction fragment length polymorphisms (RFLP) of IS sequences. Two epidemiologically linked strains would have identical or almost identical IS*6110* fingerprinting patterns. Although IS elements have a high discriminatory power in differentiating epidemiologically unrelated strains, this method requires high DNA quantity and digestion of DNA with restriction enzymes (*PvuII*) for southern blotting **(Millan-Lou *et al.*, 2012)**.

Other methods such as spoligotyping and MIRU-VNTR typing present the better technique for molecular epidemiological studies as their discriminatory power to trace ongoing transmission as similar to IS*6110* RFLP, when used in combination. Spoligotyping has been widely used in high incidence areas, and where infection and disease patterns are heterogeneous. Spoligotyping is based on the polymorphisms in the direct repeat region of MTBC. In MTBC, this region consists of multiple 36-bp direct repeats (DRs) interspersed by unique spacer DNA sequences (35 to 41 bp) **(van Embden *et al.*, 2000)**. This genotyping method is based on the evaluation of the presence or absence of 43 spacer DNA sequences between the 36 bp direct repeats in the DR region of MTBC strains. These multiple copies of 36-bp direct repeats are well conserved but the spacer sequences between those DR sequences are different. The lack of certain spacers is helpful for genotyping MTBC strains. In Netherlands, outbreak strains from epidemiologically related cases were evaluated by spoligotyping; the hybridization was identical for all except for one which was different in one spacer. All those strains had similar IS*6110* pattern. Similarly, outbreak strains from UK

which were epidemiological linked have been shown to harbor identical hybridization patterns. The conventional contact tracing suggested that patients with similar IS*6110* patterns were from the same chain of TB transmission. This suggested that these genotyping tools can be used as surrogate markers for disease transmission.

Mycobacterial interspersed repetitive units (MIRUs) in DNA elements are tandem repeats and are dispersed throughout the genome **(Allix-Beguec *et al.*, 2008)**. The MIRU-VNTR genotyping method evaluates the number of tandemly repeated sequences at different loci and whether the number of copies of the repeated sequence varies among strains. PCR amplification and comparison of the product sizes of those repeated sequences with a molecular size marker on an agarose gel is normally sufficient to find the number of copies present. The *Mycobacterium tuberculosis* H37Rv reference strain contains 41 MIRU loci, of which many are polymorphic and currently used in standard MIRU-VNTR typing **(Supply *et al.*, 2000)** . The discriminatory power of MIRU-VNTR analysis is typically proportional to the number of loci evaluated; in general, when only 12 loci are used, it is less discriminating relative to using 15 or 24 MIRU loci **(Kremer *et al.*, 2005)** .

1.6 Diagnosis, treatment and vaccination in TB

The diagnosis of TB is done by detecting MTBC from the samples collected from specific infection sites. The most common form of TB is the pulmonary TB, which is diagnosed by sputum collected from a patient with an abnormal chest x-ray. However, some active TB patients show normal X-ray, especially during early times of infection, or when harboring a low bacterial load. For extra-pulmonary TB, a biopsy sample from the infection site is collected. For example, Fine Needle Aspiration Cytology (FNAC) can be performed from suspected TB patients with enlarged lymph nodes. Smear microscopy is the most commonly used technique for the diagnosis of TB which is still considered as the “gold standard”

techniques in most developing countries. The specificity and sensitivity of microscopy is less than culture, so despite collecting three sputum samples (as NTP guidelines) from suspected TB cases, there is a high possibility that some cases will remain undetected **(Dye *et al.*, 2003)**.

In developed countries, bacterial culture is considered the gold standard for TB diagnosis, However, the conventional culture technique needs several weeks or months before a diagnosis (for e.g. for drug resistance) can be made. Hence, several rapid diagnostic tests based on molecular markers such as Xpert MTB/RIF **(Kurbatova *et al.*, 2012)** and GenoType MTBDR*sl* assay have been developed **(Ling *et al.*, 2008)**. These rapid molecular based tools are a step ahead in providing the point-of-care test needed to efficiently control TB globally. Early case detection and efficient use of those rapid tools depend on training of health staff, accessibility, and affordability by the NTP programs, which are still important challenges for many developing nations **(Ling *et al.*, 2008)**. The policy implementation of those rapid tools for the diagnosis of drug-resistant TB is yet another and particular challenge.

After introduction of anti-TB drugs as a combined therapy, much progress has been made in the control of TB **(World Health Organization, 2008)**. Rifampicin, isoniazid, pyrazinamide, and ethambutol are the most important first-line drugs used under DOTS. Resistance against these drugs, especially rifampicin and isoniazid pose a serious global threat. The genes involved in resistance of isoniazid are *katG* and *inhA*. Although there are many mutational changes characterized in those genes, Ser315Thr amino acid replacement in *katG* has been the most common mutation, occurring in about 50-93% of resistant clinical isolates. Additionally, molecular analysis to define rifampicin resistance suggests that mutations in an 81 bp drug resistance determining region (RRDR) of *rpoB* lead to different levels of resistance to rifampicin where mutations in positions 526 (H/D), 516 (D/V) and 531 (S/L) are

most common and seen worldwide. With the increase in MDR-TB cases globally, second-line drugs such as floroquinolones and aminoglycosides have been used frequently. Floroquinolones target the DNA gyrase involved in DNA replication, which is encoded by *gyrA* and *gyrB.* Point mutations at codon 90, 91, 94, and 95 of *gyrA* (a region known as "quinolones resistance determining region" (QRDR)) are associated with drug resistance. Another class of second-line anti-TB drugs, aminoglycosides, inhibits protein synthesis. Mutation from A to G at *rrs* gene position 1400 is associated with resistance to commonly used kanamycin and amikacin (Table 1). The rapid tools such as Xpert MTB/RIF exploits those mutations in RRDR region to verify drug resistance in TB and are used as a proxy for the multi-drug resistance. However, contrasting results were shown that strains with rifampicin resistance may or may not be isoniazid resistance depending on geography and treatment protocol of the local TB Control Program **(Smith *et al.*, 2012)**.

Table 1: Anti tuberculosis drugs and the gene(s) involved in drug resistance

Essential anti-TB agents (abbreviation)	Mode of action, Potency	Product (Genes involved in drug resistance)
Isoniazid (H)	Bactericidal, high	Enoyl acp reductase *(inhA)* Catalase-peroxidase *(KatG)* Alkyl hydroperoxide reductase *(ahpC)* Oxidative stress regulator *(oxyR)*
Rifampicin (R)	Bactericidal, high	RNA polymerase subunit B *(rpoB)*
Pyrazinamide (Z)	Bactericidal, low	Pyrazinamidase *(pncA)*
Streptomycin (S)	Bactericidal, low	Ribosomal protein subunit 12 *(rpsL)* 16s ribosomal RNA *(rrs)* Aminoglycoside phosphotransferase gene *(strA)*
Ethambutol (E)	Bactericidal, low	Arabinosyl transferase *(emb A, B and C)*
Fluoroquinolone	Bactericidal, low	DNA *gyrA*se A subunit and B subunit *(gyr A and B)*

In recent years, additional mutations in these resistance genes have been documented. A study by Gagneux *et al.* has showed that drug resistance in MTBC is associated with a competitive fitness cost in absence of the drug, which varies depending on the specific

resistance mutation **(Gagneux *et al.*, 2006)**. Moreover, compensatory mechanisms are able to mitigate the negative fitness effects of resistance mutations **(Comas *et al.*, 2012)**. Because of the fitness cost of drug resistance, it has been hypothesized that drug-resistant strains are less likely to transmit and result in disease. So comparing the size of molecular clusters of drug resistance versus sensitive strains may help predict transmission of drug-resistant strains compared to drug-susceptible strains. Hence, introduction of genotyping methods to drug susceptibility testing can assist drug resistance surveillance. For example, Beijing strains which are prevalent in Asia **(Li *et al.*, 2005)** have often been associated with drug resistance **(Borrell and Gagneux, 2009)**. Hence, we could expect that where many Beijing strains circulate, drug resistance, treatment failure, and relapses might be a particular problem.

The BCG vaccine offers unreliable protection against TB in adults and provides limited protection among children above 2 years of age. However, because BCG protects small children against TB meningitis, BCG continues to be one of the major constituent of the routine Expanded Program on Immunization (EPI) **(World Health Organization, 1999)**. One of the mechanisms of attenuation during the development of BCG vaccine from virulent *M. bovis* was due to the deletion of RD1 region. RD1 comprises 9 genes, including early secreted antigen type 6 (ESAT-6) and culture filtrate protein 10 (CFP-10) **(Young, 2003)**.

1.7 The Tuberculosis Control Programme in Nepal (TB situation in Nepal)[2]

Nepal is one of the least developed countries in the world, situated between the two high TB burden countries, India and China. The DOTS strategy in Nepal was adopted in 1996 as pilot project in four districts, and by 2001, it was extended to all 75 districts with an institutional nationwide coverage. By 2011, 1118 treatment centers and 3103 sub-treatment centers were offering DOTS treatment, which is complemented by many private health facilities. About 45% of the total population of Nepal is considered latently infected with MTBC, resulting in

[2] This section is adapted from "Annual Report 2010/2011- Nepal National TB Programme."

49,000 new active TB cases each year, corresponding to an annual incidence rate of 163/100,000 population. The annual death rate (deaths/100,000 population/year) was estimated at 21/100,000 in 2010, which is down from 51/100,000 before 1990 (i.e. during the pre-DOTS era). Similarly, the treatment success rate for the year 2009/10 was 90%, exceeding the global target of 85%. Since the adoption of the DOTS strategy in 1996, the DOTS population coverage has reached 100% since 2004, and the case detection rate by sputum microscopy is more than 70% **(National Tuberculosis Programme, 2011)**. Therefore, the TB control program in Nepal as a whole has been quite successful so far.

The National Tuberculosis Programme (NTP) adapts its goal, policies, and guidelines according to the international health organizations such as WHO and the Stop TB partnership, and is integrated into the Ministry of Health of Nepal. The National Tuberculosis Center (NTC) is the central governing body of NTP. The NTP activities at regional level (5 developmental regions) and district level (75 districts) are planned and coordinated by NTP. The DOTS centers and microscopy centers at the level of Health Post and Sub-health Post act as primary health care institutions and are governed by District Public Health Office at district level.

Sputum microscopy is still the gold standard method for the diagnosis of TB in Nepal. This is a free diagnostic service for patients registered in the NTP program. At present, 407 microscopy centers are providing diagnosis service under the direct NTP laboratory network coverage, while an additional 98 microscopy centers are operated through private partners. The culture and drug resistance testing services are provided by the NTC and German Nepal Tuberculosis Project (GENETUP) laboratories. Recently, the government of Nepal has planned to extend culture services in five regional hospitals. The national quality control of these laboratories is done by NTC and GENETUP, while Supranational Reference

Laboratory "Kuratorium tuberkulose in der Welt, Germany" issues the quality assurance for NTC and GENETUP.

The fixed-dose combination (FDC) for treatment of TB is a free treatment service for patients registered under NTP in Nepal. The treatment used to be eight months (2HRZE/6HE); however the standard WHO six months regimen was introduced in Nepal in 2009. The ambulatory DOTS program has been successfully implemented in selected DOTS center in Nepal **(Malla *et al.*, 2009)**. Treatment of TB is successful where early diagnosis and prompt treatment is in place. Supervision of the treatment and monitoring of possible side effects is essential for the cure of the TB patient. All the TB cases are categorized into specific disease category prior start of treatment for the homogeneity of classification of disease and to provide the standard treatment (Table 2). This also contributes to a standardized recording system.

Table 2: TB Patient Registration Category

Disease Category	Definition of Case
New	A patient who has received no or less than 28 days of anti-tuberculosis treatment. This also applies to primary resistant DR-TB confirmed after DST.
Relapse	A previously treated case whose most recent treatment outcome was "cured" or "treatment completed", and who is subsequently diagnosed with bacteriologically positive either by microscopy or culture.
Treatment after default	A patient whose previous DOTS treatment was interrupted for two or more consecutive months, and returned for treatment. The patient is bacteriologically positive either by microscopy or culture.
Treatment after failure Category I	A patient under category I treatment but is still sputum smear positive at five months or later during treatment.
Treatment after failure Category II	A patient under category II treatment but is still sputum smear positive at five months or later during treatment.
Transfer in /out	A patient who has transferred in from one DOTS treatment center to another DOTS center to continue treatment
Other	These are types of patients who may not fi t into any of the above categories. Examples include the following: sputum smear-positive patients with unknown previous treatment outcome; sputum smear-positive patients who received treatment other than Category I or II (possibly in the private sector); patients who have received several unsuccessful treatments, were considered incurable by health staff and who have lived with active TB disease with no or inadequate treatment (so-called "chronic" patients).

The treatment regimen is based on the type of the patients as categorized following the WHO standard guidelines which is outlined in Table 3.

Table 3: NTP Treatment Regimens

<table>
<tr><td>Tuberculosis Category</td><td colspan="2">Treatment Regimen</td><td>Type of patients</td></tr>
<tr><td>I</td><td colspan="2">2(HRZE)/4(HR)</td><td>New sputum smear-positive
Suggestive of TB although Sputum Negative</td></tr>
<tr><td>II</td><td colspan="2">2S(HRZE)/1(HRZE)/5(HRE)</td><td>Re-treatment TB cases including failures, relapse and return after default</td></tr>
<tr><td rowspan="3">MDR</td><td colspan="2">8(Km-Z-Lfx-Eto-Cs)/ 12(Lfx-Eto-Cs-Z)</td><td>Multi-Drug resistant Cases</td></tr>
<tr><td>Intensive Phase (8-12 months)</td><td>Continuation Phase (12 months)</td><td rowspan="2"></td></tr>
<tr><td>Kanamycin (KM)
Pyrazinamide (Z)
Levofloxacin (Ofx)
Ethionamide (Eto)
Cycloserine (Cs)</td><td>Pyrazinamide (Z)
Levofl oxacin (Ofx)
Ethionamide (Eto)
Cycloserine (Cs)</td></tr>
<tr><td rowspan="3">XDR</td><td colspan="2">Based on disease prognosis and response to anti-TB drugs and side effects</td><td>MDR cases with resistant to floroquinolone and at least one injectable.</td></tr>
<tr><td>Intensive Phase (12 months)</td><td>Continuation Phase (12 months)</td><td rowspan="2"></td></tr>
<tr><td>Capreomycin (CM)
Moxifloxacin (Mfx)
PAS
Cycloserine (Cs)
Amx/Clv
Clofazimine
Any other drug thought susceptible</td><td>Moxifloxacin (Mfx)
PAS
Cycloserine (Cs)
Amx/Clv
Clofazimine
Any other drug thought susceptible</td></tr>
</table>

A standard drug-resistant TB management programme was implemented in Nepal in 2005, after the WHO Green Light Committee (GLC) gave approval and with technical support from WHO. NTP provides fully supervised standard regimen for the treatment of MDR-TB from 12 treatment centers and 62 Sub Treatment Centers spread nationwide.

The NTP national strategic plan for the year 2010-2015 aims to detect 82% of infectious TB cases and maintain the treatment success rate at 90%. From the NTP program perspective, despite progress in case detection rate and DOTS coverage rates, the main challenges are the sustainability of the programme, which is largely dependent on foreign donors. Lack of technical expertise in surveillance of drug resistance, and for strengthening of the reference - and the regional laboratories, are key challenges. One of the objectives of the Stop TB Strategy is to "contribute to health system strengthening" by "adaptation of innovations". We believe that research on TB not limited to control activities, but covering other fields such as epidemiology, strain diversity, and drug resistance, will contribute to the control of TB in Nepal and worldwide.

Chapter 2: Rationale, Goals and Objectives

2.1 Rationale

In the developed world, molecular epidemiological studies of *M. tuberculosis* are performed in order to understand the dynamics of transmission among and local and migrant population. These findings have proved to be valuable and effective in TB control in the respective countries. Similar work from developing countries is limited despite of the endemicity of TB. In Nepal, there is currently no data about the phylogeographic distribution of MTBC. Exploring the lineage diversity of MTBC strains in Nepal is relevant as it will provide the evolutionary linkage between strains circulating in Nepal to the neighboring countries and globally. The relevance is specifically vital as the neighboring countries, India and China, are two high TB burden countries. Furthermore, the threat of drug resistance and molecular mechanisms behind emergence of such forms of disease are important for the development and effective use of new molecular diagnostic tools. In Nepal, the results of molecular epidemiological studies could assist in recommending novel disease control strategies. This will further promote research in explaining risk groups and risk factors, which is a prerequisite for an effective control program. Finally, the combination of demographic and clinical data with strain diversity data and drug resistance can provide a better picture of the evolution and transmission of TB in Nepal.

2.2 Goals

To contribute to the understanding of the phylogeography and molecular epidemiology of *Mycobacterium tuberculosis* in Nepal.

2.3 Objectives

The objectives of this research are categorized into the main and exploratory objectives, and are explained in the respective sections.

Objective 1 - To define the MTBC diversity in Nepal compared to the world

Rationale: Various studies have shown that the geographic origin of human is suggestive of the MTBC lineage. MTBC consists of 6 main lineages **(Comas *et al.*, 2009)**, and SNPs can be used as an assay to define lineages from previously unexplored geography **(Stucki *et al.*, 2012)**.

General Approach: Appropriate genotyping tools and reference information are prerequisite for performing molecular epidemiological analysis. For *M. tuberculosis*, genotyping tools based on direct repeats (DR), single nucleotide polymorphisms (SNPs) are well established, however the extent of appropriateness may slightly vary depending on geography and host factors. Taking genotype information from neighboring countries, SNPs specific to four different lineages were evaluated in samples from Nepal. For SNP typing, TaqMan and Luminex genotyping assays were used as previously described **(Stucki *et al.*, 2012)**. These methods are described in General Materials and Methods section.

In order to have good representation of the geography, MTBC strains from patients representing different regions of Nepal were enrolled in the study. The demography and clinical characteristics of the cases were diverse.

Objective 2 - To determine the distribution of drug resistance mutations and association with MTBC lineages

Rationale: Our understanding of the genetic changes conferring drug resistance and the underlying mechanisms has advanced rapidly **(World Health Organization, 2008,Ramaswamy and Musser, 1998)**. It has been observed that the major drug resistance conferring mutations are the same worldwide. Although Nepal is a TB-endemic country, limited information regarding molecular characteristics of drug resistance in TB is available. Knowledge of the molecular mechanisms of resistance also assists in the design of rapid diagnostics for detecting drug resistance. As newly developed rapid drug resistance detection kits are becoming available, a thorough understanding of the mutational sites and frequencies of mutation is critical for effective treatment and case management. Moreover, results from recent studies have shown that the drug resistance mutation and patterns are related to strain diversity among MTBC lineages **(Fenner *et al.*, 2012,Koser *et al.*, 2012)**. We assume that investigating this relation in this part of world will provide evidences to critically validate the hypothesis and help predict the susceptibility to drug resistance among MTBC lineages.

General Approach: We studied anti-TB drug resistance in new and previously treated cases. We used PCR and direct sequencing to analyze drug target genes for rifampicin, isoniazid, fluoroquinolones, and aminoglycosides including *rpoB, katG, inhA, rrs,* and *gyrA*. Phenotypic drug susceptibility tests were performed at GENETUP for first line drugs and second line drugs for selected strains. The mutations and polymorphisms in drug resistance genes were compared to global database as hosted in www.tbdreamdb.com **(Sandgren *et al.*, 2009)**.

Objective 3 – To assess the clinical and demographic characteristics of TB patients in Nepal

Rationale: The clinical and demographic characteristics of each TB patient provide critical information about TB epidemiology at a given time. The data mirror the changing trends of the disease as well as the state of TB control program. Data regarding clinical manifestations like onset of signs and symptoms relating to the disease, localized or disseminated TB, response to treatment, sputum conversion varies greatly among TB patients. Similarly, demographic variables like age, ethnicity, and patient's place of origin help understand the determinants of the disease. This descriptive epidemiological information can aid in identifying people at risk of disease, risk factors and prioritizing control programs.

General Approach A structured questionnaire was developed to collect patient variables including risk factors, clinical features, radiological presentation, and disease severity. All demographic, clinical, and epidemiological data were double-entered into a customized project database prepared in Microsoft© Access, (Copyright Microsoft© Corporation). We conducted univariate, and multivariate analyses of clinical factors to explore possible associations between patient and MTBC genotyping data.

Objective 4 - To seek association between MTBC lineages (Objectives 1) and clinical characteristics of TB patients (Objectives 3)

Rationale: Genotypically distinct pathogens have different degrees of fitness and virulence and clinical outcomes **(Visser *et al.*, 2012)**. TB patients, too, differ in terms of exposure to risk factors, place of origin, diabetes, HIV, and vaccination. A better understanding of the possible correlation between MTBC genotypes and host characteristics may identify factors predictive of ongoing transmission and shed light onto the biology of TB.

General Approach: A univariate and multivariate regression statistical approach was used to find associations between patient characteristics and MTBC lineage. We used Chi-square test to test the statistical significance of differences between groups in binary variables, and the Kruskal Wallis rank test for continuous variables. Logistic regression models were used for statistical analysis to compare patient characteristics associated with Lineages, adjusted for age, sex, treatment history, BCG vaccination status, and any drug resistance. The *p*-value less than 0.05 were considered significant. All statistical tests were performed in STATA 10.1 (Stata Corp., College Station, TX, USA).

Objective 5 - To use molecular typing tools to study the transmission of extensively drug-resistant tuberculosis

Rationale: By mid-2011, twenty-seven XDR-TB cases have been documented in Nepal. XDR-TB poses serious challenges for public health and clinical management. The emergence and transmission of XDR-TB is little known because the treatment history and case contacts are considered as the major risk factors. Moreover, the examination of second line DST preferentially among MDR failure cases has limited the case notification. It is yet unclear if the XDR-TB cases are emerging independent of each other or are transmitted from another case.

General Approach: First, to identify XDR-TB cases, we performed *rpoB* and *gyrA* DNA sequencing of all the culture positive cases. Those found resistant were then checked for *rrs* gene mutation to confirm XDR-TB. To summarize the epidemiological, clinical characteristics, and clustering of XDR-TB cases, we performed distinctive genotyping tools. Additional to SNP typing and spoligotyping methods which alone cannot predict the transmission chain due to their low discriminatory power, additional markers such as MIRU-VNTR were used.

Chapter 3: General Materials and Methods

3.1 Study settings

Nepal is a small country with 147,181 square kilometers in size populated by 26,494,504 people (www.cbs.gov.np). The country has five development regions and seventy-five districts. The Central Development Region (CDR) includes the capital city Kathmandu and is the most densely populated region with internal migration of population from other regions of the country. It also has the highest number of TB cases **(National Tuberculosis Programme, 2011)**.

We conducted a prospective, clinic based study over a three year period from 2009 (Aug-Dec), 2010 (Aug-Dec) to 2011 (Mar-Jul) in a TB reference laboratory, known as the German Nepal Tuberculosis Project (GENETUP) in Kathmandu, Nepal. GENETUP was established in 1987 with the objective of providing standard TB diagnosis and DOTS service to patients. The laboratory is certified by the Supranational Reference Laboratory "Kuratorium Tuberkulose in der Welt e.V." in Gauting, Germany. It is also a GLC and WHO-approved, treatment programme site in Nepal that provides standardized, second-line drug therapy for MDR-TB cases. The primary culture and phenotypic drug susceptibility tests for the first-line and second-line drugs are performed in collaboration with NTP. GENETUP has been involved in national surveillance of MDR and XDR among the DOTS registered patients in Nepal. Additionally, GENETUP is also the tertiary health institute, and patients suspected of drug resistance are referred for diagnosis from other microscopy centers which spread throughout the country. These centers exist under government programs or I/NGOs. GENETUB has recently introduced the molecular line probe assay for rapid diagnosis of MDR-TB. GENETUP also offers ambulatory DOTS and DOTS-plus treatment. The latest data from July 2010- July 2011 showed that there were 3568 outpatient visits by suspected TB cases, 1972 follow-up visits, and 671 newly diagnosed cases **(National Tuberculosis Programme, 2011)**. There are other private microscopy centers in Kathmandu where TB

patients may seek health advice and diagnosis, and some are referred to GENETUP for confirmation. Hence, our sampling is not population-based but represents a convenience sampling of patients visiting GENETUP.

3.2 Study population

The pulmonary TB suspects who reported with symptoms of TB such as cough for more than two weeks, hemoptysis, chest pain, night sweat and fever were subjected to microscopy. Patients already undergoing DOTS therapy were also enrolled during their follow-up visits and sputum samples were collected from them. After informed consent, we enrolled 650 sputum smear positive cases that visited GENETUP between August 2009 and June 2011. These patients included new cases as well as patients referred from other microscopy centers. A schematic view of the study processes is shown below Figure 9.

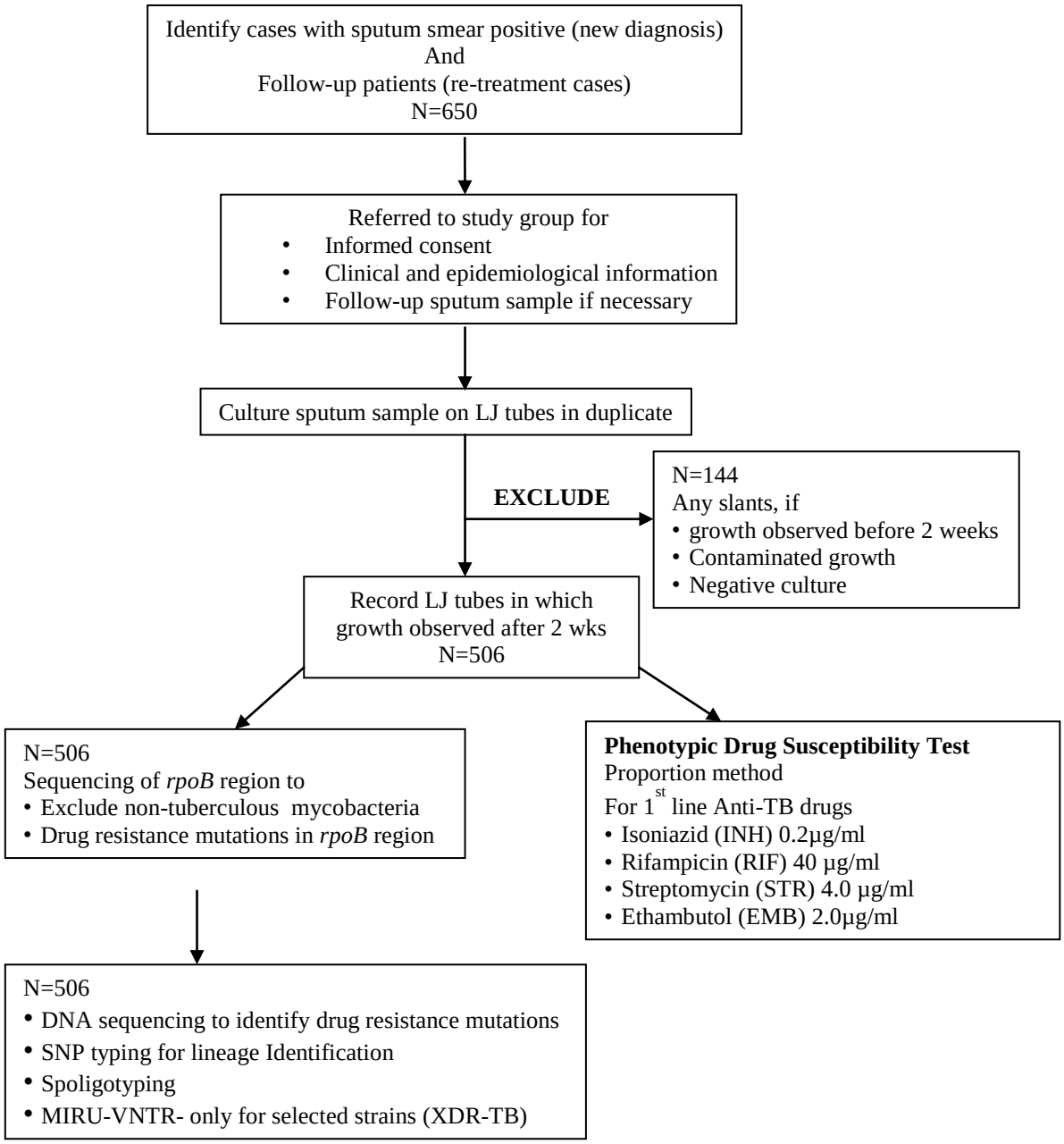

Figure 9: Summary of Study Design

We collected socio-demographic and clinical data including previous TB episodes, treatment history, HIV, and BCG vaccination status. The information was collected by physicians and trained medical and nursing staff. A new case of TB was defined as a patient who had taken anti-TB drugs for less than 28 days according to WHO guidelines **(World Health Organization, 2008**, **2008)**. A previously treated case was defined as a patient who received TB treatment for one month or more. BCG vaccination status was defined based on the presence or absence of a BCG scar. The data was double-entered in a customized and password protected Microsoft® Access database.

Sputum samples: Altogether, 3 sputum samples were collected from each patient. One was collected on the first day of visiting the GENETUP and the other two sputum samples were collected at intervals of 2 hours and on the next morning. A bacteriologic diagnosis of TB was done by microscopy using the fluorescence dye AuramineO. The sputum smear result was determined based on the WHO grading system for fluorescence microscopy.

For culture, sputum samples were decontaminated using N-acetyl-L-cysteine sodium hydroxide. The Sputum was mixed with twice of its volume with 4% NaOH in a graduated 15 ml centrifuge tube and shaken several times to digest, then left for 15 minutes at room temperature. The specimen was then centrifuged at 3000g for 15 minutes and the sediment suspended with 15ml distilled water. The tube was again centrifuged at 3000g for 15 minutes. Then, 400µl of sediment collected was used to inoculate the LJ slants. At least two sputum samples from each patient were cultured on LJ slants following standard guidelines (who reference) and were preserved in glycerol medium in -20^{o}C until further processing for DNA extraction. For positive LJ slants, the colony characteristics not resembling to MTBC colony characteristics were discarded to rule out possibility of atypical mycobacteria. Finally, 506 culture growths were collected included for genotyping and other molecular biology work. Hence, 144 cases were either sputum culture negative, or contaminated, and were excluded

from the molecular study. Patients were considered culture negative when no visible colony appeared on LJ slants, and contamination was defined if growth of other organisms such as fungi occurred or if the colony morphology was inconsistent with MTBC. Nevertheless, all of these cases showed at least one positive sputum smear and had chest radiography and symptoms suggestive of TB. Once TB was diagnosed, the cases either started DOTS at GENETUP or were "transferred out" to other DOTS center for the convenience of the patients. In summary, data from all the 650 cases were used for the epidemiological analyses, and data from 506 cases with culture-positive results were used for molecular biological analyses.

