Molecular characteristics of drug-resistant *Mycobacterium tuberculosis* in Thailand

（タイにおける薬剤耐性結核菌の遺伝学的特性）

Janisara Rudeeaneksin

CONTENTS

ABBREVIATIONS

AMR	antimicrobial resistance
AMI	amikacin
Asp	aspartate
Arg	arginine
Asn	asparagine
BDQ	bedaquiline
BCG	Bacillus Calmette-Guerin
bp	base pair
CAP	capreomycin
CIP	ciprofloxacin
DR	direct repeat
DOTS	directly observed treatment short-course
DST	drug-susceptibility testing
DNA	deoxyribonucleic acid
EMB	ethambutol
FLQ	fluoroquinolone
Glu	glutamate
Gly	glycine
HGDI	Hunter-Gaston Discriminatory Index
HIV	human immune deficiency virus
INH	isoniazid
IGRAs	interferon-γ release assays
KAN	kanamycin

LAM	lipoarabinomannan
LPA	line probe assay
Ile	isoleucine
Leu	leucine
LFX	levofloxacin
LZD	linezolid
LSP	large sequence polymorphism
MDR	multidrug-resistant
MIRU	mycobacterial interspersed repetitive unit
MST	minimum spanning tree
MTB	*Mycobacterium tuberculosis*
NAD	nicotinamide adenine dinucleotide
OFX	Ofloxacin
PCR	polymerase chain reaction
Pro	proline
QRDR	quinolone resistance determining region
RD	region of difference
RFLP	restriction fragment length polymorphism
RIF	rifampicin
RNAP	RNA polymerase
RRDR	rifampicin-resistance determining region
RR	rifampicin-resistant
STM	streptomycin
Ser	serine

SIT	spoligo-international type
SNP	single nucleotide polymorphism
TB	tuberculosis
Tyr	tyrosine
Val	valine
VNTR	variable number of tandem repeat
UPGMA	unweighted pair group method with arithmetic means mode
WGS	whole genome sequencing
WHO	World Health Organization
XDR	extensively drug-resistant

PREFACE

Tuberculosis (TB), caused by the bacillus, *Mycobacterium tuberculosis* (MTB), is a serious health-deteriorating disease and one of the leading causes of deaths worldwide (83). MTB primarily attacks the lungs causing pulmonary TB and it can cause extrapulmonary TB which affects any part of the body other than the lungs, such as lymph nodes, pleura, genitourinary tract, and the skeletal system (54). The infected individuals can either progress into the active diseases or latency stage which is asymptomatic and non-transmissible TB infection (45). It was estimated that a quarter of the global population has been latently infected with TB (83). Generally, about 5-10 percent of latent TB infection (LTBI) can progress into active TB disease during their lifetime while patients with pulmonary TB disease commonly transmit the disease to others (45). TB manifests in a dynamic spectrum from asymptomatic infection to a life-threatening disease. Active TB patients usually show general symptoms such as fever, fatigue, lack of appetite, and weight loss, and those with pulmonary TB disease can have additional signs such as night sweats, sputum production, persistent cough or coughing up blood, and chest pain. Patients with advanced diseases may have fatal outcomes by such as lung dysfunction through cavity formation (45).

Global situation of TB

The World Health Organization (WHO) reported an estimated nearly 10.6 million incidence cases (134 cases/100,000 inhabitants), 1.6 million deaths, and 0.5 incident drug-resistant (DR-TB) cases including multidrug-resistant/rifampicin-resistant TB (MDR/RR-TB), and extensively drug-resistant TB (XDR-TB) in 2021 (83) where MDR/RR-TB is defined as TB that is resistant to at least two important first-line TB drugs, isoniazid (INH) and rifampicin (RIF). XDR-TB, a more threathening TB form, is defined as MDR/RR-TB with additional resistance to any of fluoroquinolones (FLQs) and at least one group A drugs which comprises levofloxacin (LFX), moxifloxacin (MFX), bedaquiline (BDQ) and linezolid (LZD) (84). The incidence rate of TB declines slowly by about 2% each year and the disease remains a public health problem throughout the world. Figure 1 shows the global trends of the annual estimated TB incidence rate from 2010 to 2021. The cumulative reduction in the estimated TB incidence rate between 2015 and 2020 was 13.5%, lower than the first End TB Strategy target milestone of 20% reduction (83). On the other hand, the reduction rate in TB deaths from 2015 to 2019 was 14% worldwide less than the first milestone of the End TB Strategy, which targets a 35% reduction in the total number of TB deaths (82).

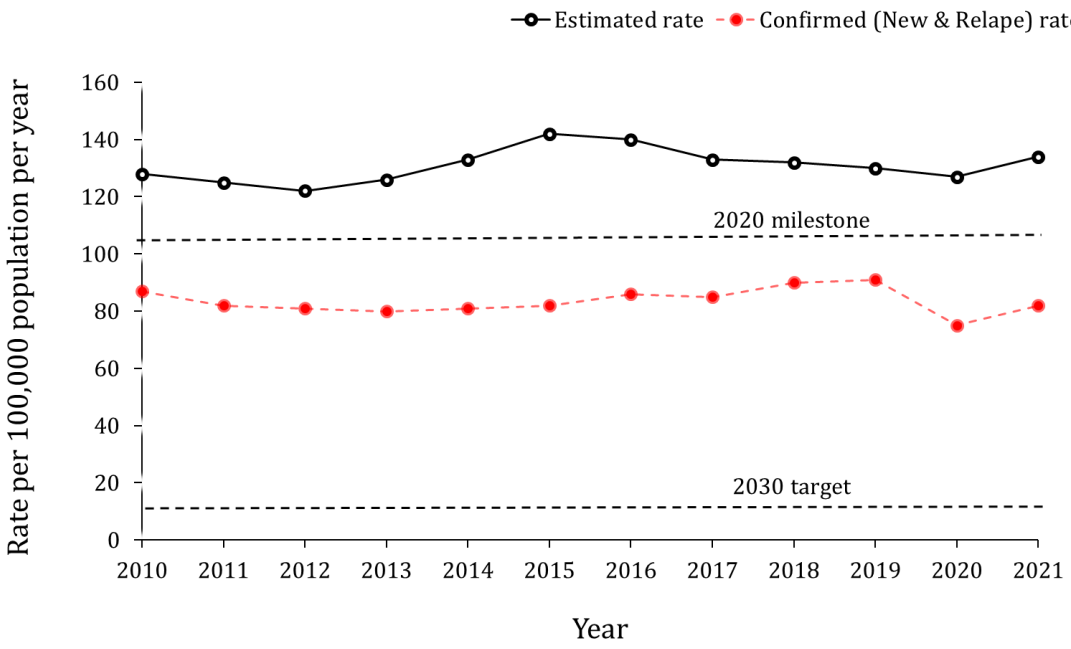


Figure 1. Global trends of the estimated TB incidence rate compared with confirmed new and relapse rates between 2010- 2021 (77-83).

TB situation in Thailand

TB continues to be an important public health problem in Thailand, a high TB burden country, with an estimated 120,000 new cases every year and ranked among the top 30 high TB burden countries. In 2021, the year of the COVID-19 pandemic, the estimated number of TB patients dramatically decreased with TB cases of 103,000 (143 cases/100,000 inhabitants), MDR-TB of 2,400 incident cases, and 11,400 mortality cases (84). Figure 2 shows the estimated TB incidence rate in Thailand from 2010 to 2021. There was a continuous decline in the TB incidence rate from 171 down to 150 cases per 100,000 inhabitants between 2015 to 2020 (77-83). Although Thailand has adopted the Operational Project of End TB to reduce the incidence of TB by 12.5% per year, from 171/100,000 inhabitants in 2014 to 88/100,000 inhabitants by the end of 2021, the trend of TB incidence rate in Thailand declined slowly only by 2.5% per year from 2015 to 2019 (73). The reduction in TB incidence rate is still far from the first milestone of the End TB Strategy.