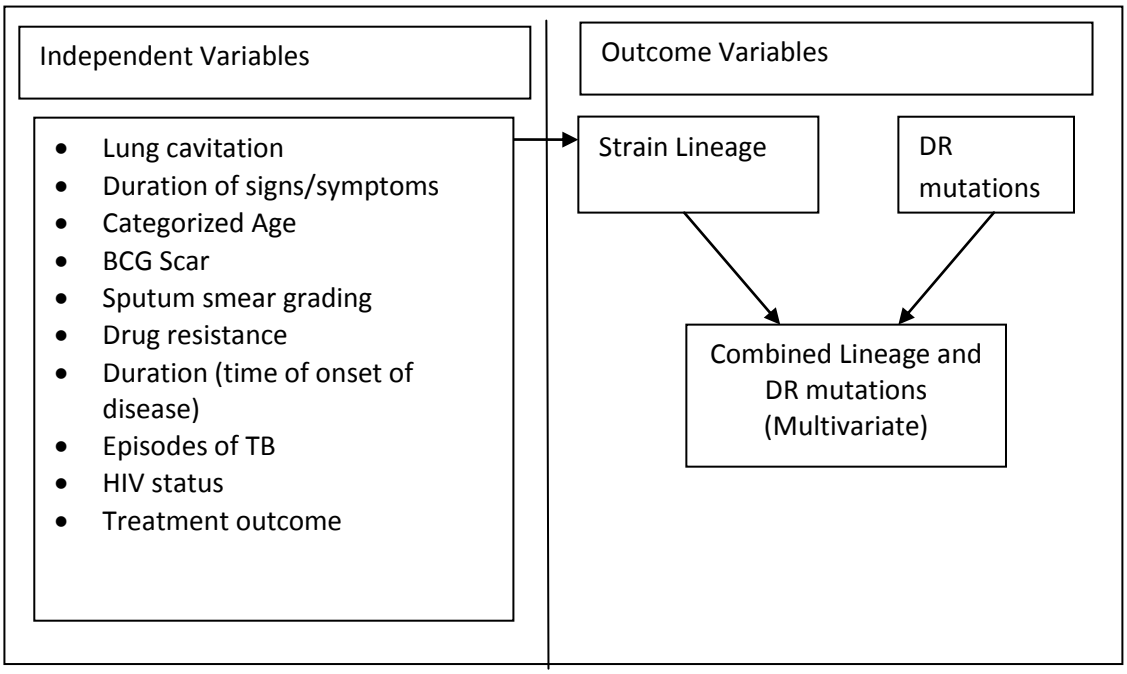

Figure 10: Outline of the multivariate analysis used in this thesis.
The clinical, epidemiological information about patient was considered as independent variable while drug susceptibility results, and strain genotype were used as dependent variables in the statistical analysis. The drug susceptibility results were categorized as any

resistance, multi-drug resistance, and extremely drug resistance for data generation and subsequent statistical analysis.

3.3 SNP typing

The list of genes, primers and probes used are mentioned in Table 4. Preliminary data analysis showed that blue and purple lineage was dominant in Nepal, so SNP for those lineages *Rv2952_0526n* and *Rv3804c_0012* respectively were tested first. Any strains not suggestive of those two lineages were then tested for red and purple SNPs (Table 4). Although many SNPs are accounted for each lineage, use of single SNPs probe as depicted in table below can assign MTBC lineage.

Table 4: List of Lineage, SNPs, Primers and Probes for Lineage typing

Lineage	SNP_Name	Primer	Primer Sequence	Probe	Probe_seq
Euro-American (red)	*katG*463	*katG*463_F	CCGAGATTGCCAGCCTTAAG	H37Rv_probe	6FAM-CAGATCCGGGCATC
		*katG*463_R	GAAACTAGCTGTGAGACAGTCAATCC	Mutant_probe	VIC-CCAGATCCTGGCATC
Blue (RD105)	Rv2952_0526n	Rv2952_F	CCTTCGATGTTGTGCTCAATGT	H37Rv_probe	6FAM-CCCAGGAGGGTAC
		Rv2952_R	CATGCGGCGATCTCATTGT	Mutant_probe	VIC-CCCAGGAAGGTACT
Pink (RV3221c)	Rv3221c_0085n	RV3221c_F	TGTCAACGAAGGCGATCAGA	H37Rv_probe	6FAM-ACAAGGGCGACGTC
		RV3221c_R	GACCGTTCCGGCAGCTT	Mutant_probe	VIC-ACAAGGGCGACATC
Purple (Rv3804c T-C)	Rv3804c_0012	Rv3804c_F	GCATGGATGCGTTGAGATGA	H37Rv_probe	6FAM-AAGAATGCAGCTTGTCGA
		Rv3804c_R	CGAGTCGACGCGACATACC	Mutant_probe	VIC-AAGAATGCAGCTTGTTGA

3.4 Ethical Consideration

This study was ethically approved by the Nepal Health Research Council (NHRC), Nepal, and the Ethics Committee of the Canton of Basel (EKBB), Switzerland. Sputum smear positive cases were required to provide written informed consent for collection and analysis

of samples and demographic/clinical data. If the patient was minor or illiterate, the parent or caretaker was requested for consent. All the information collected from cases was kept confidential. Patient names and identifiers were stripped off prior data analysis. During study, any newly diagnosed TB cases were registered to DOTS at GENETUP or were referred to nearby treatment center of their residence for the DOTS treatment as provided by NTP guidelines. Diagnosed TB cases along with MDR cases received free treatment once registered as TB patient in national TB register.

Sample transport: Following culture, DNA extraction from each MTBC isolate was done at GENETUP and then transported to SwissTPH. Live TB strains were also transported to SwissTPH following the international regulations of bio-safety **(World Health Organization, 2004)**. The transportation of both DNA and the live MTBC strains was approved by NHRC.

Results

Chapter 4: The epidemiological and clinical characteristics of TB patients in Nepal

This study included 650 TB cases from 54 of the 75 districts of Nepal, and six patients from India (Figure 11). The detail is provided in Appendix 2. More than half of the patients (56.3%) were from Kathmandu. The demographic and socioeconomic characteristics of the TB cases including age, educational status, homelessness, geographic location, previous treatment status, comorbidities (HIV and Diabetes status), Chest X-ray, substance abuse, and the treatment outcome for the current episode of TB are presented in this chapter.

As in many developing countries, direct sputum microscopy is the gold standard method for the diagnosis of pulmonary TB in Nepal **(Tuberculosis Division International Union Against Tuberculosis and Lung Disease, 2005)**. Microscopy has a varying sensitivity depending on the quality of the sputum and the skill of the microscopist. Many smear-negative TB cases yield at least one positive culture when tested for three individual sputum samples. On the other hand, smear-positive but culture-negative is seen among cases under treatment who may continue to show dead organisms in sputum samples during follow-up smear microscopy test. This is particularly true among patients with a heavy pre-treatment bacillary load, or because of delays in sputum processing. Although 650 sputum smear-positive TB cases were included in this study, not all the cases were culture-positive. The proportion of smear-positive but culture-negative cases were 22.2% (144/650) as shown in Table 5, and the distribution of these cases was similar in all age groups. This highlights that the sputum smear negativity is likely not only due to difficulty in sputum production among younger patients. Here, we briefly discuss the clinical characteristics of patients with the complete dataset of 650 cases (i.e. irrespective of the culture results).

Table 5: Characteristics of patients with sputum smear positive and with or without culture positive result

	Sputum Smear Positive Culture Negative (N=144)	Sputum Smear Positive Culture Positive (N=506)	Total
Sex of patient	n (%)	n (%)	
Male	99 (68.75)	354 (69.96)	453
Female	45 (33.25)	152 (30.03)	197
Age group (years)			
Up to 24 yrs.	40 (27.78)	169 (33.40)	209
25-34	33 (22.92)	115 (22.73)	148
35-44	21 (14.58)	80 (15.81)	101
45-54	21 (14.58)	66 (13.04)	87
55-64	17 (11.81)	47 (9.29)	64
65-74	12 (8.33)	29 (5.73)	41
Chest X-ray report			
Normal	40 (27.77)	142 (28.06)	182
Cavitary disease	104 (72.22)	364 (71.93)	468
BCG scar present	73 (57.94)	212 (47.86)	285
Signs and Symptoms			
Cough	111 (77.08)	445 (87.94)	556
Night sweat	80 (55.56)	345 (68.18)	425
Chest pain	69 (47.92)	293 (57.91)	362
Hemoptysis	27 (18.75)	139 (27.47)	166
Loss of appetite	68 (47.22)	265 (52.37)	333
HIV positive	5 (3.47)	10 (1.97)	15

The males accounted for 69.69% of the subjects (453/650) and were aged between 5 – 83 years. The patients of age group (0-24 years) were highest in proportion with 27.37% (124/453) and 43.15% (85/197) in male and female patients, respectively. The HIV status was known for 90 cases, and 16.6% (15/90) were co-infected with HIV at the time of TB diagnosis. During clinical examination, 8% (52/650) were known to be diabetic, and this proportion was higher among males (9.05%) than in females (5.58%), however with no significant difference (p=0.134). The percentage of males with habits of alcohol consumption (49.89%) was significantly higher than in females (23.86%) (p=0.000). Smoking was also significantly higher among males (48.34%) compared to females (18.27%) (p=0.00). No difference in lung cavitation was found among sexes (p=0.974). The patient characteristics such as age, sex, HIV status, Diabetic status, X-ray, clinical sign and symptoms were found to be comparable among smear positive/culture negative and smear positive/culture positive

cases (Table 5). Therefore, the rest of the analyses were focused on the smear positive/culture positive cases. Molecular characterization of the MTBC isolates allowed us to use the strain genetic background as an additional factor to relate with geography and patient characteristics from culture positive cases.

The incidence of TB in Nepal differs by geography **(National Tuberculosis Programme, 2011)**, and because the epidemiological risk factors vary in these regions, there is a need to identify the variables that might have contributed to these differences in TB prevalence. In this context, use of molecular tools that provide phylogenetic information that can be linked to the geography is useful for understanding the diversity of MTBC locally, as well as compared to the global population structure of MTBC. Here, we used Single Nucleotide Polymorphisms (SNPs) assays to analyze the distribution of MTBC lineages across Nepal. The MTBC has been divided into phylogenetically ancient and modern lineages **(Brosch *et al.*, 2002)**. The evolutionary "ancient" and "modern" lineages were originally characterized by the presence and absence of TbD1, respectively. This classification was further supported by DNA sequence analysis of multiple genes, which differentiated the global phylogeny of MTBC into six major lineages **(Hershberg *et al.*, 2008)**. The Lineages 1, 5 and 6 represent the "ancient" lineages while Lineage 2, 3, and 4 represent the "modern" lineages **(Hershberg *et al.*, 2008,Brites and Gagneux, 2012)**. The modern epidemic of TB is largely caused by evolutionarily "modern" lineages (except for particular regions such as South-India and the Philippines).

In our dataset, the "modern" lineages accounted for 449/506 (88.74%) of cases, with the proportion of "ancient" lineages being substantially lower (57/506; 11.26%), supporting the hypothesis that modern lineages might be generally more successful **(Arnvig *et al.*, 2011)**. Moreover, within the "modern" lineages, Lineage 2 (157/506; 31.02%) and Lineage 3 (206/506; 40.71%) dominated in Nepal. This result was somehow expected given the

geographic association of these lineages with the neighboring regions to the North (Tibet) and South (North-India), respectively **(Li *et al.*, 2005,Singh *et al.*, 2004,Narayanan *et al.*, 2008,Dong *et al.*, 2012)**. The map illustrates the geographic distribution of the MTBC lineages identified in this study across the 75 districts of Nepal (Figure 11). We found no obvious differences in the relative proportion of MTBC lineage distribution across the different regions of Nepal as shown in Table 6.

Table 6: Geography wise distribution of MTBC lineages

GEO_REGION	Lineage 1 n (%)	Lineage 2 n (%)	Lineage 3 n (%)	Lineage 4 n (%)	Total
Mountain (North)	1 (3.85)	11 (42.31)	11 (42.31)	3 (11.54)	26
Hill	45 (10.9)	126 (30.51)	169 (40.92)	73 (17.68)	413
Terai (South)	11 (18.03)	18 (29.51)	24 (39.34)	8 (13.11)	61
India	0 (0)	2 (33.33)	2 (33.33)	2 (33.33)	6
Total	57 (11.26)	157 (31.03)	206 (40.71)	86 (17.00)	506

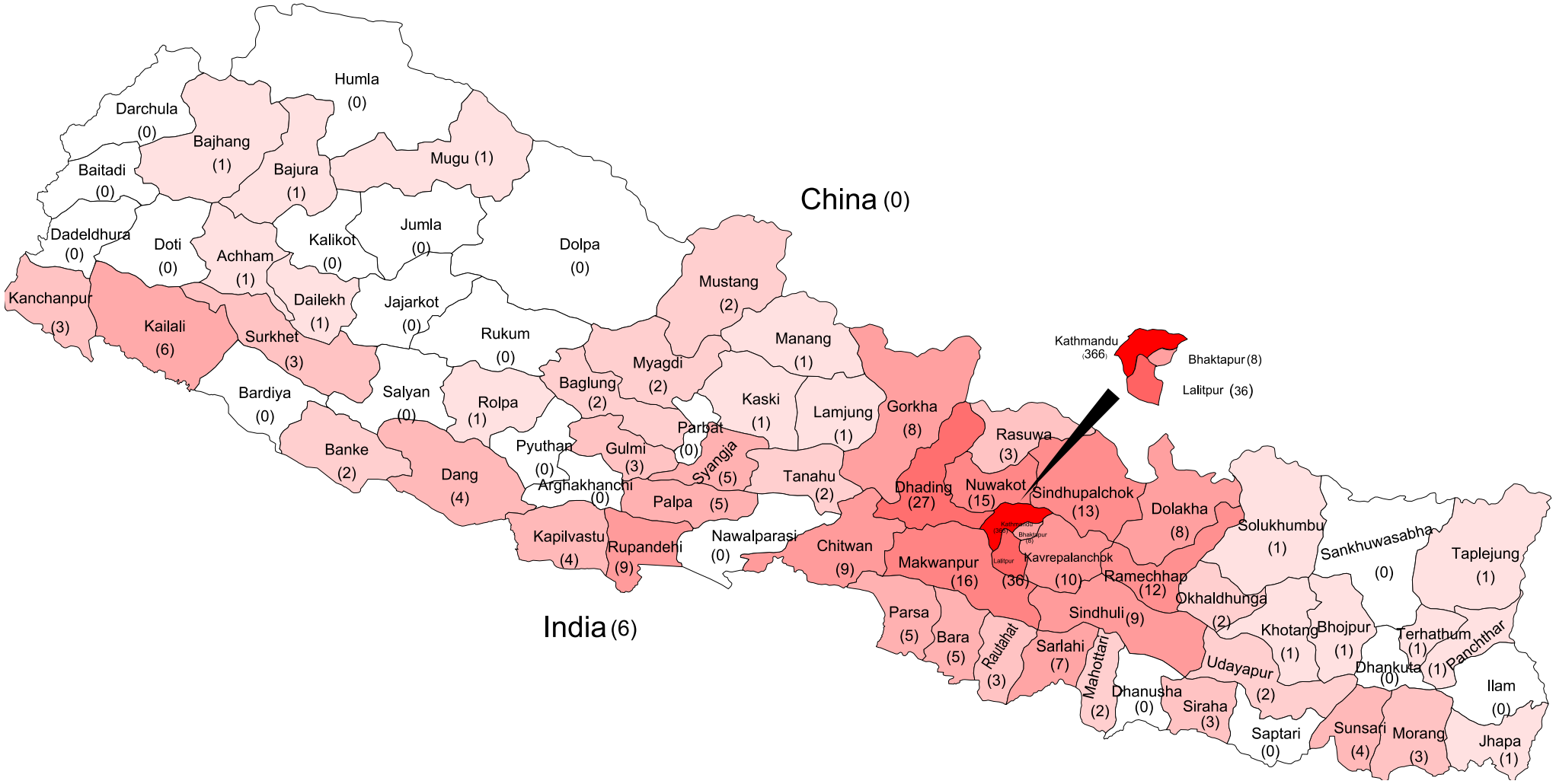

Figure 11: Origin and number of TB patients from the 75 districts in Nepal (n=650)

In addition to SNP-typing, we also performed spoligotyping on all 506 patients isolates for which the bacterial culture was available. A total of 69 different spoligotypes with unique SIT numbers were identified by comparing to the SITVITWEB website (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) **(Demay *et al.*, 2012)**. We found that 155/506 (30.63%) patient isolates showed the of "Beijing" spoligotypes (Shared International Types (SIT 1)) (Table 7). Interestingly, we observed one strain showing the typical "Beijing" spoligotypes but which clearly belonged to Lineage 3 (rather than Lineage 2). These data are further discussion in Chapter 6 **(Malla *et al.*, 2012)**.

Table 7: Lineage wise comparison of patient clinical characteristics

Patient Characteristics	Lineage 1 (n=57)	Lineage 2 (n=157)	Lineage 3 (n=206)	Lineage 4 (n=86)	total (n=506)
Sex of patients					
MALE	46 (80.70)	94 (59.87)	152 (73.79)	62 (72.09)	354 (69.96)
FEMALE	11 (19.30)	63 (40.13)	54 (26.21)	24 (27.91)	152 (30.04)
Age group (years)					
Up to 24 yrs	14 (8.28)	56 (33.14)	69 (40.83)	30 (17.75)	169 (100)
25-34	12 (10.43)	40 (34.78)	42 (36.52)	21 (18.26)	115 (100)
35-44	6 (7.50)	24 (30)	40 (50)	10 (12.50)	80 (100)
45-54	11 (16.67)	15 (22.73)	30 (45.45)	10 (15.15)	66 (100)
55-64	10 (21.28)	12 (25.53)	19 (40.43)	6 (12.77)	47 (100)
65 and more	4 (13.79)	10 (34.48)	6 (20.69)	9 (31.03)	29 (100)
TOTAL	57 (11.26)	157 (31.03)	206 (40.71)	86 (17)	506 (100)
HIV Status					
HIV positive	0	3 (1.91)	7(3.40)	0	10 (1.98)
HIV negative	12 (21.05)	16 (10.19)	21 (10.19)	12 (13.95)	61(12.06)
Unknown	45 (78.95)	138 (87.90)	178 (86.41)	74 (86.05)	435 (85.97)
DIABETES					
Yes	6 (10.53)	14 (8.92)	15 (7.28)	8 (9.30)	43 (8.50)
no or Unknown	51 (89.47)	143 (91.08)	191 (92.72)	78 (90.70)	463 (91.50)
BCG status					
scar present	23 (40.35)	63 (40.13)	86 (41.75)	40 (46.51)	212 (41.90)
Contact of TB case					
Close contact of TB case	13 (22.81)	40 (25.48)	41 (19.90)	18 (20.93)	112 (22.13)
No TB contact	41 (71.93)	113 (71.97)	157 (76.21)	65 (75.58)	376 (74.31)
Unknown	3 (5.26)	4 (2.55)	8 (3.88)	3 (3.49)	18 3.56)
Chest X-ray report					
Normal	4 (7.02)	14 (8.92)	10 (4.85)	4 (4.65)	32 (6.32)
Cavitary	39 (68.42)	113 (71.97)	147 (71.36)	65 (75.58)	364 (71.94)
Non cavitary consistent with TB	6 (10.53)	17 (10.83)	30 (14.56)	10 (11.63)	63 (12.45)

Non cavitary non consistent with TB	1 (1.75)	2 (1.27)	4 (1.94)	1 (1.16)	8 (1.58)
Not done or Unknown	7 (12.28)	11 (7.01)	15 (7.28)	6 (6.98)	39 (7.71)
Signs and symptoms					
Cough	48 (84.21)	138 (87.90)	185 (89.81)	74 (86.05)	445 (87.94)
Duration of cough (median weeks)	4 (2-10)	5 (3-12)	4 (3-12)	4 (3-12)	4 (3-12)
Night sweat	34 (59.65)	113 (71.97)	143 (69.42)	55 (63.95)	345 (68.18)
Duration of night sweat (median weeks)	4 (2-6)	4 (2-8)	4 (2-8)	4 (2-12)	4 (2-8)
Chest pain	29 (50.88)	95 (60.51)	123 (59.71)	46 (53.49)	293 (57.91)
Duration of chest pain (median weeks)	4 (2-8)	4 (2-8)	4 (2-8)	4 (2-10)	4 (2-8)
Haemoptysis	18 (31.58)	40 (25.48)	55 (26.70)	26 (30.23)	139 (27.47)
Duration of Haemoptysis (median weeks)	1 (1-4)	2.5 (1-4)	2 (1-4)	2 (1-3)	2 (1-4)
Loss of appetite	27 (47.37)	82 (52.23)	110 (53.40)	46 (53.49)	265 (52.37)
Duration of loss of appetite (median weeks)	4 (2-8)	4 (3-8)	4.5 (3-8)	8 (4-16)	3 (6-12)
Weight loss	38 (66.67)	102 (64.97)	138 (66.99)	50 (58.14)	328 (64.82)
Duration of weight loss (median weeks)	4.5 (4-8)	8 (4-12)	8 (4-16)	8 (4-14)	8 (4-14)
Breathlessness	18 (31.58)	60 (38.22)	68 (33.01)	24 (27.91)	170 (33.60)
Duration of Breathless (median weeks)	8 (4-16)	4 (2.5-10)	4 (2-12)	4 (2-12.5)	4 (2-12)
History of TB					
Yes	9 (15.79)	60 (38.22)	65 (31.55)	29 (33.72)	163 (32.21)
Number of TB episodes					
0 (None)	48 (84.21)	97 (61.78)	141 (68.45)	57 (66.28)	343 (67.79)
1	9 (15.79)	49 (31.21)	51 (24.76)	22 (25.58)	131 (25.89)
2 or more	0	9 (5.73)	14 (6.80)	7 (8.14)	30 (5.93)
missing	0	2(1.27)	0	0	2(0.39)
Persons per room share					
1 person	25 (46.30)	80 (52.63)	87 (44.16)	37 (44.05)	229 (47.02)
2-3 persons	25 (46.30)	55 (36.18)	87 (44.16)	38 (45.24)	205 (42.09)
4 or more persons	4 (7.41)	17 (11.18)	23 (11.68)	9 (10.71)	53 (10.88)
TB contact in past 2 years					
Yes	13 (22.81)	40 (25.48)	41 (19.90)	18 (20.93)	112 (22.13)
No	44 (77.19)	117 (74.52)	165 (80.10)	68 (79.07)	394 (77.87)
Drug resistance					
MDR	3 (5.26)	17 (10.83)	15 (7.20)	7 (8.14)	42 (8.30)
XDR	0	4 (2.54)	0	0	4 (0.80)
Current smoker					
yes	23 (40.35)	57 (36.31)	87 (42.23)	31 (36.05)	198 (39.13)
No	34 (59.65)	100 (63.69)	119 (57.77)	55 (63.95)	308 (60.87)
Treatment outcome					
Cured	4 (7.02)	5 (3.18)	8 (3.88)	6 (6.98)	23 (4.55)
Completed	0	0	0	1 (1.16)	1 (0.20)
Failure	0	6 (3.82)	4 (1.94)	1 (1.16)	11 (2.17)
Died	0	2 (1.27)	0	2 (2.33)	4 (0.79)
Transferred out	53 (92.98)	144 (91.72)	194 (94.17)	76 (88.37)	467 (92.29)

There is increasing evidence that in addition to host factors **(Ben-Selma *et al.*, 2012)**, the bacterial factors might also influence the outcome of TB **(Coscolla and Gagneux, 2010)**. These include genetic factors **(Weiner *et al.*, 2007,Mathema *et al.*, 2012)**, strain-specific differences in immunological recognition **(Portevin *et al.*, 2011)** and in bacterial fitness **(Borrell and Gagneux, 2009) (Chernyaeva *et al.*, 2012)**. The Table 7 shows the patient characteristics as categorized by MTBC lineage (based on all culture-positive samples: N=506). We did sample collection at a reference laboratory, so the treatment outcome result was limited to only those who had DOTS treatment at GENETUP. Here, the "transferred out" cases meant those patients who once diagnosed at GENETUP were sent to their closest DOTS center for start of DOTS treatment. Hence, the clinical outcome was unknown. Nonetheless, in-depth analyses are currently ongoing to look for lineage-specific differences among these variables; a brief summary is shown in Table 7. A previous analysis on a smaller subset of strains has been published recently and is presented in the following chapter.

Chapter 5: First insights into the phylogenetic diversity of *Mycobacterium tuberculosis* in Nepal

Bijaya Malla [1 2], David Stucki [1 2], Sonia Borrell [1 2], Julia Feldmann [1 2], Bhagwan Maharjan [3], Bhawana Shrestha [3], Lukas Fenner [4 * †], Sebastien Gagneux [1 2 * †]

1 Swiss Tropical and Public Health Institute (Swiss TPH), Basel, Switzerland

2 University of Basel, Switzerland

3 German Nepal Tuberculosis Project, Kathmandu, Nepal

4 Institute of Social and Preventive Medicine, University of Bern, Switzerland

† These authors contributed equally

This article has been published in:

PLoS One. 2012;7(12):e52297. doi: 10.1371/journal.pone.0052297. Epub 2012 Dec 26.

5.1 Abstract

Background: Tuberculosis is a major public health problem in Nepal. Strain variation in *Mycobacterium tuberculosis* may influence the outcome of TB infection and disease. To date, the phylogenetic diversity of *M. tuberculosis* in Nepal is unknown.

Methods and findings: We analyzed 261 *M. tuberculosis* isolates recovered from pulmonary TB patients recruited between August 2009 and August 2010 in Nepal. *M. tuberculosis* lineages were determined by single nucleotide polymorphisms (SNP) typing and spoligotyping. Drug resistance was determined by sequencing the hot spot regions of the relevant target genes. Overall, 164 (62.8%) TB patients were new, and 97 (37.2%) were previously treated. Any drug resistance was detected in 50 (19.2%) isolates, and 16 (6.1%) were multidrug-resistant. The most frequent *M. tuberculosis* lineage was Lineage 3 (CAS/Delhi) with 106 isolates (40.6%), followed by Lineage 2 (East-Asian lineage, includes Beijing genotype) with 84 isolates (32.2%), Lineage 4 (Euro-American lineage) with 41 (15.7%) isolates, and Lineage 1 (Indo-Oceanic lineage) with 30 isolates (11.5%). Based on spoligotyping, we found 45 different spoligotyping patterns that were previously described. The Beijing (83 isolates, 31.8%) and CAS spoligotype (52, 19.9%) were the dominant spoligotypes. A total of 36 (13.8%) isolates could not be assigned to any known spoligotyping pattern. Lineage 2 was associated with female sex (adjusted odds ratio [aOR] 2.58, 95% confidence interval [95% CI] 1.42-4.67, p=0.002), and any drug resistance (aOR 2.79; 95% CI 1.43-5.45; p=0.002). We found no evidence for an association of Lineage 2 with age or BCG vaccination status.

Conclusions: We found a large genetic diversity of *M. tuberculosis* in Nepal with representation of all four major lineages. Lineages 3 and 2 were dominating. Lineage 2 was associated with clinical characteristics. This study fills an important gap on the map of the *M. tuberculosis* genetic diversity in the Asian region.

5.2 Introduction

Tuberculosis caused by *Mycobacterium tuberculosis* remains a global health threat with an estimated nine million incident cases and 440,000 multidrug-resistant TB cases worldwide **(World Health Organization, 2010)**. The incidence of TB was 163 per 100,000 population in 2010, and multidrug resistance (MDR) occurred in 2.9% of new cases and 11.7% of previously treated cases based on the most recent drug resistance survey in 2006 **(World Health Organization, 2011,National Tuberculosis Program, 2010)**. In Nepal, the NTP adopted Directly Observed Short Course therapy (DOTS) in 1995.

Mycobacterium tuberculosis complex (MTBC) has a global phylogeographic population structure consisting of six main phylogenetic lineages **(Gagneux *et al.*, 2006,Comas *et al.*, 2009,Gutacker *et al.*, 2002)**: Lineage 1 (also known as Indo-Oceanic Lineage), Lineage 2 (East-Asian Lineage, includes the Beijing genotype), Lineage 3 (Delhi/CAS), Lineage 4 (Euro-American Lineage), and Lineages 5 and 6 (*M. africanum* West African lineages 1 and 2). These lineages are associated with specific geographic regions and human populations **(Gagneux *et al.*, 2006,Gagneux and Small, 2007,Baker *et al.*, 2004,Hirsh *et al.*, 2004)**. Lineage 2, for example, is most often isolated in countries in Asia and the former Soviet Union **(Sun *et al.*, 2011)**. There is increasing evidence that in addition to host and environmental factors, the epidemiology of TB may also be influenced by bacterial strain variation **(Coscolla and Gagneux, 2010,Caws *et al.*, 2008,Thwaites *et al.*, 2008,de Jong *et al.*, 2008,Dou *et al.*, 2008,Tho *et al.*, 2012,Hanekom *et al.*, 2011,Lari *et al.*, 2009,Click *et al.*, 2012)**. For example, Lineage 2 (includes the Beijing genotype) has been repeatedly associated with drug resistance in a wide range of settings and countries **(Borrell and Gagneux, 2009,Parwati *et al.*, 2010,Pang *et al.*, 2012,Fenner *et al.*, 2012)**, while a few studies could not find evidence for such an association **(Rajapaksa and Perera, 2011,Iwamoto *et al.*, 2008,Lasunskaia *et al.*, 2010)**.

There are several genotyping techniques to define the genetic diversity of *M. tuberculosis* **(Gagneux and Small, 2007,Malik and Godfrey-Faussett, 2005,Supply *et al.*, 2006)**. Spoligotyping is a widely used genotyping technique **(Brudey *et al.*, 2006,Kamerbeek *et al.*, 1997)**. It is based on the repetitive DNA region known as the Direct Repeat (DR) locus in *M. tuberculosis* **(Supply *et al.*, 2006)**. This region is characterized by series of direct repeats interspersed by short unique regions called "spacers". However, these spacers exhibit a high rate of change, and convergent evolution can lead to identical genetic character states in phylogenetically unrelated strains **(Comas *et al.*, 2009,Fenner *et al.*, 2011)**. By contrast, genomic deletions and single nucleotide polymorphism (SNPs) evolve more slowly. Recent advances in comparative genomics have led to the development of more robust markers to study the genetic diversity **(Gagneux and Small, 2007,Sreevatsan *et al.*, 1997,Supply *et al.*, 2003,Brudey *et al.*, 2006,Niemann *et al.*, 2009,Stucki *et al.*, 2012)**, and are therefore ideal for determining phylogenetic lineages and sub-lineages **(Coscolla and Gagneux, 2010)**.

Nepal lies between two high TB burden countries, India and China which together account for one third of the world's TB cases **(World Health Organization, 2010)**. To date, there are no data on the phylogenetic diversity of *M. tuberculosis* in Nepal. The aims of the study were to describe the main *M. tuberculosis* lineages and spoligotypes circulating in Nepal, and to explore possible associations with clinical and epidemiological characteristics.

5.3 Methods

Ethics statement

This study was approved by the Nepal Health Research Council, Nepal and the Ethics Committee of the Canton of Basel (EKBB), Switzerland. All study participants provided written informed consent. After diagnosis, the TB cases were referred to DOTS centers for treatment as provided by the Nepal Government's National TB Control Program.

Study setting

The study was based on a convenience sample of TB patients mainly representing populations from Kathmandu and the surrounding area. TB suspects who reported symptoms of TB including cough for more than two weeks, chest pain, night sweat and fever were recruited at the German Nepal Tuberculosis Project (GENETUP), Kathmandu, Nepal. Patients already undergoing DOTS therapy were also enrolled, if found smear-positive during follow-up visits. GENETUP is a national reference laboratory, technically and financially supported by "Kuratorium Tuberkulose in der Welt e. V." (Gauting, Germany), and is the main referral center for culture and drug susceptibility testing to diagnose MDR and extensively drug-resistant TB.

Study population and data collection

We included a total of 261 culture-confirmed TB cases diagnosed between August 2009 and August 2010. We collected socio-demographic and clinical data including previous TB episodes, treatment history, HIV, and BCG vaccination status. The information was collected by physicians and trained medical and nursing staff. A new case of TB was defined as a patient who had not taken anti-TB drugs for at least one month according to WHO guidelines **(World Health Organization, 2009)**. A previously treated case was defined as a patient who received TB treatment for one month or more. BCG vaccination status was defined based on the presence or absence of a BCG scar.

Culture, DNA extraction and identification of *M. tuberculosis* complex

Sputum samples were cultured on Löwenstein Jensen (LJ) growth medium following standard microbiological laboratory procedures. The DNA was extracted by re-suspension of MTBC colonies in 500 µl of sterile distilled water, heat killed at 90° C for one hour, and centrifuged. The supernatants were preserved at 4°C until further use. MTBC strains were identified by multiplex polymerase chain reaction (PCR) by targeting the *rpoB* gene region. We used the forward primers K-0155 (5'-TCCTCGATGACGCCGCTTTCT-3') and K-0209 (5'-AYATCGACCACTTCGGYAACC-3'), and the reverse primer K-0156 (5'-TCRGAGATCTTGCGCTTCTGS-3'). PCR conditions were as follows: initial denaturation step for 5 minutes at 96° C, 35 amplification cycles of 96° C for 40 secs (denaturation), 62° C for 30 secs (annealing), 72° C for 1 min (extension), and a final extension cycle of 7 minutes at 72° C. The amplicons were separated by electrophoresis on a 2% agarose gel. The PCR yielded a 849 bp amplicon in *M. tuberculosis* isolates, compared to a 1539 bp amplicon in non-tuberculous mycobacteria. All *M. tuberculosis* isolates were stored in glycerol medium at -70^{o} C.

Determination of the main *M. tuberculosis* lineages

We determined the main phylogenetic lineages of *M. tuberculosis* by real-time PCR using fluorescence-labeled probes (Taqman, Applied Biosystems, USA) targeting lineage-specific SNPs as previously described **(Gagneux *et al.*, 2006,Sreevatsan *et al.*, 1997,Stucki *et al.*, 2012)**.

Spoligotyping

Spoligotyping was performed according to the manufacturer's instructions, using commercially available kits from Isogen Bioscience BV (Maarssen, The Netherlands) **(Kamerbeek *et al.*, 1997,Lillebaek *et al.*, 2003)**. Spoligotyping patterns were defined

according to the definitions in the SITVITWEB database (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) accessed on October 22, 2012. The SITVITWEB global database has documented 7,105 spoligotyping patterns from a global collection of 53,816 strains **(Demay *et al.*, 2012)**. All patterns that could not be assigned were considered orphan spoligotypes.

Molecular drug resistance testing

As phenotypic drug susceptibility testing results were not available for all strains, we used molecular methods to detect drug resistance in our study. Molecular drug resistance testing was performed on all strains by direct sequencing of the hotspot regions of the target genes for rifampicin (*rpoB*), isoniazid (*inhA* promoter region and *katG*), and streptomycin (*rpsL*). MDR strains were then further sequenced and analyzed for ethambutol (*embB)*, fluoroquinolones (*gyrA*) and aminoglycoside (*rrs)* resistance by sequencing of the relevant gene segments. For the *rpoB* region, we used an in-house PCR assay with primer pair K-0155 and K-0209 as described above. For all other target genes, PCR primers and PCR conditions were adapted from previously published studies **(Victor *et al.*, 1999,Feuerriegel *et al.*, 2009,Brossier *et al.*, 2010)**.The sequences were analyzed with *M. tuberculosis* H37Rv as reference sequence using the Staden software package **(Staden *et al.*, 2000,Bonfield *et al.*, 1995)**, and compared to the publicly available web-based database (http://www.tbdreamdb.com/) **(Sandgren *et al.*, 2009)**. Any drug resistance was defined as resistance to isoniazid, rifampicin, streptomycin, ethambutol, fluoroquinolones, and/or aminoglycosides. MDR was defined as resistance to at least isoniazid and rifampicin.

Statistical analyses

We used Chi-square test to test the statistical significance of differences between groups in binary variables, and the Kruskal Wallis rank test for continuous variables. Logistic regression models were used to compare patient characteristics associated with Lineage 2

(includes the Beijing genotype) compared to all other lineages (Lineages 1, 3 and 4), adjusted for age, sex, treatment history, BCG vaccination status, and any drug resistance. All statistical analyses were performed in STATA 10.1 (Stata Corp., College Station, TX, USA).

5.4 Results

Patient characteristics

Of the 261 patients included in this study, 164 (62.8%) were new TB cases. Overall, 182 (69.73%) were male, and the median age was 31 years (interquartile range [IQR] 23-50). Females were significantly younger than males (median age 26 versus 35 years, $p<0.001$). HIV status was known in 26 patients; of these 8 (30.8%) were HIV-positive. Most patients originated from Kathmandu valley (153 cases, 58.6%), followed by 104 cases (39.8%) from different districts of Nepal, and four patients (1.5%) who were born in India.

Mycobacterium tuberculosis genotyping and lineage assignment

We analyzed a total of 261 *M. tuberculosis* isolates (one from each patient). The SNP-typing results showed the presence of four different *M. tuberculosis* lineages (Table 8). The most frequent lineages were Lineage 3 (includes CAS/Delhi) with 106 isolates (40.6%) and Lineage 2 (East-Asian lineage, includes Beijing genotype) with 84 isolates (32.2%). Forty one isolates (15.7%) belonged to Lineage 4 (Euro-American Lineage), and 30 isolates (11.5%) to Lineage 1 (Indo-Oceanic Lineage). Lineages 5 and 6 (*M. africanum* West African lineages) were not found in our sample.

Table 8: Description of the main *M*. tuberculosis lineages and spoligotyping patterns from Nepal (n=261)

Lineage	Spoligotyping family	Spoligotyping pattern (spacers 1-43)	SIT	n	%
1	EAI5	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■□■■■■■□□□□	138	10	33.33
1	EAI3_IND	■□□■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■□■■□□□■■■■	11	4	13.33
1	EAI1_SOM	□□□■■■■■■■■■■■□■■■■■■■■■■■■■□□□□■□■■■■■□■■■	1734	1	3.33
1	EAI1_SOM	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■□■■■■■□■■■	48	1	3.33
1	EAI2_MANILLA	■■□■■■■■■■■■■■■■■■□□■■■■■■■□□□□■□■■■■■■■■■■	19	1	3.33
1	EAI6_BGD1	■■■■■■■■■■■■■■■■■■■■■□■■■■■■□□□□■□■■□■■■■■■	292	1	3.33
1	Orphan	■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□■□■■■■■□□□□	-	1	3.33
1	Orphan	□□□■■■■■■■■■■□■■■■■■■■■■■■■□□□□■□■■□□□■■■■	-	1	3.33
1	Orphan	■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■□■■■■■□■■■	-	1	3.33
1	Orphan	■■■■■■■■■□■■■■■■■■■■■■■■■■□□□□■□■■■■■□□□□	-	1	3.33
1	Orphan	■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■□□□■■■□□□□	-	1	3.33
1	Orphan	■■■■■■■■■■■□■■■■■■■■■■■■■□□□□■□■■■■■□□□□	-	1	3.33
1	Orphan	■■■■■■■■■■■■■□□□■■■■■■■■■■□□□□■□■■■■■□□□□	-	1	3.33
1	Orphan	■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□■□□□□■■□■■■■■■	-	1	3.33
1	Orphan	■■■■■■■■■■■■■□□□■□□□□□□□■□□□□□□□■■□■■■■■■	-	1	3.33
1	Orphan	■■■■■■■■■■■■■□□□□■■■■■■■■■□□□□■□■■■■■□□□□	-	1	3.33
1	Orphan	■□□■■■■■■■■■■□■■■■■■■■■■■■■□□□□■□■■□□□■■■■	-	1	3.33
1	Orphan	■□□■■■■□□□□□□□□□□□■■■■■■■■■■■□□□□■□■■□□□■■■■	-	1	3.33
Subtotal				30	100.00
2	Beijing	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	1	82	97.62
2	Beijing	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■□■■	941	1	1.19
2	Beijing-like	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■	250	1	1.19
Subtotal				84	100.00
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□■■■■■■■■■	26	52	49.06
3	CAS	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□□■■■■■■■■	599	9	8.49
3	CAS	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□□□■■■■■■■	357	6	5.66
3	CAS2	■■■□□□□□□□■■■■■■■■■■■■□□□□□□□□□□□□■■■■■■■■■	288	3	2.83
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□■■□■■■■■■	428	3	2.83
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□■□■■■■■■■■	1091	3	2.83
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□■■□□□□□■■	2147	2	0.94
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□■■□□■■■■■	25	1	0.94
3	U (CAS_ANCESTOR)	■■■□□□□■■■■■■■■■■■■■□□■■■■■■■■□□□□■■■■■■■	27	1	0.94
3	CAS1_DELHI	■■■□□□□■■■■■■■□■■■■■■■□□□□□□□□□□□□■■■■■■■■	141	1	0.94
3	CAS	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□■■■■■■■■■	142	1	0.94
3	CAS	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□■■■■■■■■	203	1	0.94
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□■■□□□■■■■	381	1	0.94
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□■■■■■□■■■	429	1	0.94
3	CAS	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□■■■■■■■■■	1093	1	0.94
3	CAS	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□■■■■■■■■	1422	1	0.94
3	CAS	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□■■■■■■■■■	1551	1	0.94
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■□■■■□□□□□□□□□□□□■■■■■■■■	1590	1	0.94
3	CAS1_DELHI	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□□■■■■■■■	1789	1	0.94
3	Beijing [c]	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	1	1	0.94
3	Orphan	■■■□□□□■■■■■■■■■□■■■■■□□□□□□□□□□□□□□□■■■■■	-	3	0.94
3	Orphan	■■■□□□□■■■■■■■■■■■□□□□□□□□□□□□□□□□□□□□■■■■	-	1	0.94
3	Orphan	■■■□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■□■■■	-	1	0.94
3	Orphan	■■■□□□□■■■■■■■■□■■■■■■□□□□□□□□□□□□□■■■■■■■■	-	1	0.94
3	Orphan	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□□□■■■■■■■	-	1	0.94
3	Orphan	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□■■■□■■■■■	-	1	0.94
3	Orphan	■■■□□□□■■■■■■■■■■■■■□■■■■■■■■□□□□■■■■■■■■■	-	1	1.89
3	Orphan	■■□□□□□■■■■■■■■■■■■■□□□□□□□□□□□□■■□□■■■■■■	-	1	0.94
3	Orphan	■■■□□□□■■■■■■■■■□□□□□□□□□□□□□□□□□□□□□□■■■■	-	1	0.94
3	Orphan	■■■□□□□■■■■■■■■■■■□■■■■■□□□□□□□□□□□■■■□□□□□■	-	1	0.94
3	Orphan	■■■□□□□□□□■■■■■■■■■■■□□□□□□□□□□□□□□□□□□□□□□	-	1	0.94
3	Orphan	■■■□□□□■■■■■■■■■■■■■□□□□□□□□□□□□□□□□□□□□□□□	-	1	2.83
3	Orphan	■■■□□□□■■■■■■■■■■□□□□□□□□□□□□□□□□□□□□□□■■■■	-	1	0.94
Subtotal				106	100.00
4	H3	■□□■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■	655	6	14.63
4	T2	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	52	2	4.88
4	T3	■■■■■■■■■■■■□■■■■■■■■■■■■■■■■□□□□■■■■■■■	37	3	7.32
4	S	■■■■■■■■□□■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	34	2	4.88
4	T1	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	53	4	9.76
4	X2	■■■■■■■■■■■■■■■■■□■■■■■■■■■■■■■□□□□■■□□□□■	137	2	4.88
4	T1	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■□□□□■	244	2	4.88
4	H1	■■■■■■■■■■■■■■■■■■■■■■■□■□□□□□□■□□□□■■■■■■■	620	2	4.88
4	LAM9	■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■■□□□□■■■■■■■	42	1	2.44
4	H3	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■■	50	1	2.44
4	T2-T3	■■■■■■■■■■■■□■■■■■■■■■■■■■■■■□□□□■■■□■■■	73	1	2.44
4	T1	■■■■□■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	154	1	2.44
4	T1	■■■■■■■■■■■■□□□□■■■■■■■■■■■■■■■□□□□■■■■■■■	102	1	2.44
4	X1	■■■■■■■■■■■■■■■■■□■■■■■■■■■■■■■□□□□■■■■■■■	119	1	2.44
4	H1	■■■■■■■■■■■■■■■■■■■■■□□□■□□□□□□■□□□□■■■■■■■	283	1	2.44
4	T1	■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■□□	628	1	2.44
4	H3	■■■■■■■■■■■■■□■■■■■■■■■■■■■■■□□■□□□□■■■■■■■	929	1	2.44
4	Orphan	■■■■■■■■■■■■■■■■■■■■■□■□□□□■■■■■■□□□□■■■■■■■	-	1	2.44
4	Orphan	□■■■■■□■■■■■■■■■■□■■■■■■■■■■■■■□□□□■■■■■■■	-	1	2.44
4	Orphan	■■■■■■□■■□■■■■■■■■□■■■■■■■■■■■■■□□□□■■■■■■■	-	1	2.44
4	Orphan	■□■■■■■■■■■■■■■□□■■■■■■■■■■■■■■■□□□□■■■■■■■	-	1	2.44
4	Orphan	■□■■■■■■■■■■■■■■■■■■■■■■■■■■■■■□□□□■■■■■■■	-	1	2.44
4	Orphan	■■■□■■■■■□□■■■■■■■■■■■■■■■■■■■■□□□□■■■■□■■	-	1	2.44
4	Orphan	■■■■■■■■■■■■■■■■■■□■■□□□□□□□■□□□□■■■■■■■	-	1	2.44
4	Orphan	■■■■■□■■■■■■■■■■■■■□■■■■■■■■■■■■■□□□□■■■■■■■	-	1	2.44
4	Orphan	■■□■■■■■■■■■■■■■■■■■■■■■■■■■□□□□□□□□□■■■■■■■	-	1	2.44
Subtotal				41	100.00

[c] This strain was assigned to Lineage 3 (Delhi/CAS) based on alternative molecular markers, as previously published (Fenner et al., 2011)
SIT, Spoligotype International Type according to the definitions in the SITVITWEB database (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) accessed on April 2012. Lineage 1: Indo-Oceanic Lineage; Lineage 2: East-Asian Lineage (includes Beijing strains); Lineage 3: Delhi/CAS; Lineage 4: Euro-American Lineage

Based on spoligotyping, we detected 45 different spoligotypes (SITs) corresponding to 225 *M. tuberculosis* isolates (Table 8). The remaining 36 (13.8%) strains could not be assigned to any known spoligotyping pattern in the SITVITWEB database, and were therefore considered orphan spoligotypes. The spoligotyping results showed that CAS family (90, 34.5%) and Beijing (84 isolates, 32.2%) were the predominant spoligotypes in our sample (Table 8). Among the CAS family, the most prevalent spoligotype was CAS1_DELHI (SIT 26) representing 52 (19.9%) isolates, and almost all Beijing isolates (83 of 84 isolates belonging to Lineage 2) showed the classical Beijing spoligotyping pattern. Of the 41 strains belonging to Lineage 4, we found spoligotypes that have been reported before in India or Tibet (LAM9, H3, T2-T3, T1, XI, H1, and H3) according to the SITVITWEB database. Among the 30 (11.5%) Lineage 1 strains, only 18 (60.0%) matched the SITs of the East African Indian (EAI) family. Only two SIT types SIT 138 (EAI5; n=10), and SIT 11 (EAI3_IND; n=4) were represented by more than one strain. However, SIT 1734 (EAI1_SOM) present as a single isolate in our dataset was not reported before from the Indian sub-continent according to the SITVITWEB database. When comparing SNP typing with the spoligotyping results, we found one case of "pseudo-Beijing" spoligotpype as previously reported **(Fenner *et al.*, 2011)**.

Drug resistance

Overall, 50 (19.2%) *M. tuberculosis* isolates had any drug resistance and 16 (6.1%) were MDR as determined by DNA sequencing of the main target regions (Table 9). Any drug resistance was more frequently detected among previously treated TB cases (29 cases, 30.0%) compared to new cases (21 cases, 12.8%, p=0.001). Among the 16 MDR strains, 9 (56.3%) were assigned to Lineage 2 (East-Asian Lineage), 6 (37.5%) to Lineage 3 (CAS/Delhi), and one (6.2%) to Lineage 4 (Euro-American Lineage).

Association between *M. tuberculosis* lineages and patient characteristics

We observed that the proportion of female sex was different across the four main *M. tuberculosis* lineages. Lineage 2 isolates were more common among females (41.7%), compared to other lineages (range 13.3% to 27.4%, overall p=0.016, Table 9). Moreover, any drug resistance was more frequently detected in Lineage 2 isolates (31.0%) than in any other lineages (range 13.2% to 14.6%, overall p=0.011). Other patient characteristics such as age, previous treatment history, or BCG vaccination were not significantly associated with any of the four lineages (Table 9).

Table 9: Associations of patient characteristics across the four main Mycobacterium tuberculosis lineages identified in Nepal

Patient characteristics	Total n (%)	Lineage 1 (n=30)	Lineage 2 (n=84)	Lineage 3 (n=106)	Lineage 4 (n=41)	*P* value
Age, median (IQR), years	31 (23-50)	42 (24-50)	30 (23.5-50.5)	30 (23-45)	38 (23-55)	0.50
Female sex	79 (30.3)	4 (13.3)	35 (41.7)	29 (27.4)	11 (26.8)	0.016
Previously treated	97 (37.2)	8 (26.7)	39 (46.4)	35 (33.0)	15 (36.6)	0.15
BCG vaccinated	110 (42.2)	13(43.3)	31 (36.9)	46 (43.4)	20 (48.8)	0.62
Any resistance	50 (19.2)	4 (13.3)	26 (30.9)	14 (13.2)	6 (14.6)	0.011
MDR	16 (6.1)	0	9 (10.7)	6 (5.7)	1 (2.4)	0.14 [d]

Because Lineage 2 (includes Beijing genotype) has been previously associated with particular characteristics **(Caws *et al.*, 2008,Thwaites *et al.*, 2008,Parwati *et al.*, 2010,Drobniewski *et al.*, 2005)**, and because Lineage 2 was the second most common lineage in our sample, we tested whether these characteristics were also associated with Lineage 2 in our setting by comparing our Lineage 2 isolates to the other lineages combined (Table 3). Logistic regression analyses showed that Lineage 2 was associated with female sex (adjusted odds ratio [aOR] 2.58; 95% confidence interval [95%CI] 1.42-4.67, p=0.002) and any drug resistance (aOR 2.79; 95%CI 1.43-5.45, p=0.002). A history of previous TB treatment tended

[d] Fisher's exact test

BCG, Bacille Calmette Guerin; IQR, Interquartile range; MDR, Multidrug-resistant

to be associated with Lineage 2 (aOR 1.68, 95% CI 0.95-2.97, p=0.074), while BCG vaccination status was not associated with Lineage 2 (includes Beijing genotype) compared to other lineages (aOR 0.67; 95%CI 0.37-1.20, p=0.18).

5.5 Discussion

We analyzed 261 *M. tuberculosis* isolates from Nepal using SNP typing and spoligotyping. We found that four main phylogenetic lineages of *M. tuberculosis* were present in Nepal. Lineage 2 (East-Asian Lineage, includes the Beijing genotype) and Lineage 3 (CAS/Delhi) were the most frequent, while Lineage 1 (Indo-Oceanic Lineage) and Lineage 4 were less prevalent. Spoligotyping revealed a large genetic diversity with the predominant spoligotyping families being Beijing and CAS/Delhi, and nearly 14% of spoligotyping patterns previously unreported.

Because Nepal is geographically located between India and Tibet (China), we expected to observe similar *M. tuberculosis* genotypes in Nepal as in these neighboring countries. Indeed, Lineage 3 (corresponds to Delhi/CAS spoligotype), which was the most common *M. tuberculosis* genotype in our sample, was previously shown to be predominant in Northern India **(Singh *et al.*, 2004,Narayanan *et al.*, 2008,Svensson *et al.*, 2011)**. Similarly, Lineage 2 (includes Beijing), which was the second most common genotype in our study has been reported as the most frequent among TB cases from China (including Tibet) **(Dong *et al.*, 2012,Pang *et al.*, 2012,Liu *et al.*, 2011,Hu *et al.*, 2009,Guo *et al.*, 2011,Han *et al.*, 2007)**. The prevalence of the Beijing genotype of 32.2% in our study is in the range of the prevalence reported from other Asian countries, ranging from 17% in Malaysia to 72% in Japan **(European Concerted Action on New Generation Genetic Markers and Techniques for the Epidemiology and Control of Tuberculosis, 2006)**. Lineage 1 which is association with South-Indian region, Bangladesh and the Philippines was also present in our study sample **(Gagneux *et al.*, 2006)**.

We observed a discrepancy between SNP typing and spoligotyping results. Spoligotyping is based on the highly variable DR locus, and convergent evolution may therefore lead to homoplasy in spoligotyping patterns **(Comas *et al.*, 2009)** . We found a strain with a Beijing spoligotype, which was assigned to Lineage 3 (includes CAS genotype) rather than to Lineage 2 (includes Beijing) based on alternative molecular markers. We have previously published this phenomenon as "Pseudo-Beijing" **(Fenner *et al.*, 2011)**. In Asian countries with a high prevalence of Beijing spoligotypes, it is likely that this phenomenon may be observed in other settings.

We found that Lineage 2 was associated with female sex, which is in line with a previous study from Vietnam **(Buu *et al.*, 2009)**. In contrast to other studies **(Buu *et al.*, 2009,Buu *et al.*, 2009)** however, we found no evidence for an association between Lineage 2 and age. Our observation may be explained by bacterial factors or genetic host factors. Young and middle-aged women may be more likely to progress from infection to disease than men **(Holmes *et al.*, 1998,Borgdorff *et al.*, 2000)**. Alternatively, our results may be influenced by recruitment of more young females than young males into our study. Indeed, females were younger than males in our study population. Overall, our study population showed a male-to-female ratio of 2.3:1 which is similar to the global estimate of 1.9:1 reported by WHO **(World Health Organization, 2011)**, and may reflect differences in access to health care **(Connolly and Nunn, 1996,Getahun *et al.*, 2010)**. Furthermore, sex differences in TB case notification rates among males and females have been noted before in other settings **(Neyrolles and Quintana-Murci, 2009,Uwizeye *et al.*, 2011)**.

Lineage 2 was also associated with any drug resistance. This is consistent with previous studies from different settings **(Parwati *et al.*, 2010,Fenner *et al.*, 2012)**. The reasons for this association remain unknown **(Borrell and Gagneux, 2009)**, but the strain genetic background of Beijing strains **(Borrell and Gagneux, 2011)** and their interactions with the

human immune system may play a role **(Parwati *et al.*, 2010)**. Alternatively, this association might reflect higher relapse rates in patients infected with Beijing strains **(Sun *et al.*, 2006)**. Indeed, in our study, Lineage 2 included more patients that were previously treated but this association was not statistically significant. Finally, previous studies hypothesized that Beijing strain may escape the protective immunity of BCG vaccination **(Parwati *et al.*, 2010)**, but we found no evidence for such an association between Lineage 2 and BCG immunization. BCG immunization has been introduced in Nepal more than 30 years ago, with an estimated immunization coverage of 96% in 2009 **(World Health Organization, 2011)**. However, larger studies may be required for a more complete understanding of the association between previous BCG vaccination and particular *M. tuberculosis* genotypes.

Our study has several limitations. First, the study was not population-based as patients were recruited only at GENETUP (Kathmandu), and patients diagnosed at other microscopy centers during the study period could not be included. Second, patients coming from more remote areas outside of Kathmandu might be more likely to be referred as drug resistance suspects. Therefore, this may have artificially increased the proportion of drug-resistant strains in our sample. Third, although our study covered samples from forty different districts of Nepal including those bordering with India and Tibet, half of patients were from the Kathmandu area. Therefore, the study results mainly reflect the genetic diversity of the strains from the patients who visited GENETUP.

In conclusion, we found a high diversity of *M. tuberculosis* genotypes in Nepal with representation of all four main *M. tuberculosis* lineages, and showed that Lineage 2 (includes Beijing genotype) was associated with female sex and any drug resistance. This study fills the gap on the map of the genetic population structure of *M. tuberculosis* in the Asian region by providing a first insight into the phylogenetic lineages of *M. tuberculosis* circulating in Nepal.

5.6 Acknowledgements

We would like to thank all the staff members at German Nepal Tuberculosis Project (GENETUP) for their cooperation. We also thank the authorities at the National Tuberculosis Center (NTC, Nepal), and the "Kuratorium Tuberkulose in der Welt e.V." (Germany) for their support.

Presentation

This work was presented in part ["First insights into the genetic diversity of *Mycobacterium tuberculosis* in Nepal" (abstract no. OP-701-28)] at the 42th World Congress on Lung Health, Lille, France, October, 26-30, 2011.

Chapter 6: “Pseudo-Beijing”: Evidence for convergent evolution in the Direct Repeat region of *Mycobacterium tuberculosis*

Lukas Fenner [1*], Bijaya Malla [2 3], Béatrice Ninet [4], Olivier Dubuis [5], David Stucki [2 3], Sonia Borrell [2 3], Thembela Huna [6], Thomas Bodmer [7], Matthias Egger [1], Sebastien Gagneux [2 3]

1 Institute of Social and Preventive Medicine, University of Bern, Bern, Switzerland

2 Department of Medical Parasitology and Infection Biology, Swiss Tropical and Public Health Institute, Basel, Switzerland

3 University of Basel, Basel, Switzerland

4 Laboratory of Bacteriology, University Hospital of Geneva, Geneva, Switzerland

5 Viollier AG, Basel, Switzerland

6 MRC National Institute for Medical Research, London, United Kingdom

7 Mycobacteriology Unit, Institute for Infectious Diseases, University of Bern, Bern, Switzerland

This article has been published in:

PLoS One. 2011; 6(9):e24737. doi:10.1371/journal.pone.0024737. Epub2011Sep13.

6.1 Abstract

Background: *Mycobacterium tuberculosis* has a global population structure consisting of six main phylogenetic lineages associated with specific geographic regions and human populations. One particular *M. tuberculosis* genotype known as “Beijing” has repeatedly been associated with drug resistance and has been emerging in some parts of the world. “Beijing” strains are traditionally defined based on a characteristic spoligotyping pattern. We used three alternative genotyping techniques to revisit the phylogenetic classification of *M. tuberculosis* complex (MTBC) strains exhibiting the typical “Beijing” spoligotyping pattern.

Methods and Findings: MTBC strains were obtained from an ongoing molecular epidemiological study in Switzerland and Nepal. MTBC genotyping was performed based on SNPs, genomic deletions, and 24-loci MIRU-VNTR. We identified three MTBC strains from patients originating from Tibet, Portugal and Nepal which exhibited a spoligotyping patterns identical to the classical Beijing signature. However, based on three alternative molecular markers, these strains were assigned to Lineage 3 (also known as Delhi/CAS) rather than to Lineage 2 (also known as East-Asian lineage). Sequencing of the RD207 in one of these strains showed that the deletion responsible for this “Pseudo-Beijing” spoligotype was about 1,000 base pairs smaller than the usual deletion of RD207 in classical “Beijing” strains, which is consistent with an evolutionarily independent deletion event in the direct repeat (DR) region of MTBC.

Conclusions: We provide an example of convergent evolution in the DR locus of MTBC, and highlight the limitation of using spoligotypes for strain classification. Our results indicate that a proportion of “Beijing” strains may have been misclassified in the past. Markers that are more phylogenetically robust should be used when exploring strain-specific differences in experimental or clinical phenotypes.

6.2 Introduction

Mycobacterium tuberculosis complex (MTBC) adapted to humans consist of six main phylogeographical lineages**(Gagneux and Small, 2007)**. There is increasing evidence that strain diversity in MTBC plays a role in the outcome of infection and disease in tuberculosis (TB) **(Coscolla and Gagneux, 2010,Malik and Godfrey-Faussett, 2005)**. One particular MTBC genotype known as “Beijing” has repeatedly been associated with drug resistance**(Borrell and Gagneux, 2009)** and increased virulence in animal models**(Coscolla and Gagneux, 2010)**. This genotype was first described in 1995**(van Soolingen *et al.*, 1995)**, and has traditionally been defined based on a characteristic spoligotyping pattern **(Parwati *et al.*, 2010)**. More recently, phylogenetic analyses showed that the Beijing strain family belongs to Lineage 2 (known as East Asian lineage), which is one of the six main human-adapted lineages of MTBC **(Gagneux *et al.*, 2006,Comas *et al.*, 2009)**. Beijing strains are most often isolated in East- and Southeast Asia, in countries of the former Soviet Union, and have recently been emerging in South Africa **(Gagneux *et al.*, 2006,Parwati *et al.*, 2010,Cowley *et al.*, 2008,Tsolaki *et al.*, 2005)**.

Spoligotyping is based on the Clustered Regulatory Short Palindromic Repeats (CRISPR) region known as the Direct Repeat (DR) locus in MTBC. This region is characterized by series of direct repeats interspersed by short unique regions called “spacers” **(Supply *et al.*, 2006)**. The characteristic spoligotyping pattern of Beijing strains reflects the loss of the first 34 spacers of a total of 43 used in standard spoligotyping **(van Soolingen *et al.*, 1995)**. Repetitive DNA sequences like the DR locus exhibit a high rate of change, and convergent evolution can lead to identical genetic character states in phylogenetically unrelated strains; a phenomenon referred to as homoplasy **(Comas *et al.*, 2009)**.

We recently identified novel SNP markers that define the main phylogenetic lineages **(Hershberg *et al.*, 2008,Comas *et al.*, 2009)**. In contrast to spoligotyping, SNPs in MTBC exhibit almost no homoplasy **(Comas *et al.*, 2009)**. Here we used these SNPs, combined with genomic deletion and MIRU-VNTR analyses to revisit the phylogenetic classification of MTBC strains exhibiting the classical “Beijing” spoligotyping pattern.

6.3 Methods

MTBC isolates were obtained during an ongoing population-based study on the molecular epidemiology of TB in Switzerland, and from an ongoing hospital-based study in Nepal.