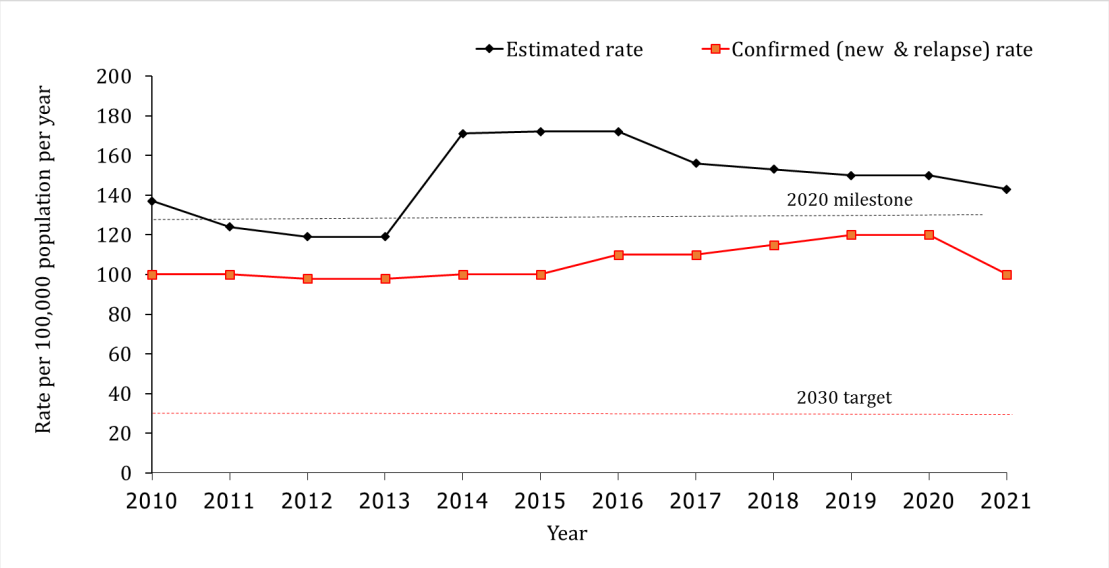


Figure 2. The trends of the estimated TB incidence rate in Thailand from 2010 to 2021 (77-81)

Diagnosis of TB

TB diagnosis depends on testing for detecting latent TB infection, active TB disease, or drug resistance (45). Two types of imunological tests; the tuberculin skin test and the interferon-γ release assays (IGRAs) has been used for detecting latency TB. The IGRAs have an advantages that the tests can distinguish MTB infection from Bacillus Calmette-Guerin (BCG) induced positive response in BCG vaccination and non-tuberculous mycobacterial infection. It can be used as a supporting tool for the diagnosis of extrapulmonary TB or some pulmonary cases with difficulty in sputum collection. For the diagnosis of active TB disease, the chest radiograph (CXR) has been used for the initial screening. Other methods for active TB diagnosis and drug resistance detection require microbiological procedures as follows.

1) Sputum microscopy; Conventional and LED fluorescence smear microscopy (endorsed by WHO in 2011) has been used for rapid screening of tubercle bacilli.

2) Culture-based techniques; Solid and liquid culture (The mycobacterial growth indicator tube (MGIT) automated culture system endorsed by WHO in 2007) are the gold standard to detect MTB and drug resistance.

3) Antigen detection technique: The lipoarabinomannan (LAM) lateral flow (endorsed by WHO in 2015) has been used to detect MTB-LAM antigen in urine for specific diagnosis of TB among individuals coinfected with human immune deficiency virus (HIV).

4) Molecular techniques (nucleic acid amplification tests):

a) Xpert MTB/RIF (endorsed by WHO in 2010); It has been used for the rapid detection of MTB and RIF-R.

b) Line probe assay (LPAs); first- and second-line probe assays (i.e., GenoType MTBDR*plus* V2.0 and GenoType MTBDR*sl* V2.0 assays endorsed by WHO in 2008 and 2016, respectively) has been used for the detection of drug resistance in which specific resistance-conferring mutations.

c) The Loopamp™ MTBC assay (endorsed by WHO in 2016) has been used for the rapid detection of MTB as an add-on or a replacement test of smear microscopy.

Treatment of TB

TB treatment has been delivered by the Directly Observed Treatment Short course (DOTS) for at least 6 months (62). The regimen with the first-line drugs in the initial phase compose a 2-month of INH, RIF, pyrazinamide (PZA), and ethambutol (EMB) to kill the majority of mycobacteria followed by the continuation phase with a 4-month sterilization of INH and RIF to eliminate the remaining persistent MTB populations (Table1). By using multi-drug regimens and the DOTS program, the treatment success rate for TB worldwide was 85%. When the bacterial strains become resistant to at least INH and RIF, second-line drugs which included fluoroquinolones and injectable aminoglycosides are applied (54). In contrast, the global success rate of the MDR-TB treatment was only 57% indicating the significant problem of drug resistance in TB (83). Recently, WHO recommended new second- drugs consisting of Group A: LFX, MDF, BDQ, LZD; Group B: clofazimine, cycloserine, or terizidone; and Group C: EMB, delamanid, PZA, imipenem-cilastatin or meropenem, amikacin (AMI) or prothionamide, *p*-aminosalicyclic (84). Also, Thailand has employed the DOTS program as a part of the TB control strategy. The treatment outcome data showed 83% for TB (new and relapse) while the MDR-TB treatment success rate was 60% in 2019 (83).

Table 1. Standard first-line treatment regimens for TB (60)

Ranking	The first phase (Initiation)	The second phase (Continuation)
Preferred drugs	INH, RIF, PZA, EMB daily for 2 months	INH, RIF daily for 4 months
	INH, RIF, PZA, EMB 3 times a week for 2 months	INH, RIF 3 times a week for 4 months
Optional drugs	INH, RIF, PZA, EMB daily for 2 months	INH, EMB daily for 6 months

DR-TB treatment and challenges

DR-TB is a cornerstone of the global antimicrobial resistance (AMR) challenge (24). The deaths from DR-TB are accounted for about one-third of all antimicrobial resistance deaths worldwide. Every year, about 7.0 millions people die of resistant infections (11). Importantly, more people die from tuberculosis than other infectious diseases; 1.6 million die of TB per year, and of these, 0.2 million die of MDR-TB (83). In 2021 WHO reported that the cumulative number of people enrolled in MDR/RR-TB treatment from 2018 to 2021 was 0.6 million worldwide, which was only 43% of the 5-year target of 1.5 million set at the UN high-level meeting on TB in 2018 (83). The annual outcome of the treatment success rate for MDR/RR-TB from 2012 to 2019 has been improved slowly from 52% to 60% (Figure 3) .

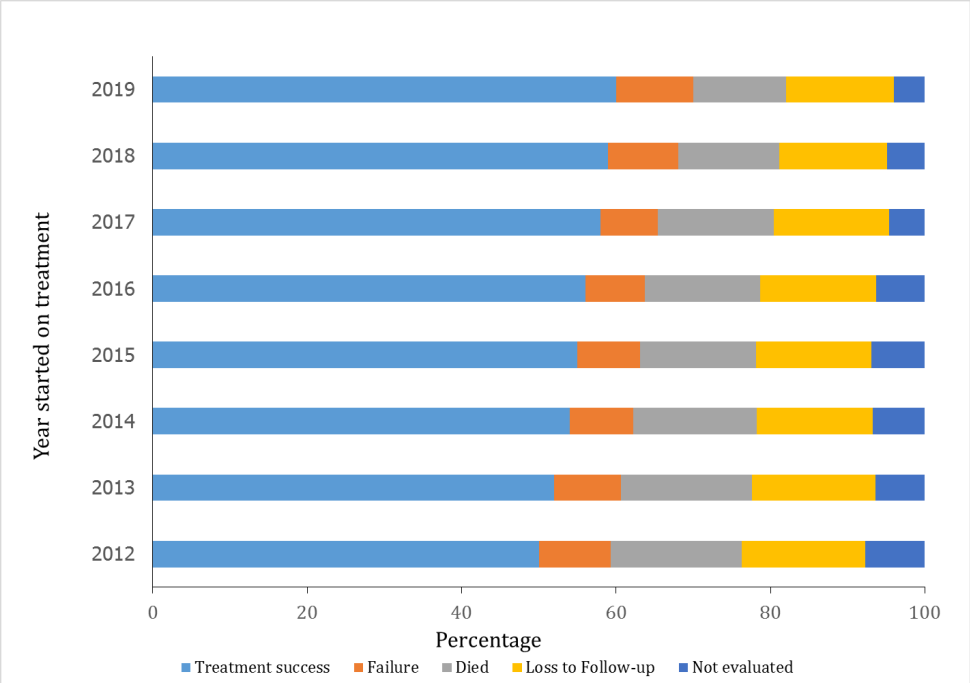


Figure 3. The global outcomes for MDR annual treatment reporting on treatment success, treatment failure, lossto follow-up of treatment, and patient deaths (82)

The treatment of TB is difficult and much more complicated in MDR/XDR-TB. The course of MDR/XDR-TB treatment is longer up to two years and requires second-line drugs that are more expensive (≥US$ 1,000 per person), poorly efficacious, and more toxic than drug regimens for drug-susceptible TB (62, 82). The treatment outcomes among patients receiving second-line drugs are associated with frequent and severe side effects and were suboptimal, resulting in increased risk for mortality, and transmission of drug resistance. Gandhi *et al.,* 2018 (17) reported a less success rate of the outcomes of MDR/XDR-TB treatment. The evidence showed that the new regimens consisting of BDQ, pretomanid, and LZD advanced and shorten the treatment duration of MDR/XDR-TB from 24-18 to 6-9 months with safety and efficacy (17, 24, 64). However, some concerns are still present in the adoption of new regimens into the clinical practice and the TB control program.

HIV infection is one of the high-risk factors for progressing all forms of TB. MDR/XDR-TB with HIV-positivity has shown poor treatment outcomes and high mortality rates due to a pill burden, drug interactions, additive adverse effects, and a long course of treatment (63). Moreover, HIV-TB patients usually exhibit drug resistance in a form of pauciacillary TB. Therefore, the diagnosis and treatment of TB in HIV-TB are challenged.

Modes of action of DR and mechanisms of DR

The modes of action of the two core first-line drugs, INH and RIF (Figure 4), and the two key second-line drugs, FLQs and injectable agents, for TB treatment were shown in Figure 4

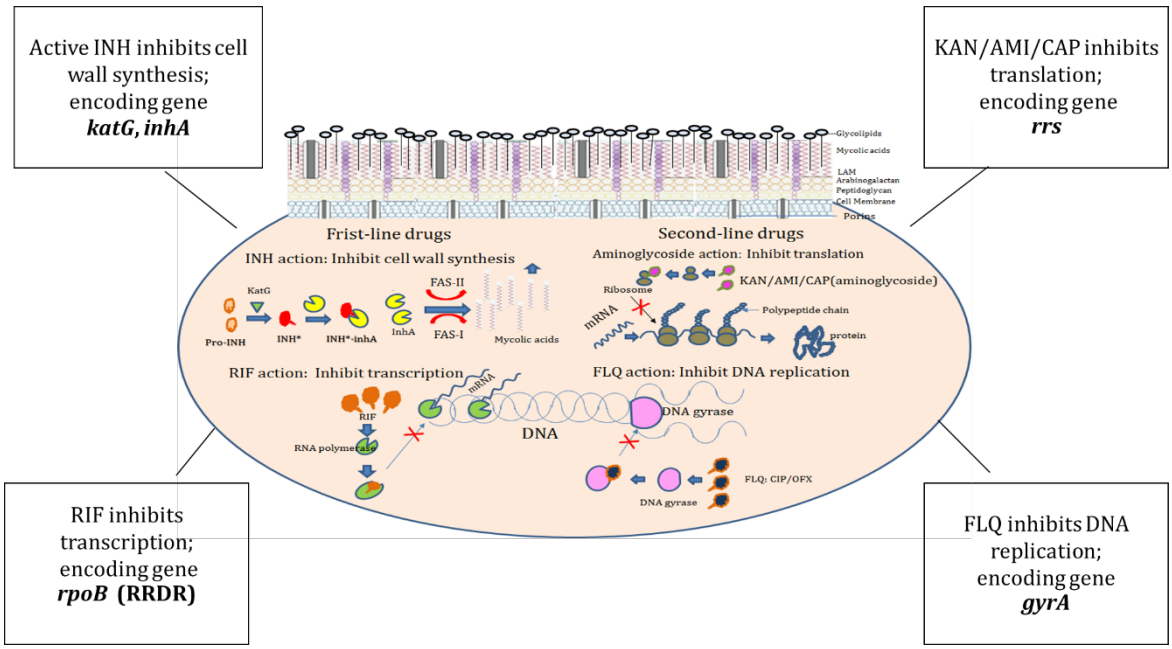


Figure 4. Modes of actions of the first-line (INH, RIF) and the second-line drugs (FLQ and injectable; KAN/AMI/CAP/VIO)

INH

INH is a pro-drug requiring the activation by the catalase/peroxidase encoded by the *katG* gene. Unissa *et al*., 2016 (74) reported that the catalase carried out the oxidation of INH. The pro-INH is metabolized to electrophilic radicals (isonicotinyl acyl radical) which can exert a toxic effect on bacilli. Another mechanism of action is the inhibition of mycolic acid synthesis via the NADH (β-Nicotinamide adenine dinucleotide)-dependent enoyl-acyl carrier protein (ACP) reductase (InhA) and beta-ketoacyl carrier protein synthase (KasA)(Figure 5) (74). The InhA is encoded by the *inhA* gene, and acts as a target for the activated INH. The activated INH or isonicotinic acyl radical also interacts with NADH in the presence of InhA and forms a ternary complex (InhA isonicotinyl-NADH) which interferes with the elongation of fatty acids and eventually blocks the biosynthesis of MA. This INH adduct, therefore, can kill bacilli by inhibiting cell wall synthesis (Figure 5).

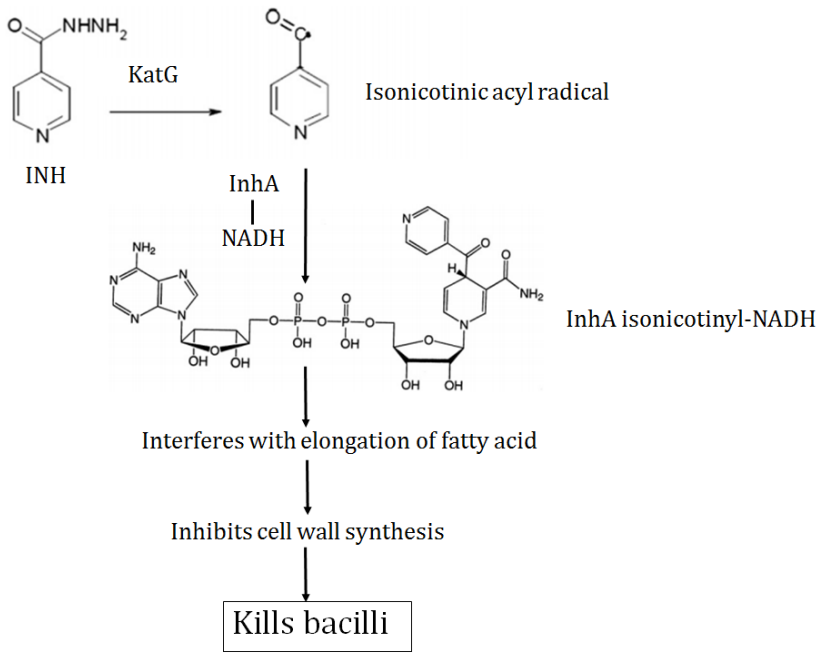


Figure 5. Mechanisms of action of INH in killing tubercle bacilli (74)

Mutations in the *katG* are the major cause of INH resistance. The *katG* mutations diminish catalase and peroxidase activity leading to loss of INH activation. It was reported that mutations in the *katG* were associated with high-level resistance to INH. The most frequent mutation in the *katG* is Ser315Thr, which leads to inefficient isoniazid–NAD production to exhibit the action of isoniazid (51, 52). In addition, point mutations in the regulatory region of the *inhA* promoter, mainly at positions -15 and -8 resulted in the over expression of *inhA*, which confers low-level resistance to INH (51, 52). Other mutations in the active region of the *inhA* gene can also lead to the modification of the InhA target resulting in a decreased affinity of the isoniazid-NAD adduct. Approximately, 75-90% of INH resistance is attributed to mutations in the *katG* gene and the *inhA* promoter (7, 40, 51). However, the molecular mechanisms of isoniazid resistance may associate with mutations in other genes such as *ahpC, kasA*, *ndh*, *iniABC*, *fadE*, *furA*, *Rv1592c,* and *Rv1772* which involve multiple biosynthetic networks or pathways (46, 74). These mutations are not common and little evidence to support the mechanisms of INH resistance. It is also possible that multiple mutations occur and combine to produce INH-phenotypic resistance.

Table 2. The distribution of mutations in the most highly mutated regions in the *katG*, and in the regulator region of the *inhA* genes involved in INH resistance (Modified from Unissa *et al.,* 2016) (74)

Genes (codon)	Protein	Nucleotide change			Amino acid Change
katG	Catalase/peroxidase enzyme				
315		AGC	→	ACC	Ser315Thr
			→	AAC	Ser315Asn
			→	ATC	Ser315Ile
			→	CGC	Ser315Arg
			→	GGC	Ser315Gly
328		TGG	→	GGG	Trp328Gly
			→	TTG	Trp328Leu
			→	TGC	Trp328Cys
350		GCT	→	TCT	Ala350Ser
381		GAC	→	GGC	Asp381Gly
inhA	NADH-enoyl-ACP-reductase				
		-8 T	→	G/A	-
		-15 C	→	T	-
		-16 A	→	G	-

RIF

RIF is an important bactericidal agent against several types of bacteria including mycobacteria. The mode of action of RIF in MTB is by binding to the β-subunit of the DNA-dependent RNA polymerase (RNAP), resulting in preventing the RNA transcription by inhibiting the elongation of messenger RNA, and the subsequent protein synthesis (51-52). The majority of RR-MTB clinical isolates harbor mutations in the *rpoB*. As a result, conformational changes occur leading to a decrease in the affinity for the drug resulting in the development of resistance. In addition, greater than 90% of RR strains are also resistant to INH (88). It was found that up to 95% of RR strains have mutations in 81 bp of the RIF resistance determining region (RRDR) in the *rpoB*, which is considered a mutation hotspot region spanning codons 426-452. (corresponding to *Escherichia coli* number 507-533) (Table 3) (42, 51, 65). Approximately 5% of clinical RR strains have no mutations in the RRDR (61, 65). The other mechanism of RIF resistance is through the efflux pump that is associated with the overexpression of specific genes. However, it is still not clear whether these individual genes involve in RIF resistance, and hence their functions are needed to be

further determined in RIF-resistant phenotype among a large number of MTB isolates (31, 33, 40).