Mycobacterial isolates were cultured and DNA extracted according to standard laboratory procedures. Spoligotyping was performed as previously described and compared to data published in SpolDB4 **(Brudey *et al.*, 2006,Kremer *et al.*, 2004)**. 24-loci MIRU-VNTR was performed as previously described **(Supply *et al.*, 2006)**, and the data analyzed using the MIRU-VNTRplus online tool (http://www.miru-vntrplus.org). Determination of the main phylogenetic MTBC lineages was performed by TaqMan real-time PCR (Taqman, Applied Biosystems, USA) using primers (Sigma-Aldrich, Buchs, Switzerland), Taqman Universal MasterMix II and Taqman minor groove binder probes (Table 10) targeting lineage-specific SNPs reported previously**(Hershberg *et al.*, 2008,Cowley *et al.*, 2008)**. Region of difference (RD) deletion PCRs were performed for RD105, RD207 and RD750 **(Tsolaki *et al.*, 2004)**. PCR products of RD207 were directly sequenced. All genotyping experiments of the three “Pseudo-Beijing” isolates described here were repeated at least twice by two independent investigators.

Table 10: Sequence information of probes and primers used in this study to detect main phylogenetic lineages of M. tuberculosis complex isolates by single nucleotide polymorphisms genotyping.

Lineage	Alternative name	SNP name *	Primer sequences	Probe sequences
2	East Asian Lineage	Rv2952_0526n	F: 5‘-CCTTCGATGTTGTGCTCAATGT-3‘	Wild type probe: FAM: 5‘-CCCAGGA**G**GGTAC-3‘
			R: 5‘-CATGCGGCGATCTCATTGT-3‘	Lineage-specific probe: VIC: 5‘-CCCAGGA**A**GGTACT-3‘
3	Delhi/CAS	Rv3804c_0012s	F: 5‘-GCATGGATGCGTTGAGATGA-3‘	Lineage-specific probe: FAM: 5‘-AAGAATGCAGCTTGT**C**GA-3‘
			R: 5‘-CGAGTCGACGCGACATACC-3‘	Wild-type probe: VIC: 5‘-AAGAATGCAGCTTGT**T**GA-3‘

* as reported in Ref. [Comas *et al.*, 2009] F: forward; R: reverse; SNP, single nucleotide polymorphisms
Probes are minor groove binder probes

The Swiss study was approved by the ethics committee of the Canton of Berne, Switzerland. Written informed consent was obtained from the patient by the treating physicians. In some cases informed consent could not be obtained because the patient could not be located or was known to have died. For these cases we obtained permission from the Federal expert commission on confidentiality in medical research (based at the Federal Office of Public Health, Bern, Switzerland) to use the data provided by the treating physician based on clinical notes. The study in Nepal was approved by the Nepal Health Research Council (NHRC), Kathmandu, Nepal, and the ethics committee of the Canton of Basel, Switzerland. Written informed consent was obtained for all Nepalese patients.

6.4 Results

Among the isolates recovered in Switzerland, we identified a total of 52 that exhibited the characteristic “Beijing” spoligotype. SNP-genotyping confirmed that 50 of these (96.2%) belonged to Lineage 2. Two isolates (3.8%) belonged to Lineage 3 (also known as Delhi/CAS). Similarly, among 55 Nepalese isolates with a Beijing spoligotype, 54 (98.2%) were confirmed as belonging to Lineage 2, while one isolate (1.8%) belonged to Lineage 3 (Table 11). The three Lineage 3 isolates were epidemiologically unrelated and were isolated from HIV-negative patients. The two strains from Switzerland were isolated in patients living in Switzerland but originating from Portugal and Tibet, and the strain from Nepal was isolated in a patient born and living in Nepal. All but one strains were pan-susceptible (one isolate with a *katG* S315T mutation). We found no evidence for a mixed-strain infection or clonal heterogeneity **(Cohen *et al.*, 2011)**based on the 24-loci MIRU-VNTR pattern.

Table 11: Spoligotyping, single nucleotide polymorphism (SNP) and region of difference (RD) PCR results from the three Mycobacterium tuberculosis isolates belonging to Lineage 3

Specimen no.	Spoligotyping pattern (spacers 1-43)	SIT [1]	Clade	SNP for Lineage 2	SNP for Lineage 3	Deletion of RD 105	Deletion of RD 750
1395	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	1	Beijing	-	+	-	+
1476	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	1	Beijing	-	+	-	+
2503	□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□■■■■■■■■	1	Beijing	-	+	-	+

[1] Spoligotype International Type (SIT) according to the definition in SpolDB4 database using SITVIT2 (http://www.pasteur-guadeloupe.fr:8081/SITVITDemo/index.jsp)
+, present; -, absent; RD, region of difference;

To confirm our SNP data, we subjected these three isolates to further molecular investigations. Lineage 2 strains, including all Beijing strains, are deleted in RD105 **(Gagneux *et al.*, 2006,Tsolaki *et al.*, 2005)**, whereas Lineage 3 strains have a deletion in RD750 **(Gagneux *et al.*, 2006)**. We found that the three strains of interest had RD105 intact and a deletion in RD750 (Table 11). In addition, the MIRU-VNTR profiles of these isolates clustered with the Lineage 3 strains (Delhi/CAS) rather than with the Beijing strains (Figure 1). The further investigations thus confirmed that these three strains belonged to Lineage 3 rather than Lineage 2.

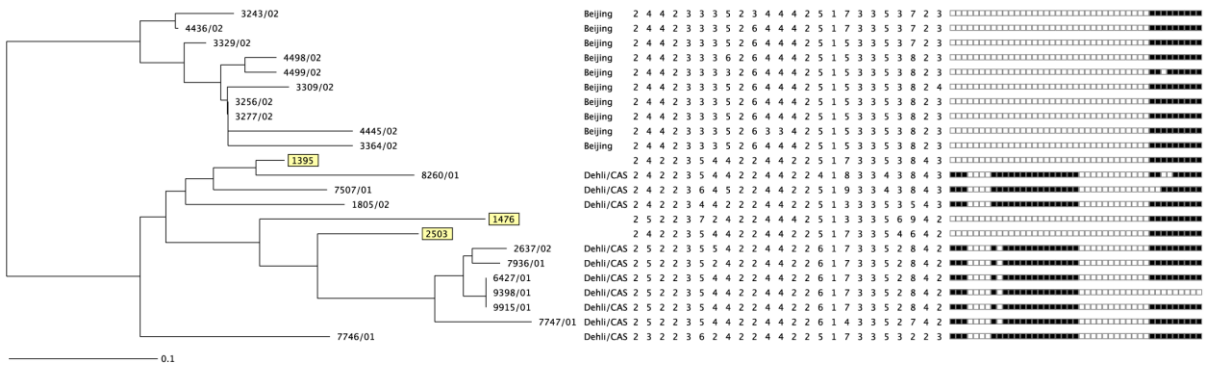

Figure 12: Neighbor-joining Dendogram based on the copy numbers of 24 MIRU-VNTR loci using the web-based MIRU-VNTRplus tool using the web-based MIRU-VNTRplus tool (http://mvpl.mlvaplus.net). The Mycobacterium tuberculosis isolates reported here are highlighted in boxes, reference strains are from the MIRU-VNTRplus database. The corresponding spoligotyping patterns are shown as a reference but were not included in the construction of the dendogram.

We explored the molecular mechanism by which these Lineage 3 isolates acquired their "Pseudo-Beijing" spoligotyping patterns. "True" Beijing strains harbor a 7,399 base pair (bp) deletion in RD207, which is responsible for the missing spacers 1-34**(Tsolaki *et al.*,**

2005). We thus hypothesized that the strains of interest might have acquired a similar but distinct deletion linked to an independent mutational event. We amplified RD207 **(Tsolaki *et al.*, 2004)** but failed to obtain a product in two of the three strains. Strain 1395 for which our PCR was successful yielded a PCR product that was about 1,000 bp larger than the corresponding product seen in true Beijing strains (Figure 13), indicating that the deletion responsible of the "Pseudo-Beijing" spoligotype in 1395 was about 1000 bp smaller than the classical deletion of RD207. Direct sequencing of the PCR product showed that the 3'-deletion boundary was 1,093 bp upstream (GenBank accession no. JF789456) of the deletion end point for RD207 published previously**(Tsolaki *et al.*, 2004)**. These results again indicate that the deletion in strain 1395 is distinct from the standard RD207 deletion. We were unable to determine the exact deletion starting point of this new deletion due to the repetitive nature of the DR region.

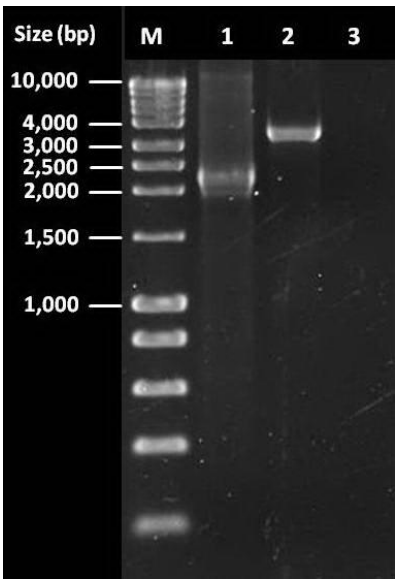

Figure 13: Results of Region of Difference (RD) 207 polymerase chain reaction
M, molecular weight marker; 1, "True" Beijing isolate; 2, "Pseudo-Beijing" isolate no. 1395; 3, Negative control.

6.5 Discussion

In this report we present three cases with MTBC isolates harboring spoligotype patterns identical to the Beijing signature, which were assigned to Lineage 3 (also known as Delhi/CAS) rather than to Lineage 2 (also known as East-Asian) based on three alternative molecular markers. We also provide evidence that these strains acquired independent deletion(s) in the DR locus of MTBC.

The fact that phylogenetically unrelated MTBC strains can harbor identical or very similar spoligotyping patterns has been observed before **(Flores *et al.*, 2007,Warren *et al.*, 2002)**. The DR locus of MTBC is highly variable, and convergent evolution can lead to homoplasy in spoligotyping patterns **(Cowley *et al.*, 2008)**. Even though strain classification based on spoligotyping will assign MTBC strains to the correct phylogenetic lineages in about 90% of the cases, some strains cannot be classified at all **(Kato-Maeda *et al.*, 2011)**, and others will be misclassified as shown here.

Misclassification of “Beijing” strains is particularly relevant given that this strain family has received increased attention over the last few years **(Parwati *et al.*, 2010)** [6]. In addition to their association with clinical drug resistance and hyper-virulence in animal models **(Coscolla and Gagneux, 2010)** [2,4], Beijing strains have been emerging in Cape Town, South Africa **(Cowley *et al.*, 2008,van der Spuy *et al.*, 2009)** [9,20], and the Canary islands **(Caminero *et al.*, 2001)**.

The data presented here suggest that a small fraction of strains traditionally referred to as “Beijing” strains might belong to another phylogenetic lineage. We stress that the prevalence of this phenomenon observed in our study of isolates from Switzerland and Nepal will not reflect the global *M. tuberculosis* genetic diversity. Of note, the three “Pseudo-Beijing” strains belonging to Lineage 3 were isolated from patients originating from three different countries. Given the mutational dynamics within the DR locus, and the

fact that strains harboring the classical Beijing spoligotyping pattern are rarely verified using independent molecular markers, it is possible that other so-called “Pseudo-Beijing” strains might turn out to belong to yet a different MTBC lineage. Further studies are needed to better define the global molecular epidemiology of these strains, including clinical phenotypes.

In conclusion, our case report illustrates an example of convergent evolution in the DR locus of MTBC, and highlights the limitation of using spoligotypes for genotypic classification. Markers that are more phylogenetically robust should be used when looking for associations between particular MTBC genotypes and experimental or epidemiological variables **(Coscolla and Gagneux, 2010)**. Future work should define the global prevalence and phenotypic characteristics of the “Pseudo-Beijing” strains.

6.6 Acknowledgements

We thank Bhawana Shrestha, director, and the laboratory staff at the German Nepal Tuberculosis Project, Kathmandu in Nepal for the contribution to this study.

Chapter 7: The genotyping and geographical analysis reveals random distribution of MTBC lineages in the disease endemic Kathmandu Valley.

This article with be submitted to: *PLoS One*

7.1 Abstract

The NTP, Nepal estimates for the nationwide prevalence of tuberculosis is high (238 per 100, 000) and at sub-national level or major cities, the risk of transmission could be higher due to crowding and recent population growth. In past 10 years, the in-migration of population from other parts of the country to Capital Kathmandu might have played a role in altering the prevalence of the disease. This gives an opportunity to portray the phylogeography of MTBC that represents genotypes from Kathmandu as well as the genotypes from many parts of the country. Understanding the nature of genetic diversity is critical for disease control programs. The city of residence of patients was used to identify the geographical location. For phylogenetic analysis, MTBC lineages were characterized by single nucleotide polymorphisms (SNPs) markers. The genetic diversity of the strains was further analyzed by the spoligotyping method. The phylogenetic data based on SNPs suggests that four major MTBC lineages were present among residents of Kathmandu and the non-residents with no significant difference. We found that the strains were randomly distributed across the geography of the Kathmandu which this study demonstrates as not different from strains circulating outside of Kathmandu. However, additional samples covering migrants from remote regions of Nepal may add to MTBC diversity yet to be discovered.

7.2 Introduction

The national data reports that the central development region including Kathmandu has the highest number of TB cases in Nepal. In the central development region, the case finding rate has decreased from 83% in the year 2009/10 to 76% in the year 2010/11. The current case finding rate (as registered at NTP) is 49% in Kathmandu **(Department of Health Service, 2011)**. This change in statistics indicates different putative factors including migration influencing the spread of infectious diseases in cities **(Wang and Wang, 2012)**. The urbanization and spur in population density has a significant impact on the prevalence and mortality due to respiratory diseases such as TB **(Barnes *et al.*, 2011)**. The human migration

in any form as temporary or permanent migration impacts on the local disease transmission and prevalence **(Ormerod, 1998)**. Thus, it is essential to understand the local distribution and type of MTBC strains circulating in Kathmandu. However, in order to understand the spectrum and burden of transmission of TB, the implementation of different genetic markers and epidemiological index case verification is essential **(Kato-Maeda *et al.*, 2011)**. The advantage would be to document the genotypes that may have disparities in virulence, in the likely hood of developing resistance to drugs, and in transmission. The Beijing family of strains is of particular interest, as experimental studies have demonstrated it as a virulent genotype. Moreover, it has been hypothesized that Beijing might show an advantage to spread over other genotypes where BCG vaccination is widely implemented through immunization program **(Parwati *et al.*, 2010,Malik and Godfrey-Faussett, 2005)**. We have previously shown the major MTBC genotypes that are circulating in Nepal using a subset of data that gave us the first insights into the diversity of MTBC in Nepal **(Malla *et al.*, 2012)** (Chapter 5).

In this chapter, we will focus on strain diversity of MTBC within the Kathmandu valley; we have attempted to use the patient residence as geographic information. The importance of using geographic information was to reveal homogeneity of strains if any, that exists among patients who share close geographical proximity inside the Kathmandu valley. This might contribute to the understanding of TB epidemiology in the Kathmandu valley and identify risk groups and risk areas within the city and allow for more effective control measures. The secondary objective of this work was to see what patient factors might be associated with bacterial genotypes.

7.3 Methods

Kathmandu, the capital of Nepal is the most densely populated city in Nepal with a population size of 1.7 million. The nearby large cities, Lalitpur, Bhaktapur, that are part of

Kathmandu valley also experienced population growth in the past decade with an estimated average household size of 3.7 to 4.4 **(Central Bureau of Statistics, 2012)**. The housing conditions in Kathmandu valley are poor **(The World Bank, 2012)**, and overcrowded with the population density per square kilometer of 2739 (Kathmandu), 877 (Lalitpur), 1895 (Bhaktapur).

TB patients were recruited at GENETUP which is located centrally in Kathmandu city, making it a primary health care unit for the surrounding population. There are 58 DOTS centers in Kathmandu city including GENETUP. The nearby districts such as Lalitpur and Bhaktpur have 23 and 20 centers, respectively, that are directly under supervision of the NTP. These DOTS centers refer patients to GENETUP for confirmation of diagnosis or for drug susceptibility tests as tertiary health care centers **(National Tuberculosis Programme, 2011)**. There are other private microscopy centers that suspected TB patients may seek health advice and diagnosis, and some are referred to GENETUP for confirmation. In short, GENETUP serves both as primary and tertiary health center for TB cases.

We prospectively studied a total of 410 pulmonary tuberculosis cases from 2009 to 2011 visiting GENETUP at Kathmandu. The patients enrolled were native Kathmandu residents as well as those who have migrated to Kathmandu at least 2 years before the date of their enrollment into the study. In other words, all patients analyzed in this chapter were residents of Kathmandu valley during the study period. This chapter focuses on defining strain genotype in relation to geography, and we analyzed patients who were culture positive (n=317). We analyzed whether there was any trend in geographic distribution of the four MTBC lineages that we found in our data set by using patient residence as a geographic variable. These data were then adjusted to other variables to compare with age and gender. We also attempted to find the geographic location of multi-drug resistance cases in our study population.

Geograpical analysis: The name of the residence of the patient was obtained from the questionnaire of demographic data. We then used the name of the city to retrieve the geographic locations by entering the name of the city into Google earth version 6.2.2.6613 (http://www.google.com/earth/index.html). The longitude and latitude in decimal degrees of the place of residence was then used in Quantum GIS v.1.8.0-Lisboa (http://www.qgis.org/) to explore the geographic link between MTBC genotypes. . The *.SHAPE file of the Kathmandu valley used as geospatial vector data was obtained from International Centre for Integrated Mountain Development (ICIMOD) (http://www.geoportal.icimod.org/Downloads/) after online registration.

7.4 Results

To assess the geographical clustering of MTBC, we used lineage data using the residences of 299 cases out of 317 culture positive cases. For the remaining 18 cases, the geographic data based on name of residence (within Kathmandu) could not be retrieved through Google earth. The total number of TB cases recruited by year and the proportion of MTBC lineages is shown in Figure 14. Similar to our previous findings based on a smaller samples (chapter 5), we found 4 main MTBC lineages while Lineage 3 and 2 being the most common with (132/317; 41.64%) and (92/317; 29.02%) respectively, followed by Lineage 4 and 1 (56/317; 17.66% and 37/317; 11.67%, respectively). We found a similar prevalence of MTBC lineages over the three year period from 2009 to 2011 (Pearson chi square test 3.22, p=0.78). Based on our geographical analyses, we saw that the MTBC lineages were spread randomly in the Kathmandu valley (Figure 15). The higher number of cases around GENETUP was probably due to the catchment area of GENETUP. The decline in cases from the regions far away from GENETUP did not affect our objective of identification of MTBC lineages/genotypes and their geo-spread in Kathmandu. The host factors such as age and gender distribution also represent the local spread of disease **(World Health Organization, 2008,Malla *et al.*, 2012)**.

The male patients represented 67.19% of total cases. The age distribution of TB cases showed that the age group of 15-25 years has the highest TB prevalence with 116 (36.59%) cases as shown in Table 12; this findings match with the national data **(National Tuberculosis Programme, 2011)**.

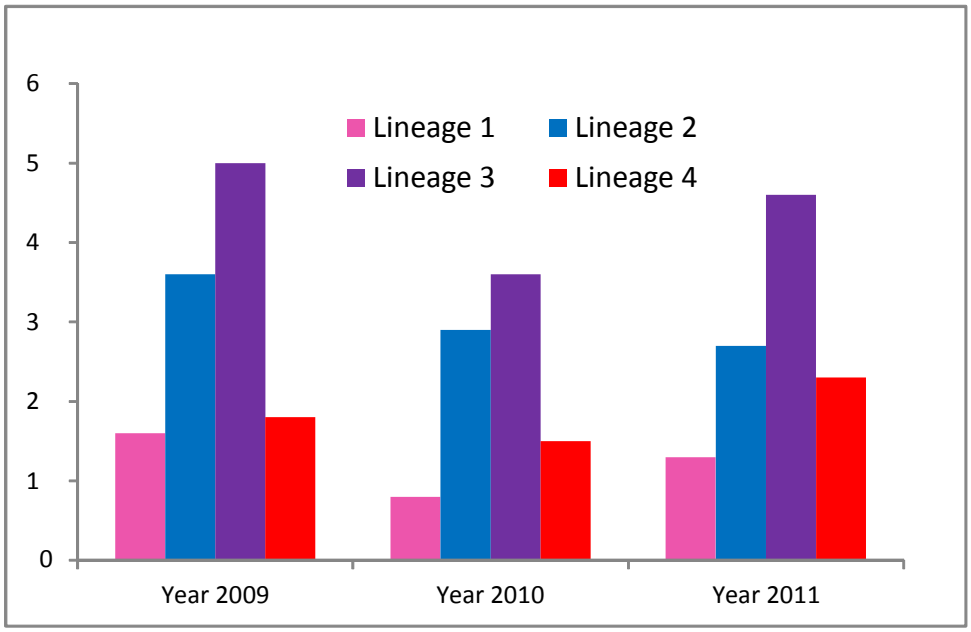

Figure 14: Study year and distribution (%) of lineages

Table 12: Patient age distribution compared to MTBC lineages

Lineages	Lineage 1	Lineage 2	Lineage 3	Lineage 4	Total	Percentage
Up to 24 yrs.	10	34	53	19	116	36.59
25-34	7	22	28	15	72	22.71
35-44	5	15	22	8	50	15.77
45-54	7	8	20	7	42	13.25
55-64	7	6	6	3	22	6.94
65 and above	1	7	3	4	15	4.73
Total	37	92	132	56	317	100.00

The spread of TB did not follow any geographic trend in Kathmandu. In other words, we found random distribution of the lineages in Kathmandu as supported by genotypes of circulating MTBC. As we previously discussed in chapter 5 **(Malla *et al.*, 2012)**, Lineage 2 and Lineage 3 were predominant lineages while Lineage 1, and Lineage 4 accounted each for less than 20% in the study. The latter two lineages also did not seem to be geographically clustered. As the SNP typing only allows defining the main phylogenetic lineages in given geographic settings **(Hershberg *et al.*, 2008)**, we used spoligotyping to further discriminate the lineages into sub-groups (Table 13). This method identified 23 distinct spoligotypes in the Kathmandu valley during the study period. Almost 64% (n=202) of the MTBC strains were either Beijing (n=91) or CAS family spoligotypes (n=111). The Beijing spoligotypes was highly uniform (as mostly SIT1 was found) and we did not perform any additional typing methods to discriminate further. However, the CAS family was represented by four different variants (Table 13). We linked the spoligopatterns with the available geographic location of the patients to explore possible geographical clustering (Figure 15) to (Figure 17). Again, we found that the CAS family was randomly distributed.

Table 13: Description of lineages and Spoligotypes of MTBC from Kathmandu (N=317)

Lineage	Spoligotyping families	N	Percentage
1	EAI6_BGD1	2	0.63
1	EAI1-SOM	4	1.26
1	EAI3-IND	5	1.58
1	EAI5	15	4.73
1	Unknown	2	0.63
2	BEIJING	91	28.71
2	Unknown	1	0.32
3	CAS-DELHI	2	0.63
3	CAS2	5	1.58
3	CAS	34	10.73
3	CAS1-DELHI	70	22.08
3	Unknown	1	0.32
4	AMBIGOUS :T3 T2	1	0.32
4	H1	1	0.32
4	H4	1	0.32
4	LAM	1	0.32
4	S	1	0.32
4	T2-T3	1	0.32
4	U(CAS_ANCESTOR)	1	0.32
4	X1	1	0.32
4	LAM9	2	0.63
4	T2	2	0.63
4	T3	2	0.63
4	T1	9	2.84
4	H3	12	3.79
--	ORPHAN	51	16.09
	TOTAL	**317**	**100.00**

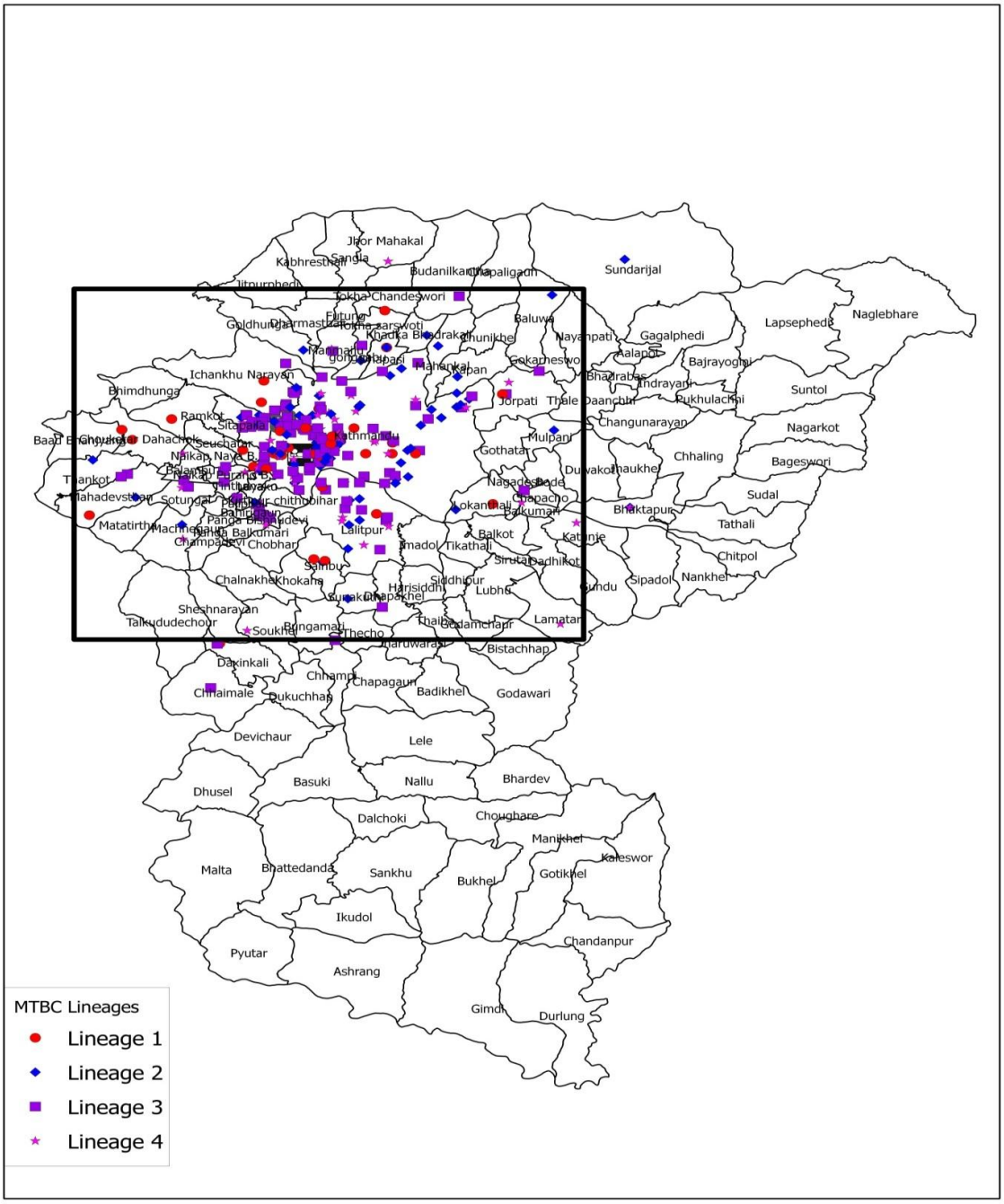

Figure 15: Geographic distribution of MTBC lineages across Kathmandu Valley

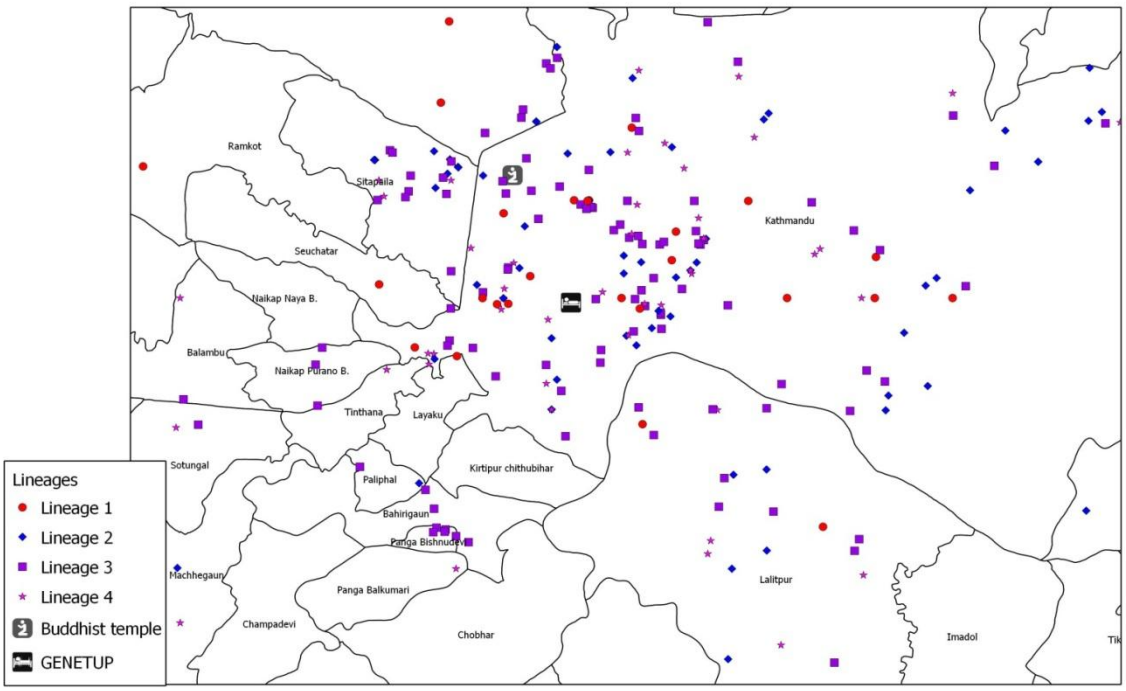

Figure 16: Geographic distribution of MTBC lineages across Kathmandu Valley (surrounding GENETUP)

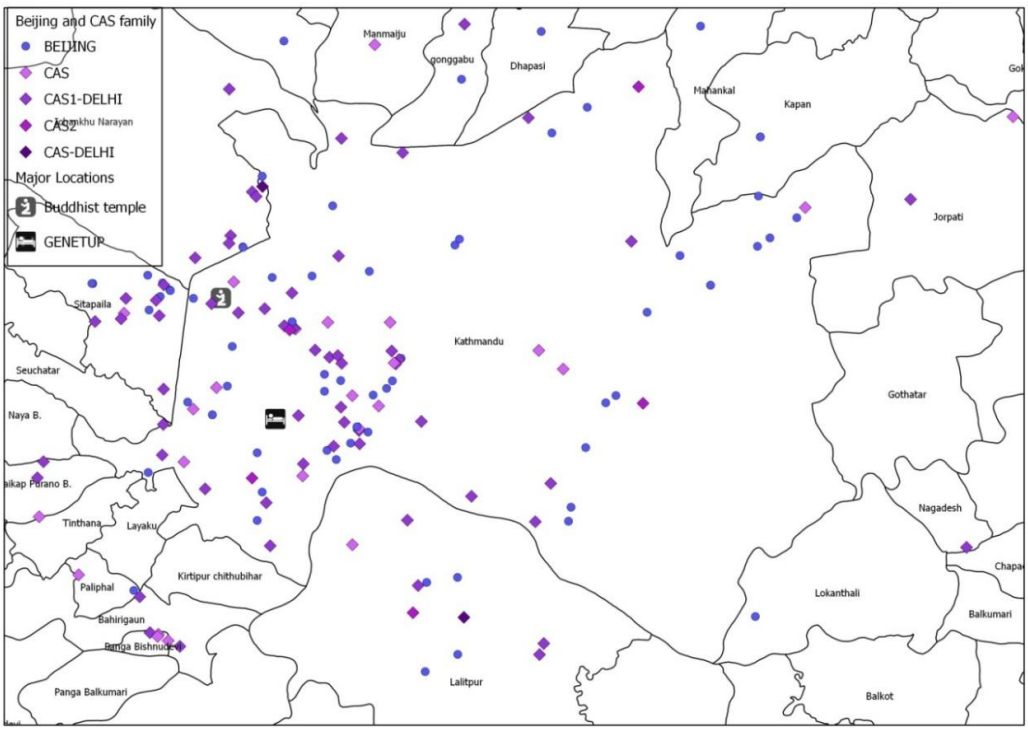

Figure 17: The Spread of Beijing and CAS family

Next, we compared the age and sex of the patients infected with the different MTBC lineages. We assumed that the movement of (economically active) male young aged population was higher than female of same age group in Nepal **(Central Bureau of Statistics, 2012)**. This was evident as we observed significant difference between the median age of male (33 years) and female cases (24 years) (p=0.04) in our sample set. In both sexes, the age (after 10 years categorization) was found to be not associated with MTBC lineages (chi2=24.70; p=0.26). Finally, we did not observe any geographical clustering of the seventeen MDR cases identified in Katmandu valley (for 2 strains the geographic information was not available).