Table 3. Mutations in *rpoB* codons 507 to 533 (*Escherichia coli* corresponding number) associated with RIF resistance (42).

Codon (MTB codon No)	Genes	Protein	Nucleotide change	Amino acid change	Codon (MTB codon)	Gene	Protein	Nucleotide change	Amino acid change
	rpoB	RNAP				*rpoB*	RNAP		
507 (426)			GGC → GAC	Gly507Asp	517 (436)			del CAG	
			del GGC ACC					del CAG AAC	
510 (429)			CAG → CAT	Gln 510His	518 (437)			AAC → CAC	Asn518His
511 (430)			CTG → CCG	Leu511Pro				del AAC	
			→ CGG	Leu511Arg				del GAACAA	
512 (431)			AGC → ACC	Ser512Thr	521 (440)			CTG → ATG	Leu521Met
			→ CGC	Ser512Arg				→ TTG	Leu521Leu
513 (432)			CAA → CTA	Gln513Leu	522 (441)			TCG → TTG	Ser522Leu
			→ AAA	Gln513Lys	526 (445)			CAC → TAC	His526Tyr
			→ CCA	Gln513Pro				→ GAC	His526Asp
			InsTTC					→ CGC	His526Arg
			InsTTC ATG					→ CTC	His526Leu
			Det CAA TTC					→ CAA	His526Pro
514 (433)			TCC → TTG	Phe514Leu				→ AAC	His526Glu
			del AATTCATTGG					→ CAG	His526Asn
515 (434)			ATG → ATA	Met515Ile				→ GGC	His526Gly
			→ GTG	Met515Val				→ ACC	His526Thr
516 (435)			GAC → GTC	Asp516Val				→ TGC	His526Cys
			→ TAC	Asp516Tyr	531 (450)			TCG → TTG	Ser531Leu
			→ GAG	Asp516Glu				→ TGG	Ser531Trp
			→ GGC	Asp516Gly				→ TGT	Ser531Cyn
			→ GCC	Asp516Ala				→ CAG	Ser531Gln
			del GAC CAG					→ TAT	Ser531Tyr
					533(452)			→ CCC	Leu533Pro

FLQ

FLQ drugs are bactericidal and have been used against various bacteria including MTB (7, 13). ciprofloxacin (CIP) and ofloxacin (OFX) are synthetic derivatives of FLQ that are currently used as second-line or alternative TB drugs. The target of FLQ is DNA gyrase which consists of two A and two B subunits encoded by the *gyrA* and the *gyrB*, respectively. The interaction between FLQ and DNA gyrase occurs in a conserved region of the *gyrA* between amino acids codons 74 to 113 and the *gyrB* between amino acids codons 500 to 538 known as the quinolone resistance-determining region (QRDR) (7, 13). Thus, FLQ acts by interfering with mycobacterial DNA replication and transcription (Figure 4). The most common mutations associated with high-level resistance to FLQs are present at codons 88, 90, 91, and 94 in the QRDR encoding subunits of the gyrase A (GyrA) (Table 4). In addition

to genetic mutations, some mechanisms of drug accumulation confer FLQ resistance, either by a decreased uptake or by an increased efflux .

Injectable drugs

KAN, AMI, and CAP are bactericidal to treat bacterial infections including TB (64) and are classified as second-line injectable agents for MDR/XDR-TB treatment. KAN and AMI are aminoglycoside antibiotics that inhibit protein synthesis by inhibiting the function of ribosomes. Mutations in *rrs* (encodes 16S rRNA) were analyzed in MTB resistant to KAN and AMI. Suzuki *et al.,* 1998 (72) described the mechanisms of KAN-resistance which caused by the modification of the rRNA. The mutations at positions 1405 or 1408 were responsible for kanamycin resistance and the other mechanism involved the nucleotide changes in the 3′ part of the 16S rRNA gene (*rrs*). The main mutations in KAN resistance had been found in the *rrs* gene, especially at positions 1400, 1401, and 1483 (7, 72). Table 4 summarizes the major mutations found in FLQ-, AMI-, KAN-, and CAP-resistant isolates. It was reported that the *rrs* mutation was found in high-level AMI-, KAN- and CAP-resistant strains (51) such as the *rrs* mutation G1484T (7). The mutations in the *rrs* gene might be involved in AMI-, KAN-, and CAP- cross-resistance (7). Another mechanism conferring CAP resistance involves mutations in the *tlyA* which encoded the rRNA methyltransferase responsible for methylation of the ribosome. However, the *tylA* mutations were identified in a few isolates (7).

Table 4. The common mutations associated with FLQ and AMK, KAN, and CAP resistance in *gyrA* and *rrs* (36, 42, 51).

Genes (codon)	Protein	Nucleotide change	Amino acid change
FLQ			
gyrA	DNA gyrase (QRDR)		
88		GGC → TGC	Gly88Cys
90		GCG → GTG	Arg90Val
91		TCG → CCG	Ser91Pro
94		GAC → AAC	Asp94Asn
		→ GGC	Asp94Gly
		→ GCC	Asp94Ala
		→ TAC	Asp94Tyr

			→	CAC	Asp94His
AMK KAN CAP					
rrs	16S rRNA				
1401		A	→	G	-
1484		G	→	T	-
1402		C	→	T	-

Molecular epidemiological analysis of TB

Molecular typing methods based on the analysis of genomic polymorphisms are widely used to differentiate MTB strains which are valuable for understanding the evolutionary history, determination of genetic diversity, and estimation of transmission including patterns of dissemination of MTB (7, 10, 14, 15, 18, 41). It is a useful tool for identifying MTB strains, investigating possible epidemiological links between TB patients, detecting suspected and/or unsuspected outbreaks including the source of transmission, and distinguishing exogenous re-infection from endogenous reactivation in relapse cases (69). Several methods have been employed for various epidemiological purposes as well as for clinical management and molecular epidemiological investigations (76). The traditionally methods have been used for genotyping such as the restriction fragment length polymorphism (RFLP), spoligotyping, and the mycobacterial interspersed repetitive units-variable number of tandem repeats (MIRU-VNTR) analyses (28).

IS*6110*-based RFLP has been considered, until recently, to be the conventional standard method of molecular epidemiologic studies owing to its discriminatory power (ability to differentiate between two unrelated strains) and its wide use since the early 1990s. IS*6110* belongs to the IS*3* family of mobile elements. It has 1,361 base pairs and is found only in organisms of the MTB complex. It has been proposed that MTB contained between 0 to 25 copies of IS*6110* (37). Although IS*6110* does not have a known target for the insertion, it is believed that the insertion sites are not random. Several preferential integration loci or hotspots have been reported (14). The polymorphisms were generated by the variations of copy number and the insertion sites of IS*6110*. However, this method is time-consuming and the result comparsions between laboratories are difficult. In addition, unclear images with many bands of different intensities, and the complexity of their reading and interpretation are limitations.

Spoligotyping

Spoligotyping is a simple and rapid genotyping method that focuses on the detection of polymorphisms based on the presence or absence of 43 unique spacer sequences in the direct repeat (DR) region(Figure 6.)(12, 28, 55, 58, 66). Each repeat makes up a constant and a variable part in genome of MTB complex. The constant part is represented by identical repeated sequences of 36 bp length (DR, direct repeats) interspaced by unique variable sequences (spacers) of 35-41 bp length that generate polymorphisms. The spoligotyping technique is based on PCR and hybridization. A single couple of primers (Dra-Drb) allows PCR amplification of a set of overlapping fragments that will be further hybridized to a set of 43 immobilized oligonucleotides on a membrane. The hybridized amplicons will be conjugated with streptavidin-peroxidase and then detected signals (28). The spoligotyping has been used to categorize MTB strains into several families, such as Beijing, East-African Indian (EAI), Central Asian (CAS), Latin American Mediterranean (LAM), Unclassified (U), and T families (8). Details of the detection of the polymorphism by spoligotyping was presented in Figure 6.