7.5 Discussion

Our molecular data showed the presence of different genotypes of MTBC in Kathmandu. Additionally, we showed that the disease was particularly prevalent among the young population. Although the number of cases varied between different age groups, the proportion of lineages of MTBC was similarly represented in all age groups. These findings suggest that all four MTBC lineages have been circulating in Kathmandu since many years.

Because the exact GIS coordinates of households were unavailable, we used the city name to find the corresponding coordinates through Google maps. Two reasons can possibly support this approach in defining the geo location of TB cases and the corresponding MTBC genotypes. First, unmanaged urbanization has led to the construction of unplanned households that were not far apart. This leads to over-crowding within and between households. Second, Kathmandu is the most important disease endemic region, and when considering the above mentioned argument, may cause transmission not localized in the household or within certain perimeter of residence of the patient. In other words, the transmission in the community (e.g. in any social or public place; as opposed to within the

household) is as likely as transmission inside the households **(Marais *et al.*, 2009)**. Additionally, the TB patients have to travel daily to the nearby DOTS center for their daily intake of drugs, and because the use of protective masks (i.e. N95) by the patients is not well adhered to, disease can be actively transmitted during their travel **(Hubad and Lapanje, 2012,Dharmadhikari *et al.*, 2012)**.

We observed the significant age difference between the male and female patients, which led us to categorize age into groups of 10-years interval. The similar proportion of MTBC lineages throughout the age groups implies that patients of all age groups were equally likely to harbor MTBC strains of any lineages. Irrespective of lineage, the male and female patients of the age group 0-24 years were nearly equal, while the female representation was comparatively lower in other age groups. The age and gender distribution of the cases showed that nearly 36% (n=116) of the patients fall in the age group of 15-24 years, and the number of cases decreased as age increased.

Studies have shown that the incidence of TB in older age is due to reactivation of infection that might have happened long time ago **(Kamper-Jorgensen *et al.*, 2012)**. However, in our study, the discrepancy in age and sex was not significant as far as the relative proportion of the four MTBC lineages is concerned. The proportion of Lineage 3 and Lineage 2 were higher in all age group Table 12. In other words, the bias towards the younger age was not influenced by MTBC genotype in Kathmandu. This can be interpreted as the geographic distribution of MTBC genotypes in Kathmandu has not been influenced by possible internal migrations, even when we consider that the younger aged population are migrating the most **(Long *et al.*, 2002)**. This then raises the question as to whether there were differences in distribution of the MTBC genotypes among non-resident Kathmanduties in comparison to the resident Kathmanduties. First, our statistical analysis comparing these two patient populations showed no differences in MTBC lineage distribution (Chi square =12; p=0.21). Second, the

migrants do not settle in any specific area of the Kathmandu valley but were largely spread inside Kathmandu. This suggests that the common genotypes in Nepal overall are more or less the equally distributed. Provided that the population migration to Kathmandu was from all region/geography of Nepal, the MTBC strain diversity in Kathmandu might in fact represent the national strain diversity.

We observed MDR strains in all four lineages, albeit in different proportions. This could be due to at least two factors. First, the MDR prevalence represents the genotype proportions indicating that the bacterial genetic background as well as the prevalent genotypes, in given geographical setting, might be influencing the emergence of drug resistance **(Fenner *et al.*, 2012)**. Similar to our findings, another study in a refugee camp found no geo-spatial association in MDR cases that were infected with strains forming the same genotypic cluster **(Oeltmann *et al.*, 2008)**. Second, MDR strains did not show geo-clustering, which is in parallel to our findings on the uneven distribution of MTBC genotypes in Kathmandu. This arguably supports the hypothesis of ongoing community transmission of MDR strains not restricted to direct household contacts. This information can be used to identify the (individual) risk factors to drug resistance such as differential adherence to treatment or acquisition of drug resistance due to primary resistance **(Lin *et al.*, 2011)** and possibly provide underlying mechanisms responsible for increased MDR transmission **(Manjourides *et al.*, 2012)**.

7.6 Limitations

Our results have several limitations as this study was conducted in one health center in Kathmandu. The limited sample size represented mostly by the patients from nearby areas of study site may not fully represent the epidemiology of TB from Kathmandu. As human migration is a dynamic process, the establishment of new strains in a community may occur at different time intervals, dependent on environmental and clinical factors. However, given

the heretofore non-existent documentation of MTBC genotypes and associated geo-information in Nepal, our data provides an outline of the contemporary MTBC strains from Kathmandu. Our hypothesis was that the MTBC strain diversity in Kathmandu remained same even after the urban-drift needs to be confirmed through detailed geographic and molecular epidemiological studies. Due to lower number of cases, the MDR transmission hotspots could not be identified, although we found many MDR cases near GENETUP.

7.8 Conclusion

We found no obvious pattern in geo-distribution of MTBC genotypes in Kathmandu. The dominance of Lineage 2 and Lineage 3 could mean transmission of local strains or less frequent import of strains in Kathmandu from other regions. Alternatively, the lineage distribution of MTBC in Nepal could be same all across the country and the migration of cases to Kathmandu does not alter the distribution of the proportion of MTBC lineages inside Kathmandu. Our study supports the previous findings that the young population represents the major risk group for TB. Targeted epidemiological studies among this age group may reveal the geo-epidemiological distribution of TB in Kathmandu.

Chapter 8: Are some molecular mechanisms of drug resistance preferred by certain MTBC lineages?

This article with be submitted to:

International Journal of Tuberculosis and Lung Disease (IJTLD)

8.1 Abstract

The drug resistance surveys conducted by the National Tuberculosis Program (NTP), Nepal found increasing numbers of drug resistance among TB cases. This has raised concerns on the effectiveness of disease control in Nepal. The previous exposure to anti-TB drugs increases the risk of drug resistance, and early detection of drug resistance among new TB cases allows for more effective treatment. However, for increasing the sensitivity and effectiveness of rapid molecular tests that detect drug resistance, molecular characterization of genes conferring resistance to anti-TB drugs is important. Additionally, different MTBC genotypes have been discussed as a potential factor leading to the heterogeneity of drug resistance mutations, which in turn could influence the disease prognosis and treatment outcomes. To characterize mutations causing rifampicin and isoniazid resistance, 506 pulmonary TB patients with or without previous history of TB were investigated at a TB reference laboratory at Kathmandu. By DNA sequencing of the targeted hotspots regions for drug resistance in the *rpoB, katG*, and *inhA* genes, we found 101/506 (19.96%) strains resistant to either one or both drugs. A mutation in *rpoB* was detected in 68/506 (13.43%) strains with codon position S531L being most the frequent (46/68; 67.64%). Seventy-six out of 506 (15.21%) strains were isoniazid resistant and 62/76 (81.57%) had a mutation at codon position S315T of *katG*. Moreover, 17/76 (22.36%) strains had a mutation in the *inhA* promotor region. We found odds of association of Lineage 4 (Euro-American lineage) to *katG* (S315T) (OR, 3.06; 95% CI, 0.94 to 9.98; p=0.38) higher compared to other lineages. Our results showed different drug resistance mutations for selected genes which seem to be favored by genetic variation of MTBC. We suggest that the variability in association needs further assessment with a larger sample size. The implementation of molecular tools as routine diagnostic procedure is advisable for identification of various mutations and strain genotyping that could identify probable drug resistant bacteria.

8.2 Introduction

The World Health Organization report published in 2011 showed varying worldwide estimates on drug resistance TB burden that ranged from 3.7% to 20% in new or previously treated cases, respectively **(World Health Organization, 2012)**. Since 1996, when the DOTS era started with implementation of standardized treatment regimen based on the 4 drugs isoniazid, rifampicin, pyrazinamide and ethambutol, TB treatment has been highly successful in producing a cure rate of more than 80% in Nepal **(National Tuberculosis Programme, 2011)**. On the other hand, the emergence and spread of drug resistance has started in parallel since around the same time. Drug resistant TB is an emerging challenge to national TB control programmes, with 2.9% MDR among new TB cases and 11.7% among pre-treated cases in Nepal **(World Health Organization, 2012)**.

MDR occurs spontaneously after exposure to anti-TB drugs where poor patient adherence, mono-therapy, inappropriate treatment occur, and through direct transmission of already drug resistant strains, which is referred to as "primary drug resistance". In recent years, conventional phenotypic drug resistance testing is slowly being replaced by more rapid molecular tools **(Bodmer and Strohle, 2012,Ignatyeva *et al.*, 2012)**. These tools interrogate mutational hotspots in *inhA, katG,* and *rpoB,* which are the genes in which most of the resistance conferring mutations occur in clinical strains resistant to isoniazid (*inhA* and *katG*) or rifampicin (*rpoB*) **(Zhang and Yew, 2009,Hillemann *et al.*, 2007)**.

A longer duration of treatment can also be a determining factor for the acquisition of drug resistance in bacterial populations. Studies have shown that multiple mutations can emerge during the standard treatment period of 6-8 months **(Rinder *et al.*, 2001,Mariam *et al.*, 2011)**. In some TB patients, not only the accumulation of mutations, but the co-existence of different mutants were found **(Mariam *et al.*, 2011)**. This means that either the same strain harbors multiple mutations, or different strains carry different mutations within single

patients with multiple infections. Alternatively, the genetic diversity of clinical MTBC strains can influences the emergence of resistance **(Koser *et al.*, 2012)**. Drug resistance mutations often cause fitness defects in absence of the drug, but low or no-cost mutations may also occur **(Comas *et al.*, 2012)**. Moreover, the strain genetic background can influence the specific Minimum Inhibitory Concentration of a given resistance conferring mutation **(Borrell and Gagneux, 2009,Fenner *et al.*, 2012)**.

Despite the rising number of drug resistant TB cases in Nepal, limited data is available on the molecular mechanisms of drug resistance. We conducted this study to obtain that information and analyze it in context of the variable strain genetic background defined as the main MTBC lineages circulating in Nepal.

8.3 Materials and methods

The MTBC strains isolated from the sputum samples provided by new and follow-up TB patients during the study period (*details are given in the General Materials and Methods chapter*). Briefly, the phenotypic DST for isoniazid, rifampicin, streptomycin, ethambutol was performed in a selected number of cases as requested by treating physician or in suspicion of drug resistance. The *inhA* promoter region and partial sequences of *katG* (*Rv1908c*) and *rpoB* (*Rv0667*) were sequenced in all strains to confirm the molecular mechanisms of drug resistance as well as to document the mutational hotspots.

8.4 Results

This study was based on laboratory based molecular epidemiological study in Nepal. A total of 650 smear positive TB cases were recruited and clinical characteristics obtained using a standardized questionnaire. Only 506 cases were culture positive for MTBC. The phenotypic drug susceptibility test (DST) results were only available for 70/506 (13.83%) cases. This was particularly due to NTP guidelines, as in Nepal DST is only performed if requested by physician in suspicion of drug resistance or if the suspected case was an MDR contact. The

DNA sequencing of the three drug resistance loci was performed in all of the 506 strains. We found that based on this DNA sequence-based DST, 101/506 (19.96%) strains showed resistance to either isoniazid or rifampicin only (mono-resistance) or both (MDR) Table 14.

Table 14: Number of resistance patterns based on DNA sequencing

Drug resistance	resistance (n)	%
INH mono-resistant	33	32.67
RIF mono-resistant	25	24.75
Multi drug resistant	43	42.57
Total	101	100.00

The isoniazid resistance was the most common drug resistance observed in our sample set Table 15. Among the 76/506 (15.01%) isoniazid resistant strains, mutations in *katG* were most common (59/76; 77.63%) with polymorphisms exclusively found at codon position *katG* 315 with amino acid substitution from S to T. Two of these strains with *katG* S315T mutation had another non-synonymous mutation (AGC 632 AAC S 211 N) not documented as drug resistance conferring polymorphisms. An *inhA* promotor mutation was found in 14/76 (18.42%) strains. Three strains (3/76; 3.94%) had both *katG* and *inhA* promotor mutations. The phenotypic DST report of isoniazid was available for 70 samples that were performed at concentration level (INH; 0.2ug/ml) in LJ slants in duplicates. Of the 52/70 (74.28%) resistant strains, 40/52 (76.92%) were genotypically confirmed, while another 12 (23.08%) resistant strains had no mutation in the *inhA* promotor or in the region of *katG* that was sequenced. On the other hand, 16/70 (22.85%) strains found susceptible by the phenotypic method showed resistance mutations suggesting these strains are in fact resistant to isoniazid. In these 16 strains, the proportion of *inhA* promotor mutation (8/16; 50 %) and *katG* mutation (8/16; 50 %) were the same. One possibility for the discrepancies between the phenotypic DST and DNA sequencing results, at least for the strains harboring an *inhA* promotor mutation, is that these strains have a low level of resistance and therefore failed to grow at the

drug concentration used for phenotypic testing (0.2ug/ml for isoniazid) **(Zhang and Yew, 2009)**.

Table 15: Spectrum of mutations obtained by DNA sequencing for isoniazid resistance conferring genes

Isoniazid resistance conferring mutations in *inhA* and *katG* genes	n	%
inhA -8	1	1.32
inhA -8 + *katG* S315T	2	2.63
inhA -15	13	17.11
inhA -15 + *katG* S315T	1	1.32
katG S315T	59	77.63
Total	76	

AGC 315 ACC S TO T [*&]
* 2 HAVE ADDITIONAL MUTATIONS
[&] AGC 315 ACC S to T; plus; AGC 210 AAC S to N

Of the 70 strains for which phenotypic DST data was available, 44 strains showed phenotypic drug resistance against rifampicin (RIF; 40ug/ml). Sequencing of the whole rifampicin resistance determining region (RRDR) of the *rpoB* gene was performed for all 506 cultures – positive MTBC strains. Overall, we found a good correlation between the phenotypic and genotypic resistance data, with 43/44 (97.73%) strains that showed phenotypic resistant to rifampicin harboring a mutations in RRDR Table 16. Overall, an RRDR mutation was detected in 68/506 (13.43%) strains at codon position 531 which was also the most frequent (46/68; 67.64%), followed by codon 516 (14/68; 20.59%). Mutations at codon position 511, 513, and 526 of *rpoB* accounted for less than 6% among rifampicin resistant strains. Except for one MDR strain which had multiple mutations in RRDR (CTG 511 GTG; L to V and GAC 516 TAC; D to Y), all other strains had a single mutation. Four strains had discrepant results showing the TCG 531 TTG S to L mutation, while being phenotypically susceptible. Several novel non-synonymous mutations within and outside of RRDR were also observed Table 16. Based on DNA sequencing data only, 43/101 (42.57%) strains were identified as MDR.

Table 16: Spectrum of mutations obtained by DNA sequencing for rifampicin rifampicin resistance conferring mutations in *rpoB* gene

Nucleotide change	amino acid change	n	%
CTG 511 GTG	L TO V	1	1.47
CTG 511 CCG	L TO P	1	1.47
CAA 513 AAA	Q TO K	1	1.47
CAA 513 CTA	Q TO L	1	1.47
GAC 516 GTC	D TO V	8	11.76
GAC 516 TAC	D TO Y	2	2.94
GAC 516 TTC	D TO F	4	5.88
CAC 526 CGC	H TO R	1	1.47
CAC 526 CTC	H TO L	2	2.94
CAC 526 GGC	H TO G	1	1.47
TCG 531 TGG	S TO W	3	4.41
TCG 531 TTG	S TO L	43	63.24
Total		68	

Additional mutations	n
(1609 to 1611) CAG to AAG Q 537 K	2
522 TCG to TTG Ser to Leu	1
TTG 524 TGG L to W; plus; ACC 525 CCC T to P; plus; C*ACA*AG 526 527 ACA IS DELETED	1
ATG 515 TTG M to L	1
CTG 533 CCG L to P	1
ATC 491 TTC I toF; plus; GCG 559 GTG A to V	1

We next compared the distribution of isoniazid resistance mutations as a function of the drug susceptibility profile. Specifically, we found that the *katG* S315T mutation was strongly associated with MDR (OR, 10.76; 95% CI, 3.7 to 31.35; P = 0.000; Table 17), occurring in 37/42 (88.10%) MDR strains. This finding is in agreement with previous findings **(Hazbon *et al.*, 2006)**.

Table 17: Mutational spectrum of preferential isoniazid resistance conferring mutations in MDR (N=42)

	katG S315T		*inhA* -15		*inhA* -8	
MDR[#]	MUT n (%)	WT n (%)	MUT n (%)	WT n (%)	MUT n (%)	WT n (%)
YES (n=42)	37 (88.10)	5 (11.90)	5 (11.90)	37 (88.10)	0 (0)	42 (100)
NO (n=56)	22 (39.39)	34 (60.71)	8 (14.29)	48 (85.71)	1 (1.79)	55 (98.21)

[#] 1 MDR with multiple mutation (both in *inhA* -8 plus *katG* 315) was excluded to reduce preferential bias.

To explore the influence of the strain genetic background on the mutational trajectory towards drug resistance, we genotyped each of the 101 drug resistant strains using SNP

markers that define the four major MTBC lineages circulating in Nepal **(Stucki *et al.*, 2012)**. The distribution of the main drug resistance mutations across the four MTBC lineages are shown in Table 18 and Table 19. Overall, the S531L was the most frequent mutation in RRDR, irrespective of lineage. Similarly, the *katG* S315T was most frequent mutation in isoniazid resistant strains, irrespective of lineage, but was somehow underrepresented in Lineage 2, particularly in comparison to Lineage 3 and 4, suggesting that other mutations (e.g. in other parts of *katG* that have not been sequenced here) might be important in the emergence of isoniazid resistance in Lineage 2 as suggested previously **(Gagneux *et al.*, 2006)**.The *inhA* -8/-15 promotor mutations were overall relatively rare (less than 20%) in all the four lineages Table 19.

Table 18: RIF-resistance conferring mutation codon position in RRDR region

Lineage	n	RIF resistant (n)	CP-511	CP-513	CP-516	CP-526	CP-531	*TOTAL*
lineage 1	7	*5*	*0*	*0*	*0*	*0*	*5*	*5*
lineage 2	45	*34*	*0*	*2*	*4*	*2*	*26*	*34*
lineage 3	28	*19*	*1*	*0*	*8*	*2*	*8*	*19*
lineage 4	21	*10*	*1*	*0*	*2*	*0*	*7*	*10*
	101	*68*	*2*	*2*	*14*	*4*	*46*	*68*
CP= Codon position								

Table 19: INH-resistance conferring mutation in *katG* and *inhA* promotor region

Lineages	n	INH resistant	*katG* -S315T	%	*inhA* promotor -15/-8	%
lineage 1	7	*5*	*4*	*80.00*	*1*	*14.29*
lineage 2	45	*29*	*22*	*75.86*	*8*	*17.78*
lineage 3	28	*24*	*20*	*83.33*	*4*	*14.29*
lineage 4	21	*18*	*16*	*88.89*	*4*	*19.05*
Total	101	*76*	*62*		*17*	

8.5 Discussion

Here, we focused on the mutations in the resistance-determining regions of genetic targets of isoniazid and rifampicin, as these mutations are most commonly found in strains resistant to first line anti-TB drugs and are epidemiologically most relevant. We show the extent and

association of different drug resistance mutations and their frequencies in different genotypes of MTBC strains in Nepal. These mutations can have serious clinical impact on patient and may lead to further resistance if not properly managed.

By analyzing the clinical and molecular data, we found that the drug resistance appeared more among patients with previous treatment history. For isoniazid and rifampicin resistance strains, the mutations conferring high level drug resistance were more frequent. Similar reports have been described among pulmonary tuberculosis cases in South Asia based on DNA sequencing and molecular probe assays **(Yadav *et al.*, 2013,Siddiqi *et al.*, 2002)**.

Altogether, we found 101/506 strains with any resistance to isoniazid and rifampicin (INH mono-resistant: 33; RIF mono-resistant: 25; MDR: 43). The prevalence of mutations at *katG* codon position 315 with amino acid change (S to T) was the most common (Table 15) which is considered as the mutation conferring high level of drug resistance and transmission **(van Soolingen *et al.*, 2000)**. The reduced virulence of INH resistant strains seen in mouse models depends on the type of mutation in *katG* gene **(Pym *et al.*, 2001,Pym *et al.*, 2002)**, and this is reflected in the primary transmission of isoniazid resistant strains. For example, several studies have shown that whiles some resistance mutations are associated with reduced transmission, the *katG* S315T is associated with successful transmission **(van Soolingen *et al.*, 1999)**.

Similar to previous reports, we found that strains with *katG* S315T mutations were associated with multi-drug resistance **(Marttila *et al.*, 1998)**. The higher prevalence of *katG* S315T mutation and its association with MDR as we observed in this study could arguably indicate that the resistance was acquired during prolonged treatment. Among the MDR cases, the proportion of getting *katG* S315T with *rpoB* S531N and *rpoB* D516V (the two most frequent mutations as in Table 20) was similar for new and previously treated cases. This raises a

question if strains harboring this particular set of mutations were more fit to persist in drug containing environment and also to transmit more efficiently than strains with other combination of mutations.

Table 20: Frequency of RIF-mutations in relation to INH-mutations among MDR strains (n=43)

Group	Mutations	L 511 N	Q 513 N	D 516 N	H 526 N	S 531 N	multiple L 511 V PLUS D 516 Y	Total (n)
katG	S 315 T	2	1	**10**	4	**20**	1	38
katG	*S 315 T plus inhA* prom -8; T to G	0	0	1	0	0	0	1
*inhA*pro	*inhA* prom -15; C TO T	0	1	2	0	2	0	5

In addition to differential acquisition of mutations and persistent combination of mutation patterns in different research settings and treatment regimens in place, studies have shown the impact of strain genetic background on emergence and fixation of drug resistance mutations **(Gagneux *et al.,* 2006,Sun *et al.,* 2012)**. The isoniazid resistance due to *katG* S315T was found associated with Euro-American Lineage (Lineage 4) and another lineage (Lineage 1) has showed predisposition to *inhA* promotor mutations **(Fenner *et al.,* 2012)**. The bacterial lineages are different depending on specific geography as we have shown in our earlier publication the variability of lineages in this study settings **(Malla *et al.,* 2012)**.

Other studies have either failed to support the association of specific drug resistance mutations with bacterial genotype **(Lavender *et al.,* 2005)**, or proposed that pre-existing isoniazid resistance induce or select the rifampicin resistance codon *in vitro*, irrespective of bacterial genotypes **(Bergval *et al.,* 2012)**. The *in vitro* study showed the shift towards acquisition of *rpoB*-S531L as the fit mutation, and suggested this kind of mutations will accumulate in subsequent bacterial population. In our clinical sample of 43 MDR strains, of the 37 strains with a *katG* S315T mutation, the proportion of *rpoB*-531 mutation was highest (20/37; 54.05%) followed by *rpoB*-516 (10/37; 27.02%).

We performed DNA sequencing of only part of *katG* (*Rv1908c*) covering 541 base pairs (bp) covering the most frequent mutation region (H37*Rv* position; 2154957 to 2154974). The *katG* gene is 2223 bp long, and other mutations in other regions of *katG* have been shown to cause isoniazid resistance **(Sandgren *et al.*, 2009,Sherman *et al.*, 1996)**. This could be one of the reasons for the inconsistent phenotypic and genotypic drug resistance test results seen in our sample-set. Moreover, the method of sampling also has significant impact on clinical and drug resistance patterns **(Piatek *et al.*, 2000)**. Hence, population based sampling or sampling from different laboratories, including more extensive DNA sequencing need to be performed in Nepal to complement our results.

8.6 Conclusion

In conclusion, our data suggest that MDR strains with a combination of *katG* S315T with *rpoB* S531L are selected in clinical settings and might propagate more efficiently. Our study did not support the hypothesis of genotype preference to specific drug resistance mutations. Further research is required to verify if this is true in a larger sample set considering other possible confounders such as geographic settings, co-morbidities. The TB control program should be vigilant and consider implementation of molecular tools for drug resistance identification and strain genotyping as a routine diagnostic tool.

Chapter 9: Molecular characterization of extremely drug resistant tuberculosis from Nepal

This article with be submitted to:

Tuberculosis

9.1 Abstract

Since the first report of extensively drug-resistant tuberculosis (XDR-TB) in 2008 in Nepal, the number of such cases is continuously rising. The fact that XDR-TB emerges due to poor management of directly observed treatment short course (DOTS)-plus treatment among multi-drug resistant TB cases is well accepted and is generally considered as an instance of acquired resistance. Although XDR cases are detected more among multi-drug resistant (MDR) cases after failure of prolonged treatment, primary resistance of XDR-TB has rarely been described. Here, we report four cases of XDR-TB; two of these were presumably acquired XDR during treatment, while the other two reported no more than three months of first line anti-TB treatment history and no TB contacts. Hence, the latter two cases could be primary XDR cases. All four XDR strains belonged to Lineage 2 and the Beijing (SIT 1) spoligotype, and MIRU-VNTR typing showed that the two putative primary XDR cases showed identical 24-loci profiles, supporting an epidemiological link. The four XDR-TB strains showed different drug resistance profiles and mutations in *rpoB* (S531L), *katG* (S315T), *gyrA* (D94G, D94N), and *rrs* (C1402A, C1402T, G1484T). These factors and young age (median age 21 years) of the patients indicate the evidence of transmission of XDR-TB in community and raise concerns on future TB epidemiology. This data revealed emergence of both acquired and primary resistance of XDR cases that may be contributing in community transmission of nearly untreatable form of TB, posing an emerging threat to TB control program of Nepal.

9.2 Introduction

WHO declared TB a global emergency in 1993, and in 1996, implemented the multi-drug therapy widely known as directly observed short course therapy (DOTS) which is a 6-8 months long treatment **(Grange and Zumla, 2002)**. In Nepal, the nationwide institutional coverage of DOTS programme was achieved only in 2001, with prime objective of diagnosis and treatment of all the TB cases. The drug resistance surveys that were first conducted in

1996 and in later years showed fluctuating numbers (prevalence) of MDR cases, which made it difficult to interpret the true prevalence of drug resistance in Nepal **(National Tuberculosis Programme, 2011)**. MDR-TB can be effectively treated with second-line drugs for the duration of 16-18 months, and latest data available from the year 2008/09 in Nepal showed a cure rate of 71.2%, while defaulter and failure rate constitute 13.2 and 8.2%, respectively **(National Tuberculosis Programme, 2011)**. Exposure to second-line anti-TB drugs among MDR defaulter and failure cases augments the risk for the emergence of extremely drug resistance (XDR) with the estimated prevalence rate ranging from 5% to 25% among MDR cases in different countries **(World Health Organization, 2008)**. XDR-TB has become a global threat to TB control programmes worldwide. Seventy-seven countries both from developed and developing countries had reported at least one case of XDR-TB by October 2011 **(World Health Organization, 2011)**. Especially the increase of MDR-TB numbers (defined as resistant to at least isoniazid and rifampicin) in high TB burden countries is a growing public health problem, as MDR-TB cases have to be treated with second-line anti-TB drugs. These drugs are more expensive, less effective and often associated with sever side-effects. In Nepal, nearly 3% of MDR cases are new TB cases, while data regarding primary XDR cases are un-available **(World Health Organization, 2012)**. As drug-resistant TB cases are not kept in isolation at health facilities, and provided there is widespread use of second-line drugs, it is possible that not only MDR but XDR-TB transmission is also on-going in the community. XDR-TB is defined as an MDR strain which is additionally resistant to fluoroquinolones, and at least one of the injectable second-line drugs **(World Health Organization, 2010)**. In Nepal, a study conducted in 2009 showed an XDR prevalence of 5% among MDR-TB cases **(National Tuberculosis Programme, 2011)**. The WHO report first published in 2007 showed 41 countries with XDR-TB cases, while the first XDR-TB case in

Nepal was documented in 2008, and by July 2011, 27 XDR cases were under treatment through NTP **(National Tuberculosis Programme, 2011)**.

Previous studies have reported primary transmission of XDR-TB among close contacts as well as in the community **(Leung *et al.*, 2012)**. This means TB cases with pre-treatment history and currently under treatment are not the only risk groups for XDR-TB. On the one hand, due to resource limitations and difficulty in performing second-line drug susceptibility testing (DST), screening of XDR is performed mostly among chronic TB cases or MDR failures, leaving many drug resistant cases undiagnosed. On the other hand, unknown drug resistant status will render current treatment ineffective, which will further worsen the situation by not being able to stop transmission **(Udwadia, 2008)**. Only few studies have reported primary transmission of XDR in the community. The molecular genotyping methods used to study TB outbreaks are key in such investigations.

The concerns about primary XDR transmission are particularly valid in low resource settings such as Nepal. The hidden burden of XDR-TB will be difficult to reveal unless the second line drug resistance testing is routinely available and used, which is economically and technically demanding. Our objective was to detect XDR-TB strains form patients with or without previous treatment history by a hospital based survey.

9.3 Methods

The study was carried out at German Nepal Tuberculosis Project (GENETUP), a tuberculosis reference laboratory in Kathmandu, Nepal which is the only laboratory for performing first-line and second-line drug susceptibility tests in Nepal. The laboratory is certified by the Supranational Reference Laboratory, "Kuratorium Tuberkulose in der Welt e.V." in Gauting, Germany **(World Health Organization, 2012)**. GENETUP is also a DOTS and DOTS-plus treatment center.