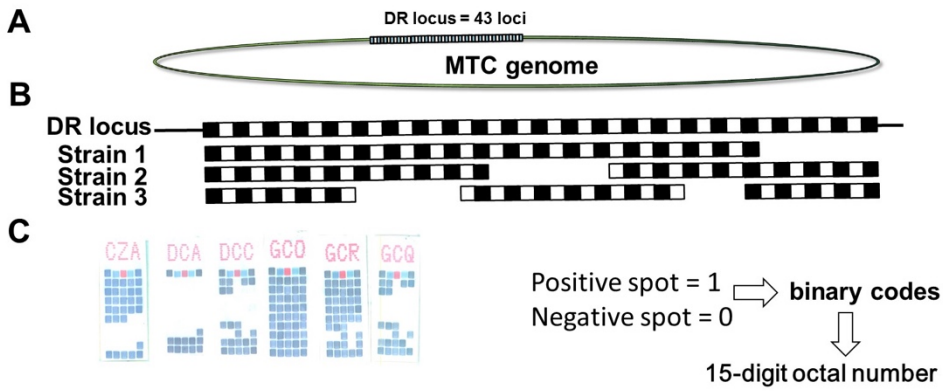


Figure 6. Spoligotyping, (A) The MTC genome and 43 unique spacer sequences in the direct repeat (DR) region, (B) Polymorphisms of spoligotyping based on the presence or absence of 43 unique spacer sequences (C) Results of spoligotyping method on a set of 43 immobilized oligonucleotides membrane; dark band indicating presence of spacer > positive spot =1 , but no band incicating the spacer's absence > negative spot = 0 for generating binary codes and the final 15-digit octal number

MIRU-VNTR typing

VNTR is a short nucleotide sequence (longer than 3) that is repeated and adjacent to each other organized as a tandem repeat, and often located in intergenic regions or coding sequences in several kinds of genomes including bacterial genomes. VNTRs can be subdivided into microsatellites (repeat sequences larger than 5 nucleotides) and minisatellites (repeat sequences lower than 5 nucleotides). The

variations in the number of VNTR alleles generate polymorphisms in different microorganisms. The mycobacterial interspersed repetitive unit-variable number of tandem repeats (MIRU-VNTR) is the repetitive units containing 40–100 base pairs in length and are located in different loci scattered throughout the genome of MTB (37, 69). MIRU-VNTR typing is a technique to analyze the repeat number of selected MIRU-VNTR per locus in MTB. The original MIRU-VNTR method includes 12-locus- or 24-locus MIRU-VNTR (15, 69). However, some investigators reported the use of a different set of MIRU-VNTR (37, 69).

The MIRU-VNTR typing is PCR-based amplification using primers specific for the flanking regions of each repeat locus. The protocol includes the amplification of each locus and the visualization of each amplified product in an electrophoresis gel (Figure 7). Because the length of the repeat unit is known, the size of the PCR product is considered for the number of copies of the repeat unit (37, 69). The result is a numerical value that reflects the number of repeats in tandem repeat at each locus. At present, there is a high-throughput MIRU-VNTR typing method based on multiplex PCR in which one primer of each primer set (Appendex A) is tagged with a different fluorescent dye. Therefore, multiple MIRU-VINTR loci could be amplified at the same time with the advanatage of primer labelling and capillary electrophoresis that allow multiple amplicons to be analyzed in one run.

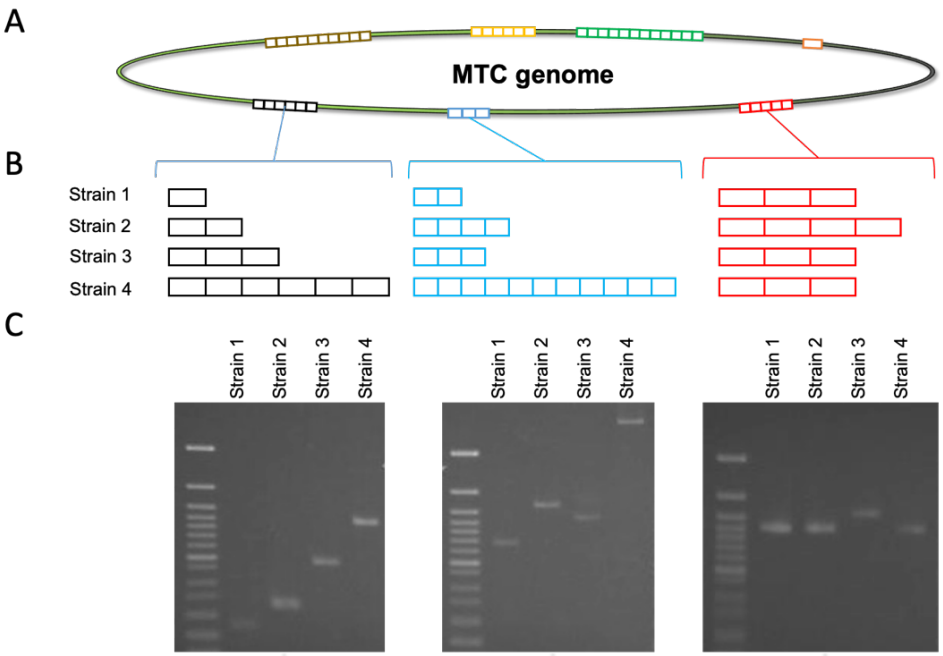


Figure 7. Differentiation of MTB strains based on MIRU-VNTR genotyping of MTB. a) MIRU-VNTR loci ongenome of MTB; b) Different sizes of amplicons of PCR product; C) Amplicons after gel electrophoresis with different sizes.

MIRU-VNTR typing is rapid, reproducible and relatively simple to perform with inexpensive cost to produce results in digital profiles. This method can be used alone or combined with other typing methods for the high resolution of genetic characterization. Generally, MIRU-VNTR typing has been used after spoliogotyping to further subtype major MTB groups. The combination of spoligotyping and the MIRU-VNTR typing method has become a practical tool providing a resolution high enough for the clear identification of MTB strains useful for monitoring the global transmission of TB (Figure 8) (4, 68-69).

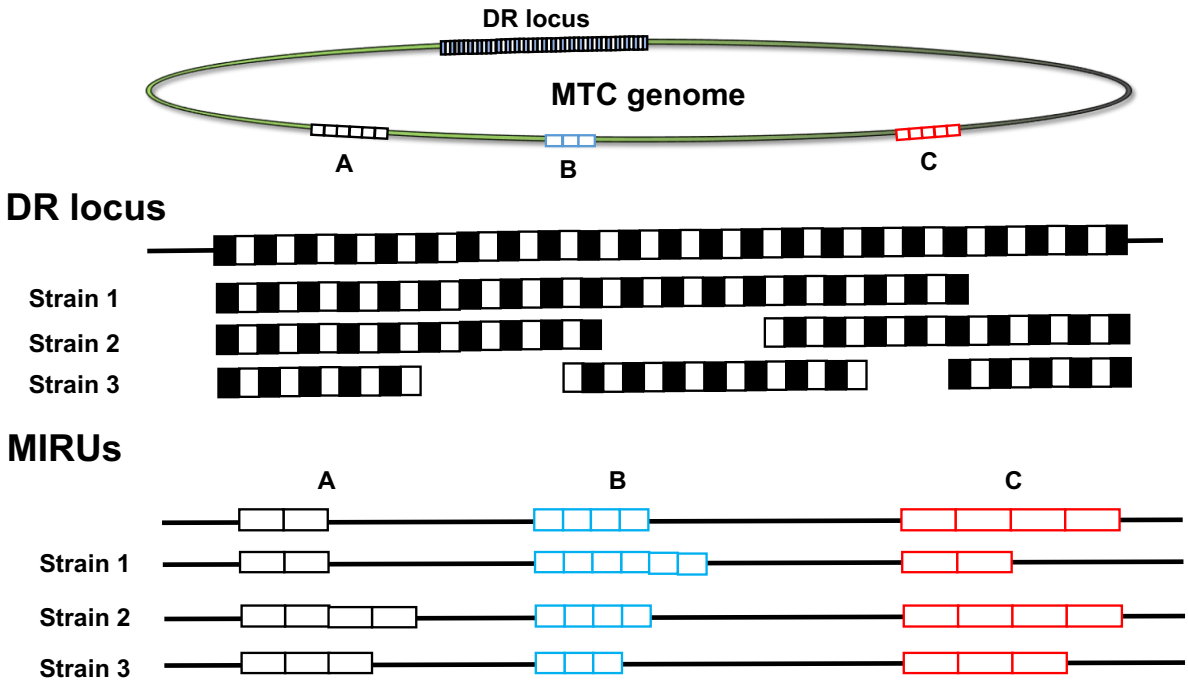


Figure 8. Diagram presenting spoligotyping and MIRU-VNTR-typing. Spoligotyping patterns are represented by color black and white squares indicating the presence or absence of direct repeat and particular spacers, respectively. The deletion of some of these 43 spacers allows differentiation of MTB strains. The MIRU-VNTR analysis relies on the identification of the different numbers of repeats at several loci scattered around the bacterial genome (marked by A, B, and C in the figure). The number of repeats at each locus is combined to generate a unique numerical code used to establish phylogenetic and epidemiological links between strains (8).

Whole genome sequencing

Recently, whole genome sequencing has been used to genetically characterize MTB (53). This method can provide comprehensive genetic information and help

deepen our understanding of TB transmission but there are some limitations mainly involved the high cost of consumable supplies, and apparatus including specialized personels and softwares for big data analysis.

Based on the significant public health problem of TB and DR-TB, the importance of TB epidemiology, the knowledge of gene mutations associated with the phenotypes of DR-TB, and the genotyping of TB as described above, genotyping and molecular characteristics of MTB and drug-resistant MTB isolated in Thailand was studied. The purposes of the research were to investigate molecular characteristics of mutations in INH-and RIF-resistant MTB strains and to determine genotypes of DR-MTB from an outbreak in Thailand. The finding can give us insight into the transmission of TB and DR-TB and help to develop new diagnostic and epidemiological tools or design the proper intervention to control TB in the study area and similar regions.

The content of this thesis is divided into two chapters. The scope and the objectives in each chapter were summarized as follows.

In Chapter I, I analyzed the mutations occurring in the *katG*, the *inhA* promoter, and the *rpoB* that are mainly associated with drug-resistant phenotypes in MTB isolated from Thai patients and compared mutations with those in the country and others.

In Chapter II, I conducted experiments to gain insight into the occurrence and transmission of drug-susceptible and -resistant TB in a TB outbreak hotspot, Thamaka district of Kanchanaburi province and neighbouring areas, in Thailand. Genotyping was conducted by spoligotyping, MIRU-VNTR typing, and SNP analysis.

Mutation characteristics in targeted genes associated with INH-, RIF-, FLQ-, and aminoglycoside resistance in MTB isolates were also analyzed.

The information of on the association between genetic mutations with drug-resistant MTB would be useful for developing or selecting the appropriate molecular diagnostic tools for the early detection of drug-resistant MTB and for designing proper management to control TB and drug-resistant TB in Thailand. The genotyping methods based on the combination of spoligotyping and MIRU-VNTR analysis would be applicable as a powerful tool for the identification of MTB strains and useful for TB monitoring and epidemiological surveillance.

CHAPTER I

Molecular characterization of mutations in isoniazid- and rifampicin-resistant *Mycobacterium tuberculosis* strains isolated in Thailand

Introduction

Tuberculosis (TB) remains one of the leading causes of deaths by infectious diseases. In 2020, the global TB burden was reported by an estimated 10.0 million cases with 1.5 million deaths. There were an estimated 0.5 million new cases of multidrug-resistant tuberculosis (MDR-TB) defined by resistance to the two most effective antituberculosis drugs, isoniazid (INH) and rifampin (RIF) which poses a major threat to TB control (82). Thailand is ranked in 30 high TB burden countries with a prevalence of 150 per 100,000 and an estimated MDR-TB of 1.7% and 10% in new TB cases and previously treated TB, respectively in 2019 (82).

The rise of MDR-TB constitutes a serious public health problem. *Mycobacterium tuberculosis* (MTB) strains can develop mono-resistance to INH- or RIF- including MDR-TB. As a further expansion of resistance in MDR strains, extensively drug-resistant (XDR) and totally drug-resistant (TDR) TB have been emerged with more threatening (29). The major mechanism of drug resistance in MTB is genetic mutations involved in the resistance to INH and RIF. The INH resistance is associated with mutations in several genes such as *katG*, *inhA* promoter, *kasA*, *ndh*, and the *oxyR-ahpC* (52). Approximately, 75-90% of INH resistance is attributed to mutations in the *katG* gene and the *inhA* promoter (7, 27, 40). INH-resistance is mainly caused by mutations in the *katG* gene which encodes a catalase-peroxidase, an enzyme whose activity is required for the INH activation (40). The *katG* gene mutations resulted in a high level of INH-resistance. Alternatively, low-level resistance to INH is caused by mutations in the regulatory region of the *inhA* gene or increasing the expression of *inhA* which encodes an enzyme involved in mycolic acid biosynthesis. About 6 to 20% of INH resistant isolates contain mutations in *inhA* promoter which locates at positions -15 and -8 (40, 58). Additional INH resistance around 10% is attributable to mutations occuring in other genes and accounts for low-level resistance to INH (7, 27). For RIF, the majority of RIF resistant strains carry mutations in a specific region of *rpoB* gene which encodes the β-subunit DNA-dependent RNA polymerase. The 81-bp of *rpoB* called the Rifampicin-Resistance Determining Region (RRDR) is considered as a

mutation hot spot region. About 90-96% of rifampicin-resistant MTB have mutations in the *rpoB* gene (7, 27, 40).

Although *in vito* drug-susceptibility testing (DST) is slow and time-consuming, its results are important and conclusive for appropriate TB treatment. Recently, rapid molecular tests for early diagnosis of drug-resistant TB have been used to provide rapid results by detecting common mutations (18, 38). Although several studies reported mutations detected, genetic determinants which are associated with drug resistance may vary base on MTB populations and distribute across the regions (1, 3, 6, 30, 31, 39, 49, 50, 70, 85, 86). DNA sequencing as the gold standard of molecular assays has been performed in many places to elucidate genetic mutations responsible for drug-resistant TB. The method is potential but less complicated than the whole genome sequencing. The frequency and mutation characteristics can be determined and may indicate the efficiency of molecular tests for the detection of drug-resistant TB. In addition, the data may be useful for developing specific mutation tests for each area. By DNA sequencing, this study investigated mutations occurring in the *katG* gene, the *inhA* promoter, and the *rpoB* gene of MTB isolated from Thai patients and compared with those in the country and others. The common and unique mutations were identified.

Materials and Methods

Bacterial strains and drug susceptibility testing (DST): DNA samples of 178 confirmed drug-resistant MTB isolated during 2013-2019 were purposefully selected and retrieved from archives of the regional TB referral laboratories; the office of Disease Prevention and Control Region 1 Chiang Mai, Disease Prevention and Control Region 5 Ratchaburi, and Disease Prevention and Control Region 12 Songkla, Department of Disease Control and the laboratory of National Institute of Health (NIH), Department of Medical Sciences. Mycobacteria were cultured previously in MGIT tubes (Becton Dickson, MD, USA) following the manufacturer's instructions. After growth, MTB isolates were identified and preserved in the bacterial collection according to routine procedures.DST was previously performed at each referral site as per routine practice with BD BACTEC™ MGIT™ automated mycobacterial detection system using the BACTEC™ MGIT™ 960 SIRE kit (Becton Dickson) for drug resistance assay following the manufacturer's recommendations. Final drug concentrations were 1.0 μg/mL for streptomycin (SEM), 0.1 μg/mL for INH, 1.0 μg/mL for RIF and 5 μg/mL for ethambutol (EMB).

Genomic DNA extraction: MTB DNA was extracted in each laboratory using lysis buffer-mixed-glass bead according to the manufacturer's protocol (Seegene, Inc., Korea). Briefly, 500 µl of bacterial suspension were centrifuged at 13,000 x rpm for 10 min, and cell pellets were re-suspended with 100 µl DNA extraction solution. The samples were mixed by vortexing for 30 s and heated at 95^{O}C for 20 min. The DNA supernatant was obtained for further analysis, and stored at -20^{O}C. The aliquots of stored DNA samples were shipped to the NIH laboratory for sequencing.