Nepal is a relatively small country and the numbers of TB and MDR-TB cases vary among regions and districts. The Central Development region (CDR) including capital city Kathmandu is the most densely populated region with internal migration of population from other regions of the country, and has the highest number of TB cases (**Department of Health Service**, **2011**). In this study, the majority (62.6%; 317/506) of patients enrolled was from Kathmandu, and others were referred to GENETUP from DOTS-center, and microscopy centers situated in various districts other regions of Nepal. This study analyzed 506 strains (one per TB case) prospectively collected from culture confirmed TB cases during 2009 to 2011 at GENETUP. The patient participation was voluntary. After written informed consent was obtained, the demographic and clinical data including geography, previous and current treatments, and TB contacts was collected by treating physician and trained medical staffs. In order to determine the prevalence of XDR-TB, we characterized the resistance to second line drugs among detected MDR-TB cases. The data was entered in password protected Microsoft® Access database.

The phenotypic DST was performed at GENETUP for isoniazid (INH; 0.2ug/ml), rifampicin (RIF; 40.0ug/ml), streptomycin (STR; 4.0 ug/ml), and ethambutol (EMB; 2.0 ug/ml) by proportion method on Lowenstein Jensen Slants in duplicates. Strains resistant to isoniazid and rifampicin, and strains from patients referred as suspected for XDR-TB were subjected to drug susceptibility test for second-line drugs; Ofloxacin, Capreomycin, and Kanamycin by proportion method **(World Health Organization, 2008**, **2001)**. The resistance was recorded when 1% of growth was observed in drug containing Lowenstein Jensen slants compared proportionally to drug-free Lowenstein Jensen slants. Molecular drug resistance testing was performed on all strains by direct sequencing of the hotspot regions of the target genes for rifampicin (*rpoB*), isoniazid (*inhA* promoter region and *katG*), and streptomycin (*rpsL*). MDR strains were then further sequenced and analyzed for ethambutol (*embB*), fluoroquinolones

(*gyrA*) and aminoglycoside (*rrs*) resistance by sequencing of the relevant gene segments as shown in Appendix 1 **(Malla *et al.*, 2012)**. The sequences were analyzed with *M. tuberculosis* H37Rv as reference sequence using the Staden software package**(Staden *et al.*, 2000)**.

SNP typing was performed as described previously **(Malla *et al.*, 2012)**. Spoligotyping was performed using commercially available kits from Isogene Bioscience BV (Maarssen, The Netherlands) following the manufacturer's instructions **(Kamerbeek *et al.*, 1997)**. The spoligotype patterns were classified according to the online SITVITWEB database (http://www.pasteur-guadeloupe.fr:8081/SITVIT_ONLINE/) accessed on October 22, 2012. The MIRU-VNTR typing of four XDR-TB Beijing strains was done at Genoscreen, France (www.genoscreen.com) following standard protocols. The MIRU-VNTR analysis was done using standard 24 loci plus 4 hypervariable loci markers **(Supply *et al.*, 2006)**. The additional 4 hypervariable loci (VNTR 1982, VNTR 3232, VNTR 3820, VNTR 4120) typing were conducted for improved clonal distinction between closely related Beijing strains allowing more discrimination **(Gopaul *et al.*, 2006)** (**Allix-Béguec et al., based on email correspondence to Genoscreen**).

This study was ethically approved by the Nepal Health Research Council, Nepal and by the Ethical Review Board (EKBB) of the Canton of Basel-Stadt, Switzerland. All patients provided written informed consent. During informed consent, all patients were informed and agreed about use of strains for research purposes.

9.4 Results

Patient characteristics: Among all the 506 patients, four (0.8%) XDR-TB cases were identified. Of the total 45 MDR-TB cases, these 4 (8.9%) XDR cases were resistant to both fluoroquinolones and aminoglycosides as summarized in Table 22. The MDR cases (excluding four XDR cases) were four folds more frequent among previously treated cases

(15.5%; 27/174) than in new TB cases (4.2%; 14/332). Patients treated for 28 days or more were grouped as previously treated cases **(World Health Organization, 2010)**. The median age of XDR-TB was 21 years and three were males (Table 21).

Patients (ID no. 2010326 and 2010408) were failure cases of previous tuberculosis treatment (more than a year), so presumably non-compliance has contributed to the acquired XDR-TB status. Remarkably, two (ID no. 200964 and 2010310) of the four XDR cases were under intensive phase of category 1 treatment regimen (2HRZE/4HR) of the NTP, Nepal, suggesting these cases might show primary XDR-TB. Both patients failed to show sputum conversion during follow-up visit of 2-3 months after the start of treatment.

Drug resistance: The phenotypic DST results for first line drugs as well as second line drugs are shown in Table 22. The drug resistance mutations were identified by DNA sequence analysis and are presented in the same table. The XDR strains exhibited different drug resistance mutations except in case of rifampicin (*rpoB* S531L). The most frequent mutation associated with isoniazid resistance was S315T in *katG* gene while in two XDR strains (ID no. 2010326 and 2010408) we were unable to identify any mutation despite being phenotypically resistant to isoniazid. The mutation in *inhA* region that is associated with low level resistance to isoniazid was absent in all strains **(Banerjee *et al.*, 1994)**. Hence these strains might harbour a mutation in another region of *katG* (which was not explored in this study) **(Kelley *et al.*, 1997)**. Except for the strain ID no.2010408 for which phenotypic DST was not performed, all other strains were found resistant to fluoroquinolones and aminoglycosides when tested phenotypically. We then analyzed the *gyrA* and *rrs* gene to confirm the phenotypic XDR status. The sequencing of the *gyrA* gene revealed that the codon position 94 (D94G, D94N) was the most frequent mutation site conferring resistance to fluoroquinolones in all four XDR strains. The strain ID NO. 2010326 showed *rrs* mutation (G1484T) which is associated with resistance to kanamycin, amikacin, viomycin,

capreomycin and hence consistent with the treatment history of patient as a case of MDR failure. The remaining three XDR strains showed different mutations in *rrs*. Moreover, these XDR strains showed varied resistance to other anti-TB drugs as shown in Table 22; two strains (ID no. 2010310 and 2010326) were resistant to all first line drugs and second line drugs tested.

The acquisition of additional mutations to mitigate the loss of fitness through compensatory mutations is observed in many organisms. Moreover, in clinical strains the fitness loss can be compensated by secondary mutations (i.e. so called compensatory mutations) **(Gagneux *et al.*, 2006,Borrell and Gagneux, 2009)**. The *rpoC* and *rpoA* genes undergo mutations to compensate the fitness lost due to the mutations occurred in *rpoB.* Drug resistance mutations often cause fitness defects in absence of the drug, but low or no-cost mutations can occur **(Comas *et al.*, 2012)**. This finding is also supported by our study, as three out of all four XDR strains had an *rpoC* mutation. We found high probability compensatory mutations (HCMs) *rpoC* V483G along with others except for one (ID 2010310). The *rpoC* (*Rv0668*) is 3951 bp long sequence which we sequenced only a fraction of. The ID 2010310 might have compensatory mutation in another region of *rpoC.*

Geographic distribution and genotyping: The XDR-TB patients were living in Kathmandu at the time of study, which might suggest an epidemiological link, although the time of initial infection was unknown. Based on the patient questionnaires, the patients' acquaintance was not known to each other, although patient ID no. 2010408 had one TB contact. The patient ID no. 200964 was permanent resident of Solukhambu and ID no. 2010310 was a permanent resident of Dolakha, both in eastern region of Nepal. Both patients migrated to Kathmandu, one ten years and the other one year before the interview, respectively.

All four strains were Lineage 2 by SNP-typing and SIT1/Beijing genotype as defined by spoligotyping. The Beijing genotype has been found associated with drug resistance and is particularly frequent in Asia **(Kamper-Jorgensen *et al.*, 2012)**. Studies from Russia **(Casali *et al.*, 2012)** and India **(Ajbani *et al.*, 2011)** showed that more than 60% of XDR strains belonged to the Beijing genotype. We further analyzed the clustering of these cases, MIRU-VNTR typing identified three unique patterns with one cluster; strains ID no. 200964 and 2010408 shared a common profile with identical alleles at all of the 24 loci screened Figure 18. The four hypervariable loci (in rectangular box) were also identical for these two strains.

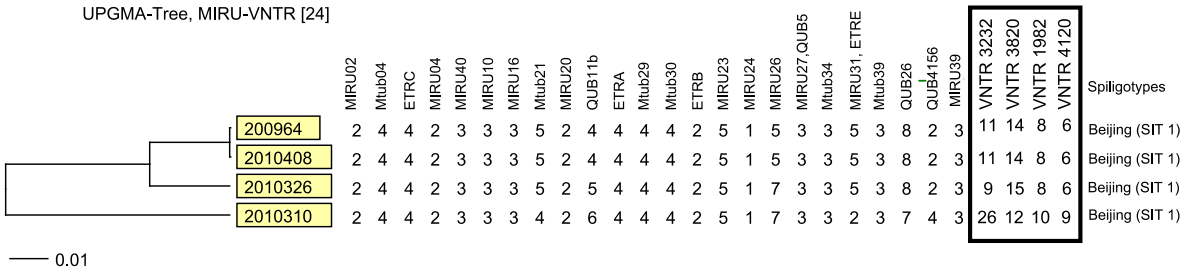

Figure 18: Dendrogram of MIRU-VNTR typing data of four XDR strains from Nepal

Table 21: Epidemiology and clinical characteristics of XDR-TB cases

Patient ID	year	Permanent Address	Year of Migration	Age/Sex	TB close contact in past 2 years	WHO case category	Sputum Microscopy Grading	BCG immunization
200964	2009	SOLUKHUMBU	more than 10 years	15/M	no	3 months follow up for first line drugs (Category I)	3+	yes
2010408	2010	KATHMANDU	no migration history	16/M	yes (Friend)	6 months follow up for first line plus streptomycin treatment (Category II)	3+	yes
2010310	2010	DOLAKHA	1 year before	21/M	no	2 months follow up for first line drugs (Category I)	3+	no
2010326	2010	KATHMANDU	no migration history	21/F	UN	MDR treatment failure case	3+	unknown

Patient treatment status at time of enrollment in study.

1.ID 200964.Sputum was examined at 3 months of Cat I,and found MDRTB,then switch on to Second Line drugs as per the rules & regulations of NTP.But unfortunately he was persistent positive and SLDST found XDR-TB.

2. ID 2010408. Sputum was examined at the follow-up month of 6 of Cat II treatment, and was found XDR-TB.

3.ID 2010310. Sputum was examined at 2 months of Cat I,and was found as MDR-TB.then switched on to second line drugs.

4.ID 2010326.this case has attended our laboratory as known XDR-TB.This was a case of MDR-TB treatment failure.

X. HIV status was unknown for all XDR-TB cases

Table 22: Phenotypic and genotypic drug resistance characterization of XDR strains

Patient ID	Lineage	Spoligotype	*katG*	*inhA*	*rpoB*	*rpsL*	*pncA*	*embB*	Phenotypic resistance to 1st line drugs	*gyrA*	*rrs*	Phenotypic resistance to 2nd line drugs
200964	Lineage 2	BEIJING (SIT 1)	AGC 315 ACC S TO T	WT	TCG 531 TTG S TO L	AAG 128 AGG K 43 R	WT	WT	HRSE	GAC 94 GGC D to G	1402 C/A*	Ofx, CM, KM
2010408	Lineage 2	BEIJING (SIT 1)	NA	WT	TCG 531 TTG S TO L	WT	WT	CAT 935 CGT H 312 R	HRSE	GAC 94 GGC D TO G	1402 C/T*	UN
2010310	Lineage 2	BEIJING (SIT 1)	AGC 315 ACC S TO T	WT	TCG 531 TTG S TO L	AAG 128 AGG K 43 R	TTG 545 TCG L 182 S	CAT 935 CGT H 312 R	HRSE	GAC 94 GGC D TO G	1402 C/T*	Ofx, CM
2010326	Lineage 2	BEIJING (SIT 1)	NA	WT	TCG 531 TTG S TO L	AAG 263 AGG K TO R	TCG 199 CCG S 67 P	CAT 935 CGT H 312 R	HRSE	GAC 94 AAC D TO N	1484 G/T**	Ofx, CM, KM

**rrs* associated resistance against kanamycin, amikacin, viomycin

***rrs* associated resistance against kanamycin, amikacin, viomycin, capreomycin

NA=NOT DONE

UN= Unknown

9.5 Discussion

Our findings suggest both the acquisition of XDR-TB during MDR-TB treatment as well as primary transmission of XDR-TB in the community. Three major findings in this paper are of particular interest. The first is that all the XDR TB cases were of young age (range 15-21 years). Young age in TB patients can be considered as a proxy for ongoing transmission. A study based on genotyping in Denmark showed higher number of clustered cases in young aged group compared to old aged population **(Kamper-Jorgensen *et al.*, 2012)**. This presumably implies reactivation of old infection among old aged population, different from young aged population that formed on-going transmission clusters.

Secondly, except for the case ID no. 2010326, the other XDR-TB cases visited health facility with the first episode of TB and reported to never have received any second-line drugs to which resistance was observed. In other words, patient ID no. 2010326 was treated more than 20 months and was a known MDR failure case, thus the multiple mutational events might have occurred resulting in accumulation of drug resistance mutation. The strain from this XDR-TB case showed resistance to other anti-TB drugs as well, including streptomycin, ethambutol, and pyrazinamide. In particular, the mutation at nucleotide position G 1484 T in *rrs* gene that confers higher level of resistance to kanamycin, amikacin, viomycin, and capreomycin was found only in this case, which suggests a long treatment history **(Maus *et al.*, 2005) (Long *et al.*, 2012)**. Patients ID no. 200964 and 2010310 received no more than three months of treatment with first line drugs. XDR-TB does not respond to the standard six-month treatment with first-line anti-TB drugs; this generally leads to treatment failure **(World Health Organization, 2010)**. The common sings of TB such as coughing and chest pain remained persistent in all four XDR cases. Contact investigations through patient interviews did not disclose any potential epidemiologic link between these XDR-TB patients.

In addition to patient characteristics, the third factor is strain genetic background that is constantly argued to be linked to the drug resistance problem **(Niemann *et al.,* 2009)**. We found that all the XDR strains belonged to the Beijing genotype which has been associated with drug resistance. The limited number of XDR-TB patients does not allow us to draw any statistical inference, however our finding are consistent with the view that Beijing might be more prone to develop drug resistance, including XDR **(Comas *et al.,* 2012,Fenner *et al.,* 2012,Lee *et al.,* 2012)**.

The geographic location of the XDR-TB patients and the non-identical mutations in drug resistance genes indicate that the strains might in fact not be epidemiologically linked. In Nepal, there is no movement restrictions for MDR- or XDR-TB patients and such patients are treated as outpatients in most of the DOTS clinics. These patients have to travel daily to the DOTS centers to take the drugs in presence of health staffs. Hence, these patients might become a source of infection given that the drug resistant cases remain infectious for longer despite under treatment **(Borrell and Gagneux, 2009)**. A study from South Africa emphasized the multi-factorial reasons for transmission of XDR-TB, that include the delay in detection, inadequate treatment, and social factors such as crowding households, and socio-economic factors **(Basu and Galvani, 2008)**. In Nepal, the establishment of hostels for some MDR cases may be beneficial for the containment of drug-resistant cases. Still, the undiagnosed cases might be moving around in community posing a risk of transmission **(National Tuberculosis Programme, 2011)**.

The 24-locus MIRU-VNTR plus additional 4 loci (highlighted in rectangle in Figure 18) analysis formed three different patterns and showed that the strains ID no. 200964 and 2010408 were genetically clustered, suggesting that these cases were epidemiologically linked. However, the drug resistant profiles (both phenotypic and genotypic) were not identical. Moreover, the year of isolation of the two strains was one year apart. So, it can be

argued that both of these patients were infected by another (unknown) XDR-TB case at different time intervals or Case ID 200964 transmitted to Case ID 2010408 after being infected from another XDR-TB case (Figure 19). The other two XDR strains had different MIRU-VNTR types, indicating that the cases were unlinked.

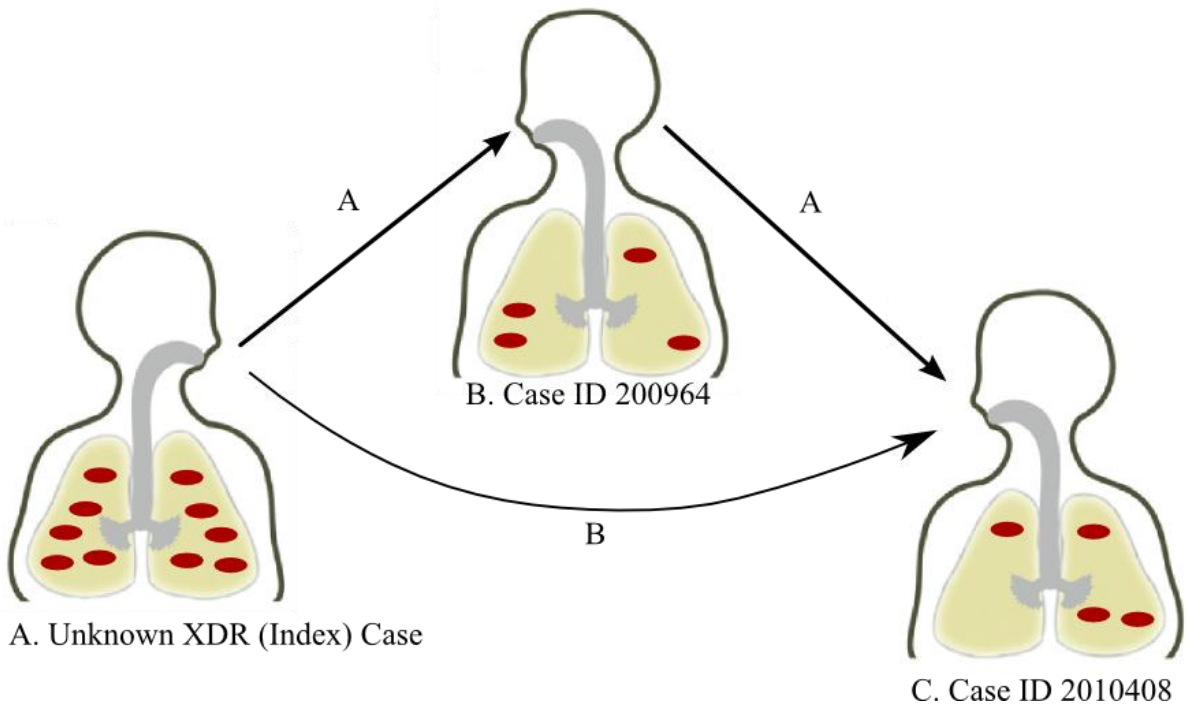

Figure 19: Proposed chain of transmission dynamics of XDR-TB in Case ID 200964 and Case ID 2010408

We proposed two routes of transmission; A- Unknown (Index) case first transmitted to Case ID 200964 and this case then transmitted to Case ID 2010408. B- Unknown (Index) case transmitted the XDR-TB to both cases. The first four digits of case ID denote year of isolation of strains.

Following current guidelines of National Tuberculosis Program of Nepal, the second line DST tests are not performed among new TB cases due to technical and resource limitations **(World Health Organization, 2001,Kim *et al.*, 2004)**. We may argue that the current NTP strategy for drug screening among MDR failure cases is not sufficient for TB case management, as indicated by the probable community transmission of XDR-TB reported here. A study from KwaZulu-Natal, South Africa reported emergence of XDR cases through exogenous reinfection as confirmed by genotyping of follow-up strains **(Andrews *et al.*,**

2008). This emphasizes the urgent need of detection of drug resistance status right at the beginning of treatment (or during follow-up visits if the patient is consistently sputum smear positive or sputum conversion is delayed). In other words, the NTP strategy should highlight the need to measure the extent of XDR-TB and not only consider MDR-TB cases as risk groups. The hidden burden of XDR-TB will be difficult to uncover unless routine second-line drug resistance test is standardized, and made widely available **(Ignatyeva *et al.*, 2012)**.

The MDR-TB cases are treatable and case management is generally successful **(National Tuberculosis Programme, 2011)**. However, XDR-TB is virtually untreatable, particularly in low-resource country such as Nepal **(Migliori *et al.*, 2012)**. Incorrect treatment may lead to worsening of the disease, and disease transmission. XDR-TB has been shown to transmit in other settings and our data suggest that it might transmit in Nepal as well **(Leung *et al.*, 2012)**.

In summary, the latest annual report from NTP, Nepal showed increasing number of XDR cases with a documented 27 XDR-TB cases under treatment by July 2011 **(National Tuberculosis Programme, 2011)**. Our study provides evidence of four XDR-TB cases which we believe is a good representation (14.8%; 4/27) to interpret the emergence and likely ongoing transmission of this severe form of TB.

9.5 Limitations

The limitation of this study is a possible sampling bias, which could be due to availability of second line DST in the reference laboratory, Kathmandu only. This might cause the geographic bias in representing the prevalence of XDR-TB. Moreover, the patient participation was voluntary, so this study does not represent all the patients who visited the reference laboratory during the study period. In addition, the low sampling fraction and the short study period may cause difficulty in defining or finding clustered cases **(Glynn *et al.*,**

1999,Houben and Glynn, 2009). One possible reason for fluoroquinolones resistance could be the use of the drugs against other diseases by these patients. The fluoroquinolones are broad spectrum antibiotics and are used empirically for treatment of other diseases **(Kakinuma *et al.*, 2012,Ginsburg *et al.*, 2003)**. The recent history of use of fluoroquinolones for any diseases by the patients was unavailable.

9.6 Conclusion

The geographical distribution of patients, variable drug resistance mutations, and MIRU-VNTR data of XDR-TB strains in this study suggest that the transmission of XDR-TB was not caused by one single outbreak of one specific Beijing (SIT 1) strain alone, but highlights likely on-going community transmission of XDR-TB in Nepal. As the number of MDR cases are increasing, both among new and pretreated cases, so are the chances of detecting XDR-TB. We provided examples showing that the treatment failure cases were not the only group at risk of becoming XDR. Our study emphasizes the necessity of drug resistance screening as routine tests at treatment onset, irrespective of treatment history. We hope that our findings will help policy makers to adopt necessary strategies to diagnose and contain the XDR-TB cases and prevent community transmission. NTP program should take serious steps and initiate the action for timely diagnosis of severe forms of TB in Nepal and globally. Future studies covering many patients from diverse region of the country will help reveal the true burden of XDR-TB in Nepal.

Chapter 10: Discussion and Conclusion

The details of this research project are discussed in the respective chapters. This chapter focuses on the general discussion of the overall project followed by the suggestions and conclusions.

In Nepal, NTP is in charge of performing prevalence surveys of TB, compiling and reporting clinical and epidemiological data, and providing patient management, all of which are core components of TB control. The main objective of this thesis was to foster the understanding of the epidemiology of TB in Nepal through novel research approaches using molecular tools. The use of molecular methods proved also to be useful in investigating the MTBC diversity circulating in Nepal compared to other geographic regions. In addition, we collected the demographic, clinical and epidemiological data from TB cases, their treatment histories, and phenotypic drug resistance data (when available) using semi-structured questionnaires. We adopted a passive case detection approach, as only one laboratory in Nepal (i.e. GENETUP) is currently certified by supranational reference laboratory to perform culture and phenotypic DST. This has resulted in a convenience sample, which is not necessarily representative of the whole of Nepal. Nevertheless, many of the cases included came from other districts (some bordering India and Tibet), allowing us to study the diversity of MTBC in these geographic areas as well. Some of the most remote areas of Nepal were only represented by a few strains. Nevertheless, this is the first report of MTBC genotypes from these areas. Another limitations of our study was that because TB has a long incubation time and DOTS treatment is 6 months long, follow-up of patients from start of diagnosis to treatment outcome was not possible due to time and logistics constraints. Moreover, because referred patients were also enrolled in this study, follow-up of all patients and retrospectively trace their contacts was not possible.

The aim of this research was to genotype the MTBC strains circulating in Nepal. The appropriateness of the genotyping tools depends on many factors. For phylogeographic analysis, the SNPs based genotyping tools are appropriate as these markers have slower molecular clocks and are therefore more stable. Using SNP-based genotyping and classification, we found four of the six main lineages of the MTBC represented in Nepal. For contact tracing, genotyping techniques with a higher discriminatory power are required **(Mathema *et al.*, 2008,Allix-Beguec *et al.*, 2008,Gutacker *et al.*, 2002,Kato-Maeda *et al.*, 2011)**. Following the replacement of IS6*110* RFLP by a combination of spoligotyping and MIRU-VNTR, the latter two methods have now become the new gold standard for molecular epidemiological investigation of TB. Some investigator use spoligotyping to identify the main phylogenetic lineages of the MTBC. In theory, different genotyping methods should identify the same groupings. However, because both spoligotyping and MIRU-VNTR are based on repetitive elements that are prone to convergent evolution generating homoplasies, relying only on these methods for phylogenetic strain classification can be misleading **(Comas *et al.*, 2009)**. This problem is illustrated by our report of the "Pseudo-Beijing" spoligotype (Chapter 6). Moreover, the limitations of the spoligotyping nomenclature also necessitate the use of other techniques. For example, we found 78 (15.22%) spoligotypes that were "orphan", which did not allow defining a specific genotype. Additionally, the discriminatory power of particular genotyping technique can also depend on the particular strain. For example, using spoligotyping to analyze the genetic diversity among Beijing strains will limit the discriminatory power because spacers 1-34 are inherently deleted **(Ferdinand *et al.*, 2004)**. In regions like Nepal, where nearly one third of MTBC strains belong to the Beijing family (Chapter 4), the use of spoligotyping can be used as a starting point, but other genotyping methods need to be used to explore diversity within the family and to differentiate "true" Beijing strains from "Pseudo-Beijing" strains.

The most common MTBC lineage observed in our study was Lineage 3 which includes the CAS family spoligotype. In terms of different spoligotypes, the CAS family was also more diverse compared to the Beijing family of strains, which was the second most frequent lineage in Nepal. Why is the CAS family spoligotype more diverse than the Beijing family? As discussed earlier, one explanation could be the lower number of spacers present in Beijing, which limits further discrimination unless other tools are used. Another reason could be that the spread of Beijing strain was more confined (or clonal) compared to the CAS family. If this is correct, then the risk factors and risk groups for the spread of these two genotypes could be different and could depend on several factors including host variables. In the future, investigation into the details of why one genotype is more frequent than another should be performed, as this may uncover some of the factors responsible for the spread of TB in Nepal and elsewhere.

Traditionally, molecular epidemiology of TB investigates the transmission of strains from one patient to another, as well as the progressive change in drug resistance within the same host **(de Jong *et al.*, 2008)**. Moreover, molecular epidemiological investigation allows differentiating between recent transmission, reactivation or relapse of previous infections, which is of particular importance for TB control programs. In contrast to most of the developed world, in places where TB is highly endemic, contact tracing by conventional processes will be difficult. In Nepal, we found Beijing strains related to pre-treatment history when adjusting for drug resistance (Chapter 6). This suggests that the exposure to drugs might have selected for one genotype over another. Molecular epidemiology can also help identifying ongoing outbreaks of TB. Despite Nepal being TB endemic country, TB outbreaks have never been documented. This is presumably due to lack of reference genotyping data and/or tools. Now, our findings filled some of this data gap and NTP in

Nepal should consider establishing genotyping facilities for MTBC, as it may also contribute to innovative control strategies.

The results from the DNA sequencing of drug resistance genes provided an overview of the mutations and mutational hotspots for resistance to first-line and second-line drugs in Nepal. These hotspots are similar to the ones seen globally; however some differences in the relative proportions were seen. We documented several additional mutations that need further verification to confirm they indeed cause drug resistance. Another approach would be to document the frequency of those mutations in a larger sample sets. From our findings we can confirm that at least for the diagnosis of MDR-TB based on the currently available molecular assays including the HAIN test and Xpert MTB/RIF will provide adequate sensitivity and specificity. Drug resistance is an emerging problem in Nepal and our findings show that it is represented mostly among previously treated cases or those infected with the Beijing genotype. An additional finding of our work is the association of female patients with Beijing genotype (Chapter 5). The NTP report of Nepal for the year 2010/11shows that 2/3 of the TB cases were male. Some Beijing strains have been associated with higher virulence in animal models and drug resistance **(de Steenwinkel *et al.*, 2012,Aguilar *et al.*, 2010)**. Hence based on our findings, female TB patients in Nepal seem to be associated with "virulent" strains of TB and drug resistance. Future research on host factors may shed lights on the host-pathogen interaction in TB and identify the reasons behind the apparent higher susceptibly of females to "virulent" TB.

Finally, we provided some evidence that XDR-TB in Nepal occurs both as transmitted (i.e. primary) and acquired (i.e. secondary) drug resistance. All four XDR-TB strains we identified were Beijing strains, which was consistent with our findings on the general association between Beijing strains and any drug resistance. Overall, Beijing represented about 30% of the samples in our study. Hence from a bacteriological point of view, up to one

third of the TB patients in Nepal will be at risk of drug resistance. Investigating the risks associated with drug resistance among new as well as pre-treated cases will be crucial to better understand the basis of the ongoing epidemics of drug-resistant TB.

Assessing the socio-economic impact of tuberculosis in Nepal

TB has been considered as a disease of poverty and affects the economically disadvantaged group of society **(Oxlade and Murray, 2012)**. Despite the fact that DOTS treatment is provided free to the patients registered to NTP, the direct financial burden to the individual and the family is substantial. The lengthy treatment (at least 6-8 months) can render patients economically inactive, causing great economic loss to both the individual and the family. Addressing the socio-economic factors has contributed to the reduction of TB and has improved treatment adherence **(Kamineni *et al.*, 2012)**. A study from China has shown that the poor quality of DOTS, treatment adherence, and a previous TB episode and poverty were associated with drug resistance **(Zhao *et al.*, 2012)**. In our study, the socio-economic factors were also addressed. For example, we have shown that TB affects young adults who are in their economically most productive age (Chapter 4).

Many studies from Africa and India have assessed the association between poverty and TB, **(Oxlade and Murray, 2012)**, but few publications are available on similar assessment focusing on the direct financial burden of TB diagnosis and treatment for individuals and their household. In order to assess the costs that the patients incur through the processes of diagnosis and treatment in Nepal, we collected additional data during our study. Generally, we found that patients could not continue with their jobs due to weakness or social stigma. The impact of TB was the largest on patients whose occupation was “daily wage earners”. The indirect cost of TB was also high, suggesting that a policy shift from “Health worker supervised DOTS” to “Home-based DOTS” or “ambulatory DOTS” could be a cost-effective

option in Nepal **(Malla *et al.*, 2009,Prado *et al.*, 2011)**. The data analysis on direct, indirect cost expenditure to the patients is still ongoing. This out of pocket expenses as stratified by economic status of patients will be used to measure the "treatment outcome" of TB. Our hypothesis is to test if the risk of treatment failure or relapse or defaulter is significantly high among economically disadvantaged group. The findings will be submitted to International Journal of Tuberculosis and Lung Disease.

References.