DNA sequencing of *rpoB*, *katG*, and *inhA* promoter region: Polymerase chain reactions (PCRs) were performed in a 20 µl mixture containing 0.25 mM (each) deoxynucleoside triphosphates, 0.5 M betaine, 1.5 mM magnesium chloride, 0.5 µM of each primer (48), 1 U of Go*Taq* DNA polymerase (Promega, WI, USA), 1 µl of DNA template. Each PCR primer pair was added to an individual reaction tube. The 392 bp DNA of the *katG* fragment (position 811 to 1170) was amplified with a primer pair; *katG*-F (5-ATGGCCATGAACGACGTCGAAAC-3) and *katG*-R (5-CGCAGCGAGAGGTCAG TGGCCAG-3). The 231 bp DNA of the *inhA* fragment (upstream -1 to -50) was amplified using primers; *inhA*-F (5-TCACACCGACAAACGTCACGAGC-3) and the *inhA*-R (5-AGCCAGCCGCTGTGCGATCGCCA-3). The DNA 278-bp of the *rpoB* fragment from 1261-1395 nucleotides was amplified using a primer pair; *rpoB*-F (5-CAGGACGTGGAGGCGATCAC-3) and *rpo*B-R (5-GAGCCGATCAGACCGATGTTGG-3). The reactions were carried out in a thermal cycler (Eppendorf AG, Germany) following conditions: denaturation at 96^{O}C for 60 s, followed by 35 cycles of amplification at 96^{O}C for 10 s, annealing at 55^{O}C for 10 s and elongation at 72^{O}C for 30 s, with a final elongation 1 cycle at 72^{O}C for 5 min. The PCR products were analyzed by 2% agarose gel electrophoresis. Thereafter, DNA fragments of interest were recovered from the agarose gel and used for sequencing using primers *katG*-F, *inhA*-F, and *rpoB*-F for *katG*, *inhA*, and *rpoB,* respectively with BigDye™ Terminator V3.1 Cycle Sequencing kit (Life Technologies Corp. CA, USA), and an ABI Prism 3130xl Genetic Analyzer (Life Technologies). The resulting sequences were compared to wild-type sequences of *M. tuberculosis* H37Rv (APPENDIX B) using Bio-Edit software version 7.0.9 (20). The identified mutations were then compared with those reported previously. The PCR amplification was carried out at NIH, Thailand, and the sequencing analysis was performed both at NIH and Hokkaido University, Japan.

Results

Phenotypic INH and RIF resistance: Phenotypic DST identified INH-resistance in 53/178 (29.8%), RIF-resistance in 10/178 (5.6%) and MDR-TB in 115/178 (64.6%). Overall, INH and RIF resistance was identified in 168 and 125 isolates, respectively. Drug-resistant profiles of strains are showed in Table 5.

Genotypic INH resistance: For *katG*, eight different types of mutations were detected in five codons (300, 315, 334+345, and 337), and gene deletion 134/168 (79.8%) INH–resistant isolates had a single mutation in the *katG* gene and the most prevalent alteration was a substitution Ser315→Thr found in 125 of 168 (74.4%) (95%CI, 67.1-80.8) isolates (Table 6). Two of the three isolates had an amino acid Ser315→Asn substitution and the other one had an amino acid Ser315→Ile substitution. Other amino acid substitutions included Trp300→Gly found in 2 isolates and a mutation at codon 337 leading to Tyr→Cys was found in one isolate. In addition, one isolate had nucleotide insertion G (InsG) and nucleotide 1003 or codon 334. A silent mutation (Lys345→Lys) was found in one isolate. Two isolates demonstrated no amplification of the *katG* fragment. No mutations in both the *katG* gene and *inhA* promoter were identified among 10 phenotypic mono-RIF resistant strains (Table 6).

For *inhA* promoter region, 27/168 (16.1%) were identified to have a single mutation (Table 6). Out of these, 25 (15%) isolates contained a C >T substitution at the position 15 upstream (-15 C >T) of the starting site of the gene. Two isolates had a conversion from T > C at the position 8 upstream (-8 T >C) of the *inhA* promoter initiation. A combination mutation in the *katG* gene and the *inhA* promoter was found in two isolates (1.2%). One of these had a C→A substitution at 31 upstream (-31 C >A) the *inhA* promoter plus the deletion in the *katG* gene. In total, the other isolate had a C > T substitution at 15 upstream *inhA* plus a Ser315→Thr substitution in the *katG* gene. In total, 5 (3.0%) of 168 INH-resistant isolates had no mutations within the target fragments.

RIF resistance: Twelve different mutations involving codons 432, 433, 435, 445, 450, and 452 were identified (Table 7). Altogether, 118/125 isolates (94.4%) harbored at least one mutation within *rpoB*. Of these, 115 (92.0%) isolates had a single mutation. Double mutations were found in 3 (2.4%) isolates affecting codons 429, 445, 450, and 460. In total, 15 resistant genotype patterns were identified in the *rpoB* gene. The most frequently mutated codons were 450, 445, and 435, with frequencies of 58.4% (73/125), 18.4% (23/125), and 9.6% (12/125), respectively. The most common

mutation was Ser450→Leu (56.8%) substitution, followed by His435→Val (8.8%), His445→Try (6.4%), and His445→Asp (5.6%) substitutions. One isolate had a universal TTC (Phe) insertion between 432 and 433 codons. Rare mutations were also detected in substitution of Ser450→Gly, Gln429→His, Glu460→Gly in each isolate. Seven isolates (5.6%) showed no mutations in this sequenced region. Among RIF susceptible isolates, 50/53 contained no mutation in the *rpo*B gene while the substitutions; Leu452→Pro and Asp435→Tyr were identified in two and one isolate (s), respectively.

Comparison of mutations in the *katG* gene, *inhA* promoter and *rpoB* gene

Among three studies in Thailand, all reported the high frequency of the *katG* gene mutation at codon 315, and Nepal reported the highest frequency rate (82.3%). The frequency range varied from 23.1-76.5% among seven Asian countries (Table 8). For the *inhA* promoter, a point mutation at -15 was observed at the highest frequency. In this region, the maximum frequency was reported by a study from Korea at 38.5% (Table 8). Among RIF-resistant strains, all countries reported codon 450 of the *rpoB* gene to be the most highly mutated. The mutation rates at this codon varied in the range of 39.2-59.2%, followed by the mutation at codon 445 except its low rate reported by the study from Korea (Table 9). The comparative results of the mutation analyses were summarized in Table 8 and 9.

Discussion

The high level of the association between the mutations of the *katG* gene, *inhA* promoter, and the *rpoB* gene and phenotypic resistance to INH and RIF were demonstrated in this population. Comparative analysis showed that the mutation rates in *katG* gene and *inhA* promoter by 81.0% (95% CI, 74.2-86.6) and 17.3% (95% CI, 11.9-23.8), respectively, were similar to those reported previously in Thailand by 80% (3) and 84.0% (6) for the *katG* gene, and by 11.2% (3) and 14.0% (6) for the *inhA* promoter. The frequency of the *katG* gene mutation in Thailand was higher than those reported in Singapore (36.2%), Myanmar (61.2%), Vietnam (76.8%), and China (70.3%) (1, 39, 86), respectively (Table 8). Overall, the frequency of the most common Ser315Thr substitution in the *katG* gene was highly identified at 76.2% compared with the high rate of those in Nepal (82.3%), south India (76.5%), Myanmar (61.2%), Vietnam (76.8%), and China (70.3%) (1, 39, 41, 49, 86) while the low frequencies of the *katG* gene mutation were documented in Singapore (36.2%) and Korea (23.1%) (30, 85).

Some INH-resistant strains contained mutations in the *inhA* promoter (1, 3, 6, 30, 39, 41, 49, 85, 86). Table 8 showed the mutation rate ranging from 10.6-38.5% (Table 8). Similarly, we found that the mutation in the *inhA* promoter at position -15 was the second most common among INH-resistant isolates. About 15.5% of isolates in this study carried *inhA* -15 mutation, which was at a higher rate compared to those reported in Nepal 10.6%, Myanmar (10.7%), Nepal (10.6%), China (8.6%) (1, 48, 86), but lower than those in Singapore (26.9%) and Korea (38.5%) (30, 85) (Table 8). Comparing the distribution between the high and the intermediate or low TB. TB burden countries showed a difference in the prevalence of mutations in the *katG* gene and *inhA* promoter. Similarly, mutations at positions -8 and -31 were rare.

Other mutations in the *katG* gene were identified, such as Tyr337 →Cys, and 1003InsG. These have not previously been reported in surrounding countries, according to the literature review (1, 3, 6, 5, 22, 30, 39, 41, 59, 85, 86). In addition, a rare mutation Trp300→Gly which was documented in Thailand, India (South), and Singapore (3, 6, 30, 41) was observed. Furthermore, no amplification of the *katG* gene was observed in the two isolates. These isolates possibly lack whole/partial *katG* gene or carry a mutation at primer binding sites. This should be confirmed by the PCR amplification of a larger region covering the current PCR amplified region in our future study. The effect of whole/partial *katG* gene deletion on a reduction of catalase activity has been reported (22, 67). As a result, INH cannot be converted into an active form against MTB (2, 67). These unique alterations have not been reported in the TB database (59), suggesting the spontaneous arise in the development of INH resistance. These mutations cannot be detected by line probe assay (LPA) using GenoType MTB /MDR*plus* according to the information for users. It might be possibly missed in the detection for INH-resistant TB by the widespread use of rapid molecular diagnostics. Therefore, the identification of a distinct mutation in the *katG* gene could improve the sensitivity of the next generation of diagnostics for TB drug resistance. Otherwise, no mutations in both the *katG* gene and *inhA* promoter were identified in low prevalence (3.0%), which were consistent to the finding from Nepal, Myanmar, Singapore, Vietnam, China, and others (Table 8).