1. World Health Organization (2012) GLOBAL TUBERCULOSIS REPORT 2012. Geneva, Switzerland: World Helath Organization. ISBN 978 92 4 156450 2 ISBN 978 92 4 156450 2. 282 p.
2. Aaron L, Saadoun D, Calatroni I, Launay O, Memain N*, et al.* (2004) Tuberculosis in HIV-infected patients: a comprehensive review. *Clin Microbiol Infect* 10: 388-398. doi:10.1111/j.1469-0691.2004.00758.x
3. Migliori GB, De Iaco G, Besozzi G, Centis R, Cirillo DM (2007) First tuberculosis cases in Italy resistant to all tested drugs. *Euro Surveill* 12: E070517 070511.
4. Velayati AA, Masjedi MR, Farnia P, Tabarsi P, Ghanavi J*, et al.* (2009) Emergence of new forms of totally drug-resistant tuberculosis bacilli: super extensively drug-resistant tuberculosis or totally drug-resistant strains in iran. *Chest* 136: 420-425. doi:10.1378/chest.08-2427
5. Udwadia ZF, Amale RA, Ajbani KK, Rodrigues C (2012) Totally drug-resistant tuberculosis in India. *Clin Infect Dis* 54: 579-581. doi:10.1093/cid/cir889
6. World Health Organization (2008) Policy guidance on drug-susceptibility testing (DST) of second-line antituberculosis drugs. Geneva, Switzerland. WHO/HTM/TB/2008.392 WHO/HTM/TB/2008.392. 20 p.
7. World Health Organization (2008) Guidelines for the programmatic management of drug-resistant tuberculosis: Emergency Update. Geneva, Switzerland: WHO. 272 p.
8. Gutierrez MC, Brisse S, Brosch R, Fabre M, Omais B*, et al.* (2005) Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. *PLoS Pathog* 1: e5. doi:10.1371/journal.ppat.0010005
9. Fabre M, Koeck JL, Le Fleche P, Simon F, Herve V*, et al.* (2004) High genetic diversity revealed by variable-number tandem repeat genotyping and analysis of hsp65 gene polymorphism in a large collection of "Mycobacterium canettii" strains indicates that the M. tuberculosis complex is a recently emerged clone of "M. canettii". *J Clin Microbiol* 42: 3248-3255. doi:10.1128/JCM.42.7.3248-3255.2004
10. Brosch R, Gordon SV, Marmiesse M, Brodin P, Buchrieser C*, et al.* (2002) A new evolutionary scenario for the Mycobacterium tuberculosis complex. *Proc Natl Acad Sci U S A* 99: 3684-3689. doi:10.1073/pnas.052548299

11. Gagneux S, DeRiemer K, Van T, Kato-Maeda M, de Jong BC*, et al.* (2006) Variable host-pathogen compatibility in Mycobacterium tuberculosis. *Proc Natl Acad Sci U S A* 103: 2869-2873. doi:10.1073/pnas.0511240103
12. Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S*, et al.* (2008) High functional diversity in Mycobacterium tuberculosis driven by genetic drift and human demography. *PLoS Biol* 6: e311. doi:10.1371/journal.pbio.0060311
13. Coscolla M, Gagneux S (2010) Does M. tuberculosis genomic diversity explain disease diversity? *Drug Discov Today Dis Mech* 7: e43-e59. doi:10.1016/j.ddmec.2010.09.004
14. Portevin D, Gagneux S, Comas I, Young D (2011) Human macrophage responses to clinical isolates from the Mycobacterium tuberculosis complex discriminate between ancient and modern lineages. *PLoS Pathog* 7: e1001307. doi:10.1371/journal.ppat.1001307
15. Brites D, Gagneux S (2012) Old and new selective pressures on Mycobacterium tuberculosis. *Infect Genet Evol* 12: 678-685. doi:10.1016/j.meegid.2011.08.010
16. Lopez B, Aguilar D, Orozco H, Burger M, Espitia C*, et al.* (2003) A marked difference in pathogenesis and immune response induced by different Mycobacterium tuberculosis genotypes. *Clin Exp Immunol* 133: 30-37.
17. Caws M, Thwaites G, Dunstan S, Hawn TR, Lan NT*, et al.* (2008) The influence of host and bacterial genotype on the development of disseminated disease with Mycobacterium tuberculosis. *PLoS Pathog* 4: e1000034. doi:10.1371/journal.ppat.1000034
18. Thwaites G, Caws M, Chau TT, D'Sa A, Lan NT*, et al.* (2008) Relationship between Mycobacterium tuberculosis genotype and the clinical phenotype of pulmonary and meningeal tuberculosis. *J Clin Microbiol* 46: 1363-1368. doi:10.1128/JCM.02180-07
19. Borgdorff MW, Van Deutekom H, De Haas PE, Kremer K, Van Soolingen D (2004) Mycobacterium tuberculosis, Beijing genotype strains not associated with radiological presentation of pulmonary tuberculosis. *Tuberculosis (Edinb)* 84: 337-340. doi:10.1016/j.tube.2003.10.002
20. de Jong BC, Hill PC, Aiken A, Awine T, Antonio M*, et al.* (2008) Progression to active tuberculosis, but not transmission, varies by Mycobacterium tuberculosis lineage in The Gambia. *J Infect Dis* 198: 1037-1043. doi:10.1086/591504
21. Dye C, Williams BG (2010) The population dynamics and control of tuberculosis. *Science* 328: 856-861. doi:10.1126/science.1185449

22. Dalton T, Cegielski P, Akksilp S, Asencios L, Caoili JC*, et al.* (2012) Prevalence of and risk factors for resistance to second-line drugs in people with multidrug-resistant tuberculosis in eight countries: a prospective cohort study. *Lancet*. doi:10.1016/S0140-6736(12)60734-X

23. Visser ME, Stead MC, Walzl G, Warren R, Schomaker M*, et al.* (2012) Baseline predictors of sputum culture conversion in pulmonary tuberculosis: importance of cavities, smoking, time to detection and W-Beijing genotype. *PLoS One* 7: e29588. doi:10.1371/journal.pone.0029588

24. Arora J, Singh UB, Suresh N, Rana T, Porwal C*, et al.* (2009) Characterization of predominant Mycobacterium tuberculosis strains from different subpopulations of India. *Infect Genet Evol* 9: 832-839. doi:10.1016/j.meegid.2009.05.008

25. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C*, et al.* (1998) Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature* 393: 537-544. doi:10.1038/31159

26. Achtman M (2008) Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens. *Annu Rev Microbiol* 62: 53-70. doi:10.1146/annurev.micro.62.081307.162832

27. Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN (2006) Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev* 19: 658-685. doi:10.1128/CMR.00061-05

28. Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbon MH*, et al.* (2006) Global phylogeny of Mycobacterium tuberculosis based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set. *J Bacteriol* 188: 759-772. doi:10.1128/JB.188.2.759-772.2006

29. Schork NJ, Fallin D, Lanchbury JS (2000) Single nucleotide polymorphisms and the future of genetic epidemiology. *Clin Genet* 58: 250-264.

30. Mathema B, Kurepina N, Fallows D, Kreiswirth BN (2008) Lessons from molecular epidemiology and comparative genomics. *Semin Respir Crit Care Med* 29: 467-480. doi:10.1055/s-0028-1085699

31. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD*, et al.* (1993) Strain identification of Mycobacterium tuberculosis by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* 31: 406-409.

32. Millan-Lou MI, Alonso H, Gavin P, Hernandez-Febles M, Campos-Herrero MI*, et al.* (2012) Rapid test for identification of a highly transmissible Mycobacterium tuberculosis Beijing strain of sub-Saharan origin. *J Clin Microbiol* 50: 516-518. doi:10.1128/JCM.06314-11

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

34. Allix-Beguec C, Harmsen D, Weniger T, Supply P, Niemann S (2008) Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of Mycobacterium tuberculosis complex isolates. *J Clin Microbiol* 46: 2692-2699. doi:10.1128/JCM.00540-08

35. Supply P, Mazars E, Lesjean S, Vincent V, Gicquel B*, et al.* (2000) Variable human minisatellite-like regions in the Mycobacterium tuberculosis genome. *Mol Microbiol* 36: 762-771.

36. Kremer K, Au BK, Yip PC, Skuce R, Supply P*, et al.* (2005) Use of variable-number tandem-repeat typing to differentiate Mycobacterium tuberculosis Beijing family isolates from Hong Kong and comparison with IS6110 restriction fragment length polymorphism typing and spoligotyping. *J Clin Microbiol* 43: 314-320. doi:10.1128/JCM.43.1.314-320.2005

37. Dye C, Watt CJ, Bleed DM, Williams BG (2003) What is the limit to case detection under the DOTS strategy for tuberculosis control? *Tuberculosis (Edinb)* 83: 35-43.

38. Kurbatova EV, Kaminski DA, Erokhin VV, Volchenkov GV, Andreevskaya SN*, et al.* (2012) Performance of Cepheid((R)) Xpert MTB/RIF ((R)) and TB-Biochip ((R)) MDR in two regions of Russia with a high prevalence of drug-resistant tuberculosis. *Eur J Clin Microbiol Infect Dis*. doi:10.1007/s10096-012-1798-0

39. Ling DI, Zwerling AA, Pai M (2008) GenoType MTBDR assays for the diagnosis of multidrug-resistant tuberculosis: a meta-analysis. *Eur Respir J* 32: 1165-1174. doi:10.1183/09031936.00061808

40. Ling DI, Zwerling AA, Pai M (2008) Rapid diagnosis of drug-resistant TB using line probe assays: from evidence to policy. *Expert Rev Respir Med* 2: 583-588. doi:10.1586/17476348.2.5.583

41. World Health Organization (2008) Anti-tuberculosis drug resistance in the world Report no.4. Geneva, Switzerland. WHO/HTM/TB/2008.394 WHO/HTM/TB/2008.394. 153 p.

42. Smith SE, Kurbatova EV, Cavanaugh JS, Cegielski JP (2012) Global isoniazid resistance patterns in rifampin-resistant and rifampin-susceptible tuberculosis. *Int J Tuberc Lung Dis* 16: 203-205. doi:10.5588/ijtld.11.0445

43. Gagneux S, Long CD, Small PM, Van T, Schoolnik GK*, et al.* (2006) The competitive cost of antibiotic resistance in Mycobacterium tuberculosis. *Science* 312: 1944-1946. doi:10.1126/science.1124410

44. Comas I, Borrell S, Roetzer A, Rose G, Malla B*, et al.* (2012) Whole-genome sequencing of rifampicin-resistant Mycobacterium tuberculosis strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet* 44: 106-110. doi:10.1038/ng.1038

45. Li WM, Wang SM, Li CY, Liu YH, Shen GM*, et al.* (2005) Molecular epidemiology of Mycobacterium tuberculosis in China: a nationwide random survey in 2000. *Int J Tuberc Lung Dis* 9: 1314-1319.

46. Borrell S, Gagneux S (2009) Infectiousness, reproductive fitness and evolution of drug-resistant Mycobacterium tuberculosis. *Int J Tuberc Lung Dis* 13: 1456-1466.

47. World Health Organization (1999) Issues relationg to the use of BCG in immunization programmes: A discussion document. Geneva, Switzerland. WHO/V&B/99.23 WHO/V&B/99.23. 45 p.

48. Young DB (2003) Building a better tuberculosis vaccine. *Nat Med* 9: 503-504. doi:10.1038/nm868

49. National Tuberculosis Programme (2011) Annual report 2010/2011: Nepal National Tuberculosis Programme. Kathmandu, Nepal: National Tuberculosis Programme. 252 p.

50. Malla P, Kanitz EE, Akhtar M, Falzon D, Feldmann K*, et al.* (2009) Ambulatory-based standardized therapy for multi-drug resistant tuberculosis: experience from Nepal, 2005-2006. *PLoS One* 4: e8313. doi:10.1371/journal.pone.0008313

51. Comas I, Homolka S, Niemann S, Gagneux S (2009) Genotyping of genetically monomorphic bacteria: DNA sequencing in Mycobacterium tuberculosis highlights the limitations of current methodologies. *PLoS One* 4: e7815. doi:10.1371/journal.pone.0007815

52. Stucki D, Malla B, Hostettler S, Huna T, Feldmann J*, et al.* (2012) Two new rapid SNP-typing methods for classifying Mycobacterium tuberculosis complex into the main phylogenetic lineages. *PLoS One* 7: e41253. doi:10.1371/journal.pone.0041253

53. Ramaswamy S, Musser JM (1998) Molecular genetic basis of antimicrobial agent resistance in Mycobacterium tuberculosis: 1998 update. *Tuber Lung Dis* 79: 3-29. doi:10.1054/tuld.1998.0002

54. Fenner L, Egger M, Bodmer T, Altpeter E, Zwahlen M*, et al.* (2012) Effect of mutation and genetic background on drug resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 56: 3047-3053. doi:10.1128/AAC.06460-11

55. Koser CU, Feuerriegel S, Summers DK, Archer JA, Niemann S (2012) Importance of the genetic diversity within the Mycobacterium tuberculosis complex for the development of novel antibiotics and diagnostic tests of drug resistance. *Antimicrob Agents Chemother* 56: 6080-6087. doi:10.1128/AAC.01641-12

56. Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM*, et al.* (2009) Tuberculosis drug resistance mutation database. *PLoS Med* 6: e2. doi:10.1371/journal.pmed.1000002

57. World Health Organization (2004) Transport of Infectious Substance. Geneva, Switzerland. WHO/CDS/CSR/LYO/2004.9 WHO/CDS/CSR/LYO/2004.9. 28 p.

58. Tuberculosis Division International Union Against Tuberculosis and Lung Disease (2005) Tuberculosis bacteriology--priorities and indications in high prevalence countries: position of the technical staff of the Tuberculosis Division of the International Union Against Tuberculosis and Lung Disease. *Int J Tuberc Lung Dis* 9: 355-361.

59. Arnvig KB, Comas I, Thomson NR, Houghton J, Boshoff HI*, et al.* (2011) Sequence-based analysis uncovers an abundance of non-coding RNA in the total transcriptome of Mycobacterium tuberculosis. *PLoS Pathog* 7: e1002342. doi:10.1371/journal.ppat.1002342

60. Singh UB, Suresh N, Bhanu NV, Arora J, Pant H*, et al.* (2004) Predominant tuberculosis spoligotypes, Delhi, India. *Emerg Infect Dis* 10: 1138-1142.

61. Narayanan S, Gagneux S, Hari L, Tsolaki AG, Rajasekhar S*, et al.* (2008) Genomic interrogation of ancestral Mycobacterium tuberculosis from south India. *Infect Genet Evol* 8: 474-483. doi:S1567-1348(07)00139-6 [pii]

10.1016/j.meegid.2007.09.007

62. Dong H, Shi L, Zhao X, Sang B, Lv B*, et al.* (2012) Genetic Diversity of Mycobacterium tuberculosis Isolates from Tibetans in Tibet, China. *PLoS One* 7: e33904. doi:10.1371/journal.pone.0033904

PONE-D-11-14065 [pii]

63. Demay C, Liens B, Burguiere T, Hill V, Couvin D*, et al.* (2012) SITVITWEB--a publicly available international multimarker database for studying Mycobacterium tuberculosis genetic diversity and molecular epidemiology. *Infect Genet Evol* 12: 755-766. doi:S1567-1348(12)00031-7 [pii]

10.1016/j.meegid.2012.02.004

64. Malla B, Stucki D, Borrell S, Feldmann J, Maharjan B*, et al.* (2012) First Insights into the Phylogenetic Diversity of Mycobacterium tuberculosis in Nepal. *PLoS One* 7: e52297. doi:10.1371/journal.pone.0052297

65. Ben-Selma W, Harizi H, Letaief M, Boukadida J (2012) Age- and gender-specific effects on NRAMP1 gene polymorphisms and risk of the development of active tuberculosis in Tunisian populations. *Int J Infect Dis* 16: e543-550. doi:10.1016/j.ijid.2011.11.016

66. Weiner M, Burman W, Luo CC, Peloquin CA, Engle M*, et al.* (2007) Effects of rifampin and multidrug resistance gene polymorphism on concentrations of moxifloxacin. *Antimicrob Agents Chemother* 51: 2861-2866. doi:10.1128/AAC.01621-06

67. Mathema B, Kurepina N, Yang G, Shashkina E, Manca C*, et al.* (2012) Epidemiologic consequences of microvariation in Mycobacterium tuberculosis. *J Infect Dis* 205: 964-974. doi:10.1093/infdis/jir876

68. Chernyaeva E, Dobrynin P, Pestova N, Matveeva N, Zhemkov V*, et al.* (2012) Molecular genetic analysis of Mycobacterium tuberculosis strains spread in different patient groups in St. Petersburg, Russia. *Eur J Clin Microbiol Infect Dis* 31: 1753-1757. doi:10.1007/s10096-011-1497-2

69. World Health Organization (2010) Multidrug and extensively drug-resistant tb (m/xdr-tb); 2010 Global Report on Surveillance and Response. Geneva, Switzerland: World Health Organization. 71 p.

70. World Health Organization (2011) Global tuberculosis control: WHO report 2011. Geneva, Switzerland: World Health Organization. 258 p.

71. National Tuberculosis Program (2010) NTP Annual Report 2009/2010. In: Program NTC, editor. Bhaktapur, Nepal.

72. Gutacker MM, Smoot JC, Migliaccio CA, Ricklefs SM, Hua S*, et al.* (2002) Genome-wide analysis of synonymous single nucleotide polymorphisms in Mycobacterium tuberculosis

complex organisms: resolution of genetic relationships among closely related microbial strains. *Genetics* 162: 1533-1543.

73. Gagneux S, Small PM (2007) Global phylogeography of Mycobacterium tuberculosis and implications for tuberculosis product development. *Lancet Infect Dis* 7: 328-337. doi:S1473-3099(07)70108-1 [pii]

10.1016/S1473-3099(07)70108-1

74. Baker L, Brown T, Maiden MC, Drobniewski F (2004) Silent nucleotide polymorphisms and a phylogeny for Mycobacterium tuberculosis. *Emerg Infect Dis* 10: 1568-1577. doi:10.3201/eid1009.040046

75. Hirsh AE, Tsolaki AG, DeRiemer K, Feldman MW, Small PM (2004) Stable association between strains of Mycobacterium tuberculosis and their human host populations. *Proc Natl Acad Sci U S A* 101: 4871-4876. doi:10.1073/pnas.0305627101

76. Sun JR, Dou HY, Lee SY, Chiueh TS, Lu JJ (2011) Epidemiological studies of Beijing strains of Mycobacterium tuberculosis from Taipei and other Asian cities based on MIRU profiles. *APMIS* 119: 581-587. doi:10.1111/j.1600-0463.2011.02790.x

77. Dou HY, Tseng FC, Lin CW, Chang JR, Sun JR*, et al.* (2008) Molecular epidemiology and evolutionary genetics of Mycobacterium tuberculosis in Taipei. *BMC Infect Dis* 8: 170. doi:1471-2334-8-170 [pii]

10.1186/1471-2334-8-170

78. Tho DQ, Torok ME, Yen NT, Bang ND, Lan NT*, et al.* (2012) Influence of Antituberculosis Drug Resistance and Mycobacteruim tuberculosis lineage on Outcome in HIV-associated Tuberculous Meningitis. *Antimicrob Agents Chemother*. doi:AAC.00319-12 [pii]

10.1128/AAC.00319-12

79. Hanekom M, Gey van Pittius NC, McEvoy C, Victor TC, Van Helden PD*, et al.* (2011) Mycobacterium tuberculosis Beijing genotype: a template for success. *Tuberculosis (Edinb)* 91: 510-523. doi:S1472-9792(11)00130-2 [pii]

10.1016/j.tube.2011.07.005

80. Lari N, Rindi L, Cristofani R, Rastogi N, Tortoli E*, et al.* (2009) Association of Mycobacterium tuberculosis complex isolates of BOVIS and Central Asian (CAS) genotypic lineages with extrapulmonary disease. *Clin Microbiol Infect* 15: 538-543. doi:CLM2712 [pii]

10.1111/j.1469-0691.2009.02712.x

81. Click ES, Moonan PK, Winston CA, Cowan LS, Oeltmann JE (2012) Relationship between Mycobacterium tuberculosis phylogenetic lineage and clinical site of tuberculosis. *Clin Infect Dis* 54: 211-219. doi:10.1093/cid/cir788

82. Parwati I, van Crevel R, van Soolingen D (2010) Possible underlying mechanisms for successful emergence of the Mycobacterium tuberculosis Beijing genotype strains. *Lancet Infect Dis* 10: 103-111. doi:10.1016/S1473-3099(09)70330-5

83. Pang Y, Zhou Y, Zhao B, Liu G, Jiang G*, et al.* (2012) Spoligotyping and Drug Resistance Analysis of Mycobacterium tuberculosis Strains from National Survey in China. *PLoS One* 7: e32976. doi:10.1371/journal.pone.0032976

PONE-D-11-15171 [pii]

84. Fenner L, Egger M, Bodmer T, Altpeter E, Zwahlen M*, et al.* (2012) Effect of mutation and genetic background on drug resistance in Mycobacterium tuberculosis. *Antimicrob Agents Chemother*. doi:AAC.06460-11 [pii]

10.1128/AAC.06460-11

85. Rajapaksa US, Perera AJ (2011) Sublineages of Beijing Strain of Mycobacterium tuberculosis in Sri Lanka. *Indian J Microbiol* 51: 410-412. doi:10.1007/s12088-011-0150-1

86. Iwamoto T, Yoshida S, Suzuki K, Wada T (2008) Population structure analysis of the Mycobacterium tuberculosis Beijing family indicates an association between certain sublineages and multidrug resistance. *Antimicrob Agents Chemother* 52: 3805-3809. doi:10.1128/AAC.00579-08

87. Lasunskaia E, Ribeiro SC, Manicheva O, Gomes LL, Suffys PN*, et al.* (2010) Emerging multidrug resistant Mycobacterium tuberculosis strains of the Beijing genotype circulating in Russia express a pattern of biological properties associated with enhanced virulence. *Microbes Infect* 12: 467-475. doi:10.1016/j.micinf.2010.02.008

88. Malik AN, Godfrey-Faussett P (2005) Effects of genetic variability of Mycobacterium tuberculosis strains on the presentation of disease. *Lancet Infect Dis* 5: 174-183. doi:10.1016/S1473-3099(05)01310-1

89. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S*, et al.* (2006) Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol* 44: 4498-4510. doi:10.1128/JCM.01392-06

90. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A*, et al.* (2006) Mycobacterium tuberculosis complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. *BMC Microbiol* 6: 23. doi:1471-2180-6-23 [pii]
10.1186/1471-2180-6-23
91. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D*, et al.* (1997) Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. *J Clin Microbiol* 35: 907-914.
92. Fenner L, Malla B, Ninet B, Dubuis O, Stucki D*, et al.* (2011) "Pseudo-Beijing": evidence for convergent evolution in the direct repeat region of Mycobacterium tuberculosis. *PLoS One* 6: e24737. doi:10.1371/journal.pone.0024737
PONE-D-11-11180 [pii]
93. Sreevatsan S, Pan X, Stockbauer KE, Connell ND, Kreiswirth BN*, et al.* (1997) Restricted structural gene polymorphism in the Mycobacterium tuberculosis complex indicates evolutionarily recent global dissemination. *Proc Natl Acad Sci U S A* 94: 9869-9874.
94. Supply P, Warren RM, Banuls AL, Lesjean S, Van Der Spuy GD*, et al.* (2003) Linkage disequilibrium between minisatellite loci supports clonal evolution of Mycobacterium tuberculosis in a high tuberculosis incidence area. *Mol Microbiol* 47: 529-538. doi:10.1046/j.1365-2958.2003.03315.x
95. Brudey K, Filliol I, Ferdinand S, Guernier V, Duval P*, et al.* (2006) Long-term population-based genotyping study of Mycobacterium tuberculosis complex isolates in the French departments of the Americas. *J Clin Microbiol* 44: 183-191. doi:44/1/183 [pii]
10.1128/JCM.44.1.183-191.2006
96. Niemann S, Koser CU, Gagneux S, Plinke C, Homolka S*, et al.* (2009) Genomic diversity among drug sensitive and multidrug resistant isolates of Mycobacterium tuberculosis with identical DNA fingerprints. *PLoS One* 4: e7407. doi:10.1371/journal.pone.0007407
97. Stucki D, Malla B, Hostettler S, Huna T, Feldmann J*, et al.* (2012) Two New Rapid SNP-Typing Methods for Classifying Mycobacterium tuberculosis Complex into the Main Phylogenetic Lineages. *PLoS One* 7: e41253. doi:10.1371/journal.pone.0041253
98. World Health Organization (2010) The GlobalPlan to stop tb 2011–2015: transforming the fight towards elimination of tuberculosis.

99. World Health Organization (2009) Guidelines for surveillance of drug resistance in tuberculosis WHO/HTM/TB/2009.422 WHO/HTM/TB/2009.422.

100. Lillebaek T, Andersen AB, Dirksen A, Glynn JR, Kremer K (2003) Mycobacterium tuberculosis Beijing genotype. *Emerg Infect Dis* 9: 1553-1557.

101. Victor TC, Jordaan AM, van Rie A, van der Spuy GD, Richardson M*, et al.* (1999) Detection of mutations in drug resistance genes of Mycobacterium tuberculosis by a dot-blot hybridization strategy. *Tuber Lung Dis* 79: 343-348. doi:10.1054/tuld.1999.0222

102. Feuerriegel S, Cox HS, Zarkua N, Karimovich HA, Braker K*, et al.* (2009) Sequence analyses of just four genes to detect extensively drug-resistant Mycobacterium tuberculosis strains in multidrug-resistant tuberculosis patients undergoing treatment. *Antimicrob Agents Chemother* 53: 3353-3356. doi:AAC.00050-09 [pii]

10.1128/AAC.00050-09

103. Brossier F, Veziris N, Aubry A, Jarlier V, Sougakoff W (2010) Detection by GenoType MTBDRsl test of complex mechanisms of resistance to second-line drugs and ethambutol in multidrug-resistant Mycobacterium tuberculosis complex isolates. *J Clin Microbiol* 48: 1683-1689. doi:10.1128/JCM.01947-09

104. Staden R, Beal KF, Bonfield JK (2000) The Staden package, 1998. *Methods Mol Biol* 132: 115-130.

105. Bonfield JK, Smith K, Staden R (1995) A new DNA sequence assembly program. *Nucleic Acids Res* 23: 4992-4999.

106. Drobniewski FA, Balabanova YM, Ruddy MC, Graham C, Kuznetzov SI*, et al.* (2005) Tuberculosis, HIV seroprevalence and intravenous drug abuse in prisoners. *Eur Respir J* 26: 298-304. doi:26/2/298 [pii]

10.1183/09031936.05.00136004

107. Svensson E, Millet J, Lindqvist A, Olsson M, Ridell M*, et al.* (2011) Impact of immigration on tuberculosis epidemiology in a low-incidence country. *Clin Microbiol Infect* 17: 881-887. doi:10.1111/j.1469-0691.2010.03358.x

CLM3358 [pii]

108. Liu RX, Li QZ, Xing LL, Peng Z, Zhu CM (2011) [Genotyping of 210 Mycobacterium tuberculosis strains with Spoligotyping and MIRU-VNTR among pediatric tuberculosis patients in Chongqing]. *Zhonghua Liu Xing Bing Xue Za Zhi* 32: 593-597.

109. Hu Y, Jiang WL, Zhao Q, Wang WB, Xu B (2009) [The combined application of multiple genotyping methods in identifying genotypes of Mycobacterium tuberculosis strain circulating in rural China]. *Zhonghua Jie He He Hu Xi Za Zhi* 32: 576-580.

110. Guo YL, Liu Y, Wang SM, Li CY, Jiang GL*, et al.* (2011) Genotyping and drug resistance patterns of Mycobacterium tuberculosis strains in five provinces of China. *Int J Tuberc Lung Dis* 15: 789-794. doi:10.5588/ijtld.10.0403

111. Han H, Wang F, Xiao Y, Ren Y, Chao Y*, et al.* (2007) Utility of mycobacterial interspersed repetitive unit typing for differentiating Mycobacterium tuberculosis isolates in Wuhan, China. *J Med Microbiol* 56: 1219-1223. doi:56/9/1219 [pii]

10.1099/jmm.0.47005-0

112. 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. *Emerg Infect Dis* 12: 736-743. doi:10.3201/eid1205.050400

113. Buu TN, Huyen MN, Lan NT, Quy HT, Hen NV*, et al.* (2009) The Beijing genotype is associated with young age and multidrug-resistant tuberculosis in rural Vietnam. *Int J Tuberc Lung Dis* 13: 900-906.

114. Buu TN, Huyen MN, Lan NN, Quy HT, Hen NV*, et al.* (2009) Mycobacterium tuberculosis genotype and case notification rates, rural Vietnam, 2003-2006. *Emerg Infect Dis* 15: 1570-1577. doi:10.3201/eid1510.090170

115. Holmes CB, Hausler H, Nunn P (1998) A review of sex differences in the epidemiology of tuberculosis. *Int J Tuberc Lung Dis* 2: 96-104.

116. Borgdorff MW, Nagelkerke NJ, Dye C, Nunn P (2000) Gender and tuberculosis: a comparison of prevalence surveys with notification data to explore sex differences in case detection. *Int J Tuberc Lung Dis* 4: 123-132.

117. Connolly M, Nunn P (1996) Women and tuberculosis. *World Health Stat Q* 49: 115-119.

118. Getahun H, Gunneberg C, Granich R, Nunn P (2010) HIV infection-associated tuberculosis: the epidemiology and the response. *Clin Infect Dis* 50 Suppl 3: S201-207. doi:10.1086/651492

119. Neyrolles O, Quintana-Murci L (2009) Sexual inequality in tuberculosis. *PLoS Med* 6: e1000199. doi:10.1371/journal.pmed.1000199

120. Uwizeye CB, De Serres G, Gilca R, Schwartzman K, Gasana M (2011) Tuberculosis may be underestimated in Rwandan women. *Int J Tuberc Lung Dis* 15: 776-781. doi:10.5588/ijtld.10.0454

121. Borrell S, Gagneux S (2011) Strain diversity, epistasis and the evolution of drug resistance in Mycobacterium tuberculosis. *Clin Microbiol Infect* 17: 815-820. doi:10.1111/j.1469-0691.2011.03556.x

122. Sun YJ, Lee AS, Wong SY, Paton NI (2006) Association of Mycobacterium tuberculosis Beijing genotype with tuberculosis relapse in Singapore. *Epidemiol Infect* 134: 329-332. doi:10.1017/S095026880500525X

123. World Health Organization (2011) National Immunization Program. http://www.nep.searo.who.int/EN/Section4/Section29/Section89.htm. Accessed 2012 October 26.

124. van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H*, et al.* (1995) Predominance of a single genotype of Mycobacterium tuberculosis in countries of east Asia. *J Clin Microbiol* 33: 3234-3238.

125. Cowley D, Govender D, February B, Wolfe M, Steyn L*, et al.* (2008) Recent and rapid emergence of W-Beijing strains of Mycobacterium tuberculosis in Cape Town, South Africa. *Clin Infect Dis* 47: 1252-1259. doi:10.1086/592575

126. Tsolaki AG, Gagneux S, Pym AS, Goguet de la Salmoniere YO, Kreiswirth BN*, et al.* (2005) Genomic deletions classify the Beijing/W strains as a distinct genetic lineage of Mycobacterium tuberculosis. *J Clin Microbiol* 43: 3185-3191. doi:10.1128/JCM.43.7.3185-3191.2005

127. Kremer K, Glynn JR, Lillebaek T, Niemann S, Kurepina NE*, et al.* (2004) Definition of the Beijing/W lineage of Mycobacterium tuberculosis on the basis of genetic markers. *J Clin Microbiol* 42: 4040-4049. doi:10.1128/JCM.42.9.4040-4049.2004
42/9/4040 [pii]

128. Tsolaki AG, Hirsh AE, DeRiemer K, Enciso JA, Wong MZ*, et al.* (2004) Functional and evolutionary genomics of Mycobacterium tuberculosis: insights from genomic deletions in 100 strains. *Proc Natl Acad Sci U S A* 101: 4865-4870. doi:10.1073/pnas.0305634101
0305634101 [pii]

129. Cohen T, Wilson D, Wallengren K, Samuel EY, Murray M (2011) Mixed-strain Mycobacterium tuberculosis infections among patients dying in a hospital in KwaZulu-Natal, South Africa. *J Clin Microbiol* 49: 385-388. doi:10.1128/JCM.01378-10

130. Flores L, Van T, Narayanan S, DeRiemer K, Kato-Maeda M*, et al.* (2007) Large sequence polymorphisms classify Mycobacterium tuberculosis strains with ancestral spoligotyping patterns. *J Clin Microbiol* 45: 3393-3395. doi:JCM.00828-07 [pii]
10.1128/JCM.00828-07

131. Warren RM, Streicher EM, Sampson SL, van der Spuy GD, Richardson M*, et al.* (2002) Microevolution of the direct repeat region of Mycobacterium tuberculosis: implications for interpretation of spoligotyping data. *J Clin Microbiol* 40: 4457-4465.

132. Kato-Maeda M, Gagneux S, Flores LL, Kim EY, Small PM*, et al.* (2011) Strain classification of Mycobacterium tuberculosis: congruence between large sequence polymorphisms and spoligotypes. *Int J Tuberc Lung Dis* 15: 131-133.

133. van der Spuy GD, Kremer K, Ndabambi SL, Beyers N, Dunbar R*, et al.* (2009) Changing Mycobacterium tuberculosis population highlights clade-specific pathogenic characteristics. *Tuberculosis (Edinb)* 89: 120-125. doi:10.1016/j.tube.2008.09.003

134. Caminero JA, Pena MJ, Campos-Herrero MI, Rodriguez JC, Garcia I*, et al.* (2001) Epidemiological evidence of the spread of a Mycobacterium tuberculosis strain of the Beijing genotype on Gran Canaria Island. *Am J Respir Crit Care Med* 164: 1165-1170.

135. Department of Health Service (2011) Annual Report; Department of Health Services (2010/2011).

136. Wang L, Wang X (2012) Influence of temporary migration on the transmission of infectious diseases in a migrants' home village. *J Theor Biol* 300: 100-109. doi:10.1016/j.jtbi.2012.01.004

137. Barnes I, Duda A, Pybus OG, Thomas MG (2011) Ancient urbanization predicts genetic resistance to tuberculosis. *Evolution* 65: 842-848. doi:10.1111/j.1558-5646.2010.01132.x

138. Ormerod LP (1998) Is new immigrant screening for tuberculosis still worthwhile? *J Infect* 37: 39-40.

139. Kato-Maeda M, Metcalfe JZ, Flores L (2011) Genotyping of Mycobacterium tuberculosis: application in epidemiologic studies. *Future Microbiol* 6: 203-216. doi:10.2217/fmb.10.165

140. Central Bureau of Statistics (2012) Population census preliminary report - 2011. Kathmandu: Central Bureau of Statistics, Nepal. 18 p.

141. The World Bank (2012) World Development Indicators 2012. World Bank. 463 p.

142. Marais BJ, Hesseling AC, Schaaf HS, Gie RP, van Helden PD*, et al.* (2009) Mycobacterium tuberculosis transmission is not related to household genotype in a setting of high endemicity. *J Clin Microbiol* 47: 1338-1343. doi:10.1128/JCM.02490-08

143. Hubad B, Lapanje A (2012) Inadequate hospital ventilation system increases the risk of nosocomial Mycobacterium tuberculosis. *J Hosp Infect* 80: 88-91. doi:10.1016/j.jhin.2011.10.014

144. Dharmadhikari AS, Mphahlele M, Stoltz A, Venter K, Mathebula R*, et al.* (2012) Surgical face masks worn by patients with multidrug-resistant tuberculosis: impact on infectivity of air on a hospital ward. *Am J Respir Crit Care Med* 185: 1104-1109. doi:10.1164/rccm.201107-1190OC

145. Kamper-Jorgensen Z, Andersen AB, Kok-Jensen A, Bygbjerg IC, Thomsen VO*, et al.* (2012) Characteristics of non-clustered tuberculosis in a low burden country. *Tuberculosis (Edinb)* 92: 226-231. doi:10.1016/j.tube.2012.02.001

146. Long R, Sutherland K, Kunimoto D, Cowie R, Manfreda J (2002) The epidemiology of tuberculosis among foreign-born persons in Alberta, Canada, 1989-1998: identification of high risk groups. *Int J Tuberc Lung Dis* 6: 615-621.

147. Oeltmann JE, Varma JK, Ortega L, Liu Y, O'Rourke T*, et al.* (2008) Multidrug-resistant tuberculosis outbreak among US-bound Hmong refugees, Thailand, 2005. *Emerg Infect Dis* 14: 1715-1721. doi:10.3201/eid1411.071629

148. Lin H, Shin S, Blaya JA, Zhang Z, Cegielski P*, et al.* (2011) Assessing spatiotemporal patterns of multidrug-resistant and drug-sensitive tuberculosis in a South American setting. *Epidemiol Infect* 139: 1784-1793. doi:10.1017/S0950268810002797

149. Manjourides J, Lin HH, Shin S, Jeffery C, Contreras C*, et al.* (2012) Identifying multidrug resistant tuberculosis transmission hotspots using routinely collected data. *Tuberculosis (Edinb)* 92: 273-279. doi:10.1016/j.tube.2012.02.003

150. Bodmer T, Strohle A (2012) Diagnosing pulmonary tuberculosis with the Xpert MTB/RIF test. *J Vis Exp*: e3547. doi:10.3791/3547

151. Ignatyeva O, Kontsevaya I, Kovalyov A, Balabanova Y, Nikolayevskyy V*, et al.* (2012) Detection of resistance to second-line antituberculosis drugs by use of the genotype

MTBDRsl assay: a multicenter evaluation and feasibility study. *J Clin Microbiol* 50: 1593-1597. doi:10.1128/JCM.00039-12

152. Zhang Y, Yew WW (2009) Mechanisms of drug resistance in Mycobacterium tuberculosis. *Int J Tuberc Lung Dis* 13: 1320-1330.

153. Hillemann D, Rusch-Gerdes S, Richter E (2007) Evaluation of the GenoType MTBDRplus assay for rifampin and isoniazid susceptibility testing of Mycobacterium tuberculosis strains and clinical specimens. *J Clin Microbiol* 45: 2635-2640. doi:10.1128/JCM.00521-07

154. Rinder H, Mieskes KT, Loscher T (2001) Heteroresistance in Mycobacterium tuberculosis. *Int J Tuberc Lung Dis* 5: 339-345.

155. Mariam SH, Werngren J, Aronsson J, Hoffner S, Andersson DI (2011) Dynamics of antibiotic resistant Mycobacterium tuberculosis during long-term infection and antibiotic treatment. *PLoS One* 6: e21147. doi:10.1371/journal.pone.0021147

156. Hazbon MH, Brimacombe M, Bobadilla del Valle M, Cavatore M, Guerrero MI*, et al.* (2006) Population genetics study of isoniazid resistance mutations and evolution of multidrug-resistant Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 50: 2640-2649. doi:10.1128/AAC.00112-06

157. Gagneux S, Burgos MV, DeRiemer K, Encisco A, Munoz S*, et al.* (2006) Impact of bacterial genetics on the transmission of isoniazid-resistant Mycobacterium tuberculosis. *PLoS Pathog* 2: e61. doi:10.1371/journal.ppat.0020061

158. Yadav R, Sethi S, Dhatwalia SK, Gupta D, Mewara A*, et al.* (2013) Molecular characterisation of drug resistance in Mycobacterium tuberculosis isolates from North India. *Int J Tuberc Lung Dis* 17: 251-257. doi:10.5588/ijtld.12.0319

159. Siddiqi N, Shamim M, Hussain S, Choudhary RK, Ahmed N*, et al.* (2002) Molecular characterization of multidrug-resistant isolates of Mycobacterium tuberculosis from patients in North India. *Antimicrob Agents Chemother* 46: 443-450.

160. van Soolingen D, de Haas PE, van Doorn HR, Kuijper E, Rinder H*, et al.* (2000) Mutations at amino acid position 315 of the katG gene are associated with high-level resistance to isoniazid, other drug resistance, and successful transmission of Mycobacterium tuberculosis in the Netherlands. *J Infect Dis* 182: 1788-1790. doi:10.1086/317598

161. Pym AS, Domenech P, Honore N, Song J, Deretic V*, et al.* (2001) Regulation of catalase-peroxidase (KatG) expression, isoniazid sensitivity and virulence by furA of Mycobacterium tuberculosis. *Mol Microbiol* 40: 879-889.

162. Pym AS, Saint-Joanis B, Cole ST (2002) Effect of katG mutations on the virulence of Mycobacterium tuberculosis and the implication for transmission in humans. *Infect Immun* 70: 4955-4960.

163. van Soolingen D, Borgdorff MW, de Haas PE, Sebek MM, Veen J*, et al.* (1999) Molecular epidemiology of tuberculosis in the Netherlands: a nationwide study from 1993 through 1997. *J Infect Dis* 180: 726-736. doi:10.1086/314930

164. Marttila HJ, Soini H, Eerola E, Vyshnevskaya E, Vyshnevskiy BI*, et al.* (1998) A Ser315Thr substitution in KatG is predominant in genetically heterogeneous multidrug-resistant Mycobacterium tuberculosis isolates originating from the St. Petersburg area in Russia. *Antimicrob Agents Chemother* 42: 2443-2445.

165. Sun G, Luo T, Yang C, Dong X, Li J*, et al.* (2012) Dynamic population changes in Mycobacterium tuberculosis during acquisition and fixation of drug resistance in patients. *J Infect Dis* 206: 1724-1733. doi:10.1093/infdis/jis601

166. Lavender C, Globan M, Sievers A, Billman-Jacobe H, Fyfe J (2005) Molecular characterization of isoniazid-resistant Mycobacterium tuberculosis isolates collected in Australia. *Antimicrob Agents Chemother* 49: 4068-4074. doi:10.1128/AAC.49.10.4068-4074.2005

167. Bergval I, Kwok B, Schuitema A, Kremer K, van Soolingen D*, et al.* (2012) Pre-existing isoniazid resistance, but not the genotype of Mycobacterium tuberculosis drives rifampicin resistance codon preference in vitro. *PLoS One* 7: e29108. doi:10.1371/journal.pone.0029108

168. Sherman DR, Mdluli K, Hickey MJ, Arain TM, Morris SL*, et al.* (1996) Compensatory ahpC gene expression in isoniazid-resistant Mycobacterium tuberculosis. *Science* 272: 1641-1643.

169. Piatek AS, Telenti A, Murray MR, El-Hajj H, Jacobs WR, Jr.*, et al.* (2000) Genotypic analysis of Mycobacterium tuberculosis in two distinct populations using molecular beacons: implications for rapid susceptibility testing. *Antimicrob Agents Chemother* 44: 103-110.

170. Grange JM, Zumla A (2002) The global emergency of tuberculosis: what is the cause? *J R Soc Promot Health* 122: 78-81.

171. World Health Organization (2011) Global map and information on XDR-TB. Available at: http://www.who.int/tb/challenges/mdr/drs_maps_oct2011.pdf. Accessed on: 2012 November 06.

172. World Health Organization (2012) Tuberculosis in the South-East Asia Region: The Regional Report:2012. SEA/TB/338 SEA/TB/338. 156 p.

173. World Health Organization (2010) Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 global report on surveillance and response. World Health Organization, Geneva, Switzerland: World Health Organization. WHO/HTM/TB/2010.3 WHO/HTM/TB/2010.3. 71 p.

174. Leung EC, Leung CC, Kam KM, Yew WW, Chang KC*, et al.* (2012) Transmission of multidrug-resistant and extensively drug-resistant tuberculosis in a metropolitan city. *Eur Respir J*. doi:10.1183/09031936.00071212

175. Udwadia ZF (2008) XDR-TB in India : when will we heed the alarm? *J Assoc Physicians India* 56: 409-410.

176. World Health Organization (2012) List of Supranational Reference laboratories. Available at: http://www.stoptb.org/wg/gli/assets/documents/srl_network_mar10.pdf. Accessed on: 2012 November 09.

177. World Health Organization (2001) Guidelines for drug susceptibility testing for second-line anti-tuberculosis drugs for DOTS-PLUS. Geneva, Switzerland. WHO/CDS/TB/2001.288 WHO/CDS/TB/2001.288. 17 p.

178. Gopaul KK, Brown TJ, Gibson AL, Yates MD, Drobniewski FA (2006) Progression toward an improved DNA amplification-based typing technique in the study of Mycobacterium tuberculosis epidemiology. *J Clin Microbiol* 44: 2492-2498. doi:10.1128/JCM.01428-05

179. Banerjee A, Dubnau E, Quemard A, Balasubramanian V, Um KS*, et al.* (1994) inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis. *Science* 263: 227-230.

180. Kelley CL, Rouse DA, Morris SL (1997) Analysis of ahpC gene mutations in isoniazid-resistant clinical isolates of Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 41: 2057-2058.

181. Casali N, Nikolayevskyy V, Balabanova Y, Ignatyeva O, Kontsevaya I*, et al.* (2012) Microevolution of extensively drug-resistant tuberculosis in Russia. *Genome Res* 22: 735-745. doi:10.1101/gr.128678.111

182. Ajbani K, Rodrigues C, Shenai S, Mehta A (2011) Mutation detection and accurate diagnosis of extensively drug-resistant tuberculosis: report from a tertiary care center in India. *J Clin Microbiol* 49: 1588-1590. doi:10.1128/JCM.00113-11

183. Maus CE, Plikaytis BB, Shinnick TM (2005) Molecular analysis of cross-resistance to capreomycin, kanamycin, amikacin, and viomycin in Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 49: 3192-3197. doi:10.1128/AAC.49.8.3192-3197.2005

184. Long Q, Li W, Du Q, Fu Y, Liang Q*, et al.* (2012) gyrA/B fluoroquinolone resistance allele profiles amongst Mycobacterium tuberculosis isolates from mainland China. *Int J Antimicrob Agents* 39: 486-489. doi:10.1016/j.ijantimicag.2012.02.015

185. World Health Organization (2010) Drug-resistant tuberculosis now at record levels. Available at: http://www.who.int/mediacentre/news/releases/2010/drug_resistant_tb_20100318/en/index.html. Accessed on: 2012 November 05.

186. Lee JH, Ammerman NC, Nolan S, Geiman DE, Lun S*, et al.* (2012) Isoniazid resistance without a loss of fitness in Mycobacterium tuberculosis. *Nat Commun* 3: 753. doi:10.1038/ncomms1724

187. Basu S, Galvani AP (2008) The transmission and control of XDR TB in South Africa: an operations research and mathematical modelling approach. *Epidemiol Infect* 136: 1585-1598. doi:10.1017/S0950268808000964

188. Kim SJ, Espinal MA, Abe C, Bai GH, Boulahbal F*, et al.* (2004) Is second-line anti-tuberculosis drug susceptibility testing reliable? *Int J Tuberc Lung Dis* 8: 1157-1158.

189. Andrews JR, Gandhi NR, Moodley P, Shah NS, Bohlken L*, et al.* (2008) Exogenous reinfection as a cause of multidrug-resistant and extensively drug-resistant tuberculosis in rural South Africa. *J Infect Dis* 198: 1582-1589. doi:10.1086/592991

190. Migliori GB, Sotgiu G, Gandhi NR, Falzon D, Deriemer K*, et al.* (2012) Drug resistance beyond XDR-TB: results from a large individual patient data meta-analysis. *Eur Respir J*. doi:10.1183/09031936.00136312

191. Glynn JR, Bauer J, de Boer AS, Borgdorff MW, Fine PE*, et al.* (1999) Interpreting DNA fingerprint clusters of Mycobacterium tuberculosis. European Concerted Action on Molecular Epidemiology and Control of Tuberculosis. *Int J Tuberc Lung Dis* 3: 1055-1060.

192. Houben RM, Glynn JR (2009) A systematic review and meta-analysis of molecular epidemiological studies of tuberculosis: development of a new tool to aid interpretation. *Trop Med Int Health* 14: 892-909. doi:10.1111/j.1365-3156.2009.02316.x

193. Kakinuma Y, Maeda Y, Mason C, Goldsmith CE, Coulter WA*, et al.* (2012) Molecular characterisation of the quinolone resistance-determining regions (QRDR) including gyrA, gyrB, parC and parE genes in Streptococcus pneumoniae. *Br J Biomed Sci* 69: 123-125.

194. Ginsburg AS, Grosset JH, Bishai WR (2003) Fluoroquinolones, tuberculosis, and resistance. *Lancet Infect Dis* 3: 432-442.

195. Ferdinand S, Valetudie G, Sola C, Rastogi N (2004) Data mining of Mycobacterium tuberculosis complex genotyping results using mycobacterial interspersed repetitive units validates the clonal structure of spoligotyping-defined families. *Res Microbiol* 155: 647-654. doi:10.1016/j.resmic.2004.04.013

196. de Steenwinkel JE, ten Kate MT, de Knegt GJ, Verbrugh HA, Aarnoutse RE*, et al.* (2012) Consequences of noncompliance for therapy efficacy and emergence of resistance in murine tuberculosis caused by the Beijing genotype of Mycobacterium tuberculosis. *Antimicrob Agents Chemother* 56: 4937-4944. doi:10.1128/AAC.00124-12

197. Aguilar D, Hanekom M, Mata D, Gey van Pittius NC, van Helden PD*, et al.* (2010) Mycobacterium tuberculosis strains with the Beijing genotype demonstrate variability in virulence associated with transmission. *Tuberculosis (Edinb)* 90: 319-325. doi:10.1016/j.tube.2010.08.004

198. Oxlade O, Murray M (2012) Tuberculosis and poverty: why are the poor at greater risk in India? *PLoS One* 7: e47533. doi:10.1371/journal.pone.0047533

199. Kamineni VV, Wilson N, Das A, Satyanarayana S, Chadha S*, et al.* (2012) Addressing poverty through disease control programmes: examples from Tuberculosis control in India. *Int J Equity Health* 11: 17. doi:10.1186/1475-9276-11-17

200. Zhao P, Li XJ, Zhang SF, Wang XS, Liu CY (2012) Social behaviour risk factors for drug resistant tuberculosis in mainland China: a meta-analysis. *J Int Med Res* 40: 436-445.

201. Prado TN, Wada N, Guidoni LM, Golub JE, Dietze R*, et al.* (2011) Cost-effectiveness of community health worker versus home-based guardians for directly observed treatment of tuberculosis in Vitoria, Espirito Santo State, Brazil. *Cad Saude Publica* 27: 944-952**.**

Appendix 1: The PCR primer sets used for detection of mutations in respective drug target genes

Target Gene (Drugs)	Primer Direction	Primer Sequence in direction 5' to 3'	Size of PCR product (bp)
rpoB (RIF)	F	TCCTCGATGACGCCGCTTTCT	849
	R	TCR GAG ATC TTG CGC TTC TGS; R=G/A, S=G/C	
	R	AYATCGACCACTTCGGYAACC; Y=C/T	
katG (INH)	F	CCAGCGGCCCAAGGTATC	850
	R	GCTGTGGCCGGTCAAGAAGAAGTA	
inhA (INH)	F	GGCACGTACACGTCTTTATGTA	478
	R	GGTGCTCTTCTACCGCCGTGAA	
rpsL (STR)	F	CGTGAAAGCGCCCAAGATAG	375
	R	GAACCGCGGATGATCTTGTAG	
embB (EMB)	F	CGGCATGCGCCGGCTGATTC	260
	R	TCCACAGACTGGCGTCGCTG	
pncA (PZA)	F	GGCTGCCGCGTCGGTAGG	652
	R	GCCGCCAACAGTTCATCCC	
gyrA (FQ)	F	CAGCTACATCGACTATGCG	320
	R	GGCTTCGGTGTACCTCATC	
rrs (AMK)	F	CGTTCCCTTGTGGCCTGTG	547
	R	GGCGTTTTCGTGGTGCTCC	

Appendix 2: District wise frequency of TB cases enrolled in the study (N=650)

S. No	Origin of Patients (Districts)	n	Percent
1	KATHMANDU	366	56.31
2	LALITPUR	36	5.54
3	DHADING	27	4.15
4	MAKWANPUR	16	2.46
5	NUWAKOT	15	2.31
6	SINDHUPALCHOK	13	2.00
7	RAMECHHAP	12	1.85
8	KAVREPALANCHOK	10	1.54
9	CHITWAN	9	1.38
10	RUPANDEHI	9	1.38
11	SINDHULI	9	1.38
12	BHAKTAPUR	8	1.23
13	DOLAKHA	8	1.23
14	GORKHA	8	1.23
15	SARLAHI	7	1.08
16	KAILALI	6	0.92
17	BARA	5	0.77
18	PALPA	5	0.77
19	PARSA	5	0.77
20	SYANGJA	5	0.77
21	DANG	4	0.62
22	KAPILVASTU	4	0.62
23	SUNSARI	4	0.62
24	GULMI	3	0.46
25	KANCHANPUR	3	0.46
26	MORANG	3	0.46
27	RASUWA	3	0.46
28	RAUTAHAT	3	0.46
29	SIRAHA	3	0.46
30	SURKHET	3	0.46
31	BAGLUNG	2	0.31
32	BANKE	2	0.31
33	MAHOTTARI	2	0.31
34	MUSTANG	2	0.31
35	MYAGDI	2	0.31
36	OKHALDHUNGA	2	0.31
37	TANAHU	2	0.31
38	UDAYAPUR	2	0.31
39	ACHHAM	1	0.15

40	BAJHANG	1	0.15
41	BAJURA	1	0.15
42	BHOJPUR	1	0.15
43	DAILEKH	1	0.15
44	JHAPA	1	0.15
45	KASKI	1	0.15
46	KHOTANG	1	0.15
47	LAMJUNG	1	0.15
48	MANANG	1	0.15
49	MUGU	1	0.15
50	PANCHTHAR	1	0.15
51	ROLPA	1	0.15
52	SOLUKHUMBU	1	0.15
53	TAPLEJUNG	1	0.15
54	TERHATHUM	1	0.15
55	**INDIA (foreign)**	6	0.92
	TOTAL	650	99.95

Curriculum Vitae

Personal Information	
First Name/ Sur-Name	**BIJAYA MALLA**
Date of Birth	1978-Jan- 09
Nationality	Nepalese
E-mail	bijaya.malla@outlook.com
Education	
2009 --	**PhD in Natural Science**
Expected completion date	Swiss Tropical and Public Health Institute, University of Basel
2013	"Molecular Epidemiology of *Mycobacterium tuberculosis* in Nepal"
	Supervisor:
	Prof. Sebastien Gagneux
	Co- referee:
	Prof. Douglas B. Young
02/2008 - 05/2008	**Post-Graduate Diploma in International Health**
	Swiss Tropical and Public Health Institute, University of Basel
	Scholarships awarded:
	Amt für Ausbildungsbeiträge, Canton Basel-Stadt, Switzerland
2001 – 2003	**Master of Science in Microbiology**
	Tribhuvan University, Nepal
	Thesis on "Study on Nasopharyngeal Pneumococcal Carriage and Entero-Parasitic Infestations in Children"
1997 - 2000	**Bachelor of Science in Microbiology, Nepal**
1995 -1996	**Proficiency Certificate in Science (Biology), Nepal**
1993	**School Leaving Certificate, Nepal**
Work Experiences	
04/2006 - 01/2008	**Teaching Assistant**, National College, Tribhuvan University, Nepal
04/2005 – 01/2008	**Teaching Assistant**, Central Department of Microbiology, Tribhuvan University, Nepal
01/2008 – 01/2008	**Survey Coordinator (as a Consultant team member)** for baseline survey of community responsive antenatal, delivery, and life essentials (CRADLE) project funded by CARE, Nepal
10/2005 – 12/2008	**Research Officer** at Institute of Medicine, Tribhuvan University Teaching Hospital, Nepal
10/2004 – 09/2005	**Microbiologist** at Health Research laboratory, **Tribhuvan University Teaching Hospital**, Nepal
Trainings	
30/05/2011-08/06/2011	As **Trainer**; Training seminar on "Mycobacterium tuberculosis culture and drug susceptibility tests" at **National Reference Laboratory, German Nepal Tuberculosis Project, Nepal**
19 -- 28/05/2011	As **Trainer**; Training seminar on "Primary Culture of Mycobacterium tuberculosis" at **National Reference Laboratory, German Nepal Tuberculosis Project, Nepal**
23 -- 28/01/2011	Wellcome Trust Advanced course: Genomics and Clinical Microbiology at **Wellcome Trust Genome Campus, Hinxton, Cambridge, UK**
	Bursary Award
13 -- 15/12/2010	"Tuberculosis Diagnostic Research-beyond the basics" **, funded by SEARO, WHO,** at Tuberculosis Research Center, **Chennai, India**
	Discussion on implementation of new diagnostics at South Asia region

31/05 -08/06/2007	Training workshop on Epidemiology methods at Central Department of Microbiology, Tribhuvan University, **Nepal**
27-31/08/2006	Tuberculosis laboratory Modular Training based on National Tuberculosis Program module, **Nepal**
29/12/2003 – 02/01/2004	"Research Methodology and Ethics" at Nepal Health Research Council, **Nepal**

Participation Workshops/conferences	
29/10 –02/11/2012	Comparative genomics workshop, at **SIB bioinformatics, Lausanne, Switzerland**
29 – 31/08/2012	LSS2012—Global Health meets Infection Biology at **EPFL, Lausanne, Switzerland** **Poster Presentation on:** Previous treatment history and drug resistance is associated with the Beijing genotypes of *Mycobacterium tuberculosis*
27 – 29/08/2012	International BioCamp 2012: Novartis Biotechnology Leadership Camp **Novartis, Basel, Switzerland** **Scholarship award**
21 – 22/06/2012	Joint Annual Meeting, SSI, SSHH, SSM, SSTMP at St. Gallen, Switzerland **Poster Presentation on**: *Mycobacterium tuberculosis* complex Beijing genotype showed association with drug resistance and relapse in Nepal
22 – 24/11/2011 13 – 15/12/2011	Essentials in Drug Development and Clinical Trials at Swiss TPH, **Basel, Switzerland**
15 – 16/11/2011	Doctoral meeting of Swiss Society of Tropical Medicine and Parasitology (SSTMP) at **Basel, Switzerland** **Presented on** : Molecular Epidemiology of tuberculosis and Drug resistance in Nepal
26 -- 30/10/2011	42nd Union World Conference on Lung Health organized by **International Union Against Tuberculosis and Lung Disease** in **Lille, France** **Presented on:** "First insights into the genetic diversity of *Mycobacterium tuberculosis* in Nepal"
20 -- 22/09/2011	"BioValley life Sciences Week 2011" organized at Congress Center **Basel, Switzerland**
09 – 13/08/2011	"Multinational Influenza Seasonal Mortality Study (MISMS) South Asia" organized by **Fogarty International Center of the U.S. National Institutes of Health (NIH)** in **Nepal**
26/03 – 03/04/2007	International workshop on Kala-azar organized by **WHO/TDR** & **Banaras Hindu University, Varanasi, India**
28/02 – 02/03/2006	**Proposal writing** workshop on "Cost effective Integrated Vector Management as a contribution to the Visceral Leishmaniasis Elimination Initiative on Indian Subcontinent" in **Kolkota, India,** organized by **WHO/TDR, Geneva, and Institute of Medicine, Nepal**
14 – 17/12/2004	First SAARC Conference on TB, HIV/AIDS & Respiratory Diseases Organized by SAARC Tuberculosis Centre, **Nepal**

Languages	
Newari	Mother tongue
Nepali	Mother tongue
English	Fluent

Computer skills	
Windows7/MS Office	Very Good
STATA 10 / R Package	Working proficiency

Linux	Limited working proficiency
Bioinformatics	Experiences on Bioinformatics
Extra-curricular activities	Community service, Debate activity, Music
18 -- 20/04/2012	Participated at "Geneva Health Forum, Geneva" to **promote Swiss TPH;** organized with Medicus Mundi, Switzerland

Publications

- **Malla B**, Fenner L, Gagneux S., *et al*, "First insights into the phylogenetic diversity of *Mycobacterium tuberculosis* in Nepal." ***PLoS One*** Dec 2012.
- Stucki D., **Malla B**, Gagneux S., *et al*,"Two new rapid SNP-typing methods for classifying *Mycobacterium tuberculosis* complex into the main phylogenetic lineages." ***PLoS One***. 2012;7(7):e41253. Epub 2012 Jul 20**.**
- Comas I., Borrell S., **Malla B.**, Gagneux S., *et al*, "Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes." ***Nature Genetics*** 44, 106–110 (2012) doi:10.1038/ng.1038
- Fenner L, **Malla B**, *et al*. "Pseudo-Beijing": evidence for convergent evolution in the direct repeat region of *Mycobacterium tuberculosis*. " ***PLoS One***. 2011;6(9):e24737.
- D.R. Joshi, D. B. Khatry, K. Rosyara, **B. Malla**. "Do Behaviour Patterns of University Students of Nepal Make Them Vulnerable to HIV/AIDS?" ***JIOM***, 2007;29(3): 13-17
- K.P. Rosyara, **B. Malla**. "Community Health Programmes for Health Promotion in Rural Community of Nepal." ***JIOM*** 2007; 29(3): 32-35
- **Malla B**., Editorial: "Integrated Vector Management as a Contribution to Vector Borne Disease Elimination in Nepal "***J NHRC****, Nepal."* 2006; 4:2
- **Malla B**, Sherchand J. B., BC Rajendra Kumar., Ghimire P, Rijal BP., "Prevalence of Pneumococcal Carriage and Antimicrobial Susceptibility Pattern of Streptococcus Pneumoniae Isolates." ***SAARC J TB L DIS & HIV/AIDS*** 2005; 2: 6-8.
- **Malla B**, Sherchand J. B., Ghimire P, BC Rajendra Kumar., Gauchan P.,"Prevalence of Enteric Parasitic Infections and Malnutrition among Children in a Rural Community of Sarlahi, Nepal". *J NHRC, Nepal.* 2004; 2: 55-58.