For RIF, the *rpoB* mutations were highly associated with RIF resistance, and the mutations in the 81-bp RRDR region accounted for 71.7-94.4% (1, 31, 39, 41, 70, 85, 86). The mutations at positions 450, 445, and 435 were the most frequent in RIF-resistant strains, which is in agreement with previous studies (6, 50). The highest frequency mutation at codon 450 was 59.2% reported by this study. This result agreed

with previous reports from Nepal (58.7%) (49), Myanmar (48.3%) (1), China (54.7%) (86), Cambodia (46.5%) (70), and Singapore (54.9%) (31) (Table 9). The finding of one TTC insertion was similar to those isolated in Nepal (49). The Gln432Lys substitution is another rare mutation similar to one isolated in Nepal, but different from those in Singapore (Gln432Glu) (31). Double mutation type was rare and observed differently in each country. One strain with double mutations at codon 445/460 showed similarity to those earlier reported in Thailand (6). The rare mutations at either codon 429 or 460 reported here have been identified in Myanmar (1). Meanwhile, the presence of no mutation in the RRDR (5.6%) was in a similar range to other reports except the two studies from Myanmar (28.7%) and Korea (22%) (1, 85). The discordance in genotypic and phenotypic was presented in a small proportion and could be explained in that the RIF resistance was attributed to the occurrence of mutations outside the RRDR or in other loci in the genome. Otherwise, the resistance might be mediated by other mechanisms. Nonetheless, this study detected substitution mutations in 3 RIF-susceptible isolates which are not associated with RIF resistance. Specific mutations at codons 430, 432, 437 and 441 conferring lower-level resistance to RIF were reported (50). It might turn into RIF susceptibility in drug-resistant phenotypic analysis. These mutants were not detected in this study. In addition, it was unclear whether a mutation at codon 452 was related to RIF resistance (6). Our finding of the mutation at codon 452 in 5 RIF-resistant isolates indicated the association with RIF resistance. Since RIF resistance is often considered as a surrogate marker for MDR-TB, the detection of *rpoB* mutation is important. Based on WHO report, the proportion of MDR/RR slowly declined in Thailand. The proportion of MDR among RR-TB has been declined from 86% in 2016 to 74% in 2019 (83). In this respect, the detection of INH resistance is necessary.

Several studies reported mutations in *katG*, *inh*A promoter and *rpoB*. The mutation patterns and frequencies in specific country may vary by MTB populations. In Thailand, mutation characteristics seem to be in consistent patterns but the frequency varied according to recent studies (2, 44, 56, 57). Some advantages and update data from those studies involved the use of whole genome sequencing and nationwide sampling (44), the correlation of mutations conferring drug-resistance and MTB genotypes (56, 57), the analyses of mutations in other genes involved in drug-resistant TB (71), the distribution of *katG* and *InhA* mutations with the potential of the diagnostic assay for MDR-TB (2). This study was performed in accordance to the purpose on characterization of resistance-associated mutations in major genes. Usual

and distinct mutations were highlighted. The limitations of this study were no clinical data, no data of MTB genotypes, and the limitation of DNA sequencing targeting specific sequences. However, DNA sequencing exhibited applicability for the detection of INH and RIF resistance. In this study, the sensitivity of the detection of the *katG* gene and *inhA* promoter mutations for detecting INH resistance 97% [95% Confidential interval (CI); 88.4-97.5], and 100%, repectively. For RIF resistance, the *rpoB* sequencing showed 94.4 % [95% CI 88.8-97.7] sensitivity.

Overall, we showed a high rate mutation in the *katG* gene, the *inhA* promoter, and the *rpoB* gene with the most common mutation of *katG* Ser315Thr, *inhA* promoter 15C>T, and *rpoB* Ser450Leu. Therefore, most INH- and RIF- resistance in this setting could be detected by mutation analysis. Rare mutations and distinct mutations raised a concern for possible misdetection by molecular tests. The data generated would be useful for developing new molecular tests and designing proper management for effective control of TB drug resistance.

Summary

Drug-resistant tuberculosis (TB) is a great challenge in TB control. The frequency and mutation characteristics can imply the efficiency of molecular tests for the rapid detection of TB drug resistance. This study examined the existence of mutations in *katG* and *inhA* promoter for isoniazid (INH) resistance, and *rpoB* for rifampicin (RIF) resistance. A total of 178 drug-resistant *Mycobacterium tuberculosis* (MTB) isolates were analyzed. Mutations in *katG* encoding and *inhA* regulatory regions were detected in 136/168 (81.0%) and 29/168 (17.3%), respectively, with the most prominent mutation of Ser315Thr substitution in *katG* in 126/168 (75.0%), and -15 C > T substitution in the regulatory region of the *inhA* (26/168; 15.5%). Two distinct *katG* mutations (Tyr337Cys, 1003InsG) were identified. Of 125 RIF-resistant isolates, 118 (94.4%) carried mutations affecting the 81-bp RIF resistance-determining region (RRDR) with the most commonly affected codons 450, 445, and 435 identified in 74 (59.2%), 26 (20.8%) and 12 (9.6%) isolates, respectively. The genetic mutations were highly associated with phenotypic INH and RIF resistance, and the majority shared similarities with those in previous studies in Thailand and other Asian countries. The data is useful for guiding the use and improvement of molecular tests for TB-drug resistance.

Table 5. Frequency of INH- and RIF-resistant MTB strains isolated in Thailand

TB drug-resistance types	Drug resistant profile				No. of isolates (%)
	INH	RIF	EMB	SEM	
- INH-resistance					**53 (29.8)**
	R	S	S	S	38
	R	S	S	R	9
	R	S	ND	ND	6
- RIF-resistance					**10 (5.6)**
	S	R	S	S	7
	S	R	S	R	1
	S	R	ND	ND	2
- Multidrug resistance					**115 (64.6)**
	R	R	S	S	20
	R	R	S	R	27
	R	R	R	S	7
	R	R	R	R	50
	R	R	ND	ND	11
Total					**178**

Abbreviation: INH; isoniazid, RIF; rifampicin, EMB; ethambutol, SEM; streptomycin, S; susceptible, R; resistance, ND; no data.

Table 6. Distribution of mutations in *katG* encoding region and *inhA* regulatory region of 168 INH-resistant isolates (115 of MDR and 53 of INH-resistance) in Thailand.

Mutation site	***katG* encoding region**		***inhA* regulatory region**		**No. of isolates**	**%**
	Nucleotide change (s)	**Amino acid change (s)**	**Mutation site**			
300	TGG to GGG	Trp to Gly			2	1.2
315	AGC to ACC	Ser to Thr			125	74.4
	AGC to AAC	Asn			2	1.2
	AGC to ATC	Ile			1	0.6
334+345*	1003 InsG+AAG to AAA	frameshift /synonymous			1	0.6
337*	TAC to TGC	Tyr to Cys			1	0.6
gene deletion	Δ*katG* gene				2	1.2
	Subtotal (*katG* encoding region)				**136****	**81.0**
			-8	T to C	2	1.2
			-15	C to T	25	14.9
	Subtotal (*inhA* regulatory region)				**29****	**17.3**
315	AGC to ACC	Ser to Thr	-15	C to T	1	0.6
gene deletion	Δ*katG* gene		-31	C to A	1	0.6
Total					**163**	**97.0**
	No mutation			No mutation	5	3.0

Note: * Distinct mutation, ** The values include isolates with mutations at combination mutation (*katG*+*Inh*A)

Abbreviation: INH, isoniazid; Trp, Tryptophan; Ser, Serine; Asn, Asparagine; Ile, Isoleucine; Tyr, Tyrosine; Cys, Cysteine**;** Thr,Threonin

Table 7. Distribution of mutations in *rpoB* of 125 RIF-resistant isolates (115 MDR and 10 RIF-resistant MTB) in Thailand

Codon (s)	Mutation		Amino acid change (s)	No. of isolates	%
Single mutation					
432-433	TTC Insertion		Phe insertion	1	0.8
432	CAA to AAA		Gln to Lys	1	0.8
435	GAC to TAC		Asp to Tyr	1	0.8
	GTC		Val	11	8.8
445	CAC to CGC		His to Arg	4	3.2
	GAC		Asp	7	5.6
	CCC		Pro	4	3.2
	TAC		Tyr	8	6.4
450	TCG to GGG		Ser to Gly	1	0.8
	TTG		Leu	71	56.8
	TGG		Trp	1	0.8
452	CTG to CCG		Leu to Pro	5	4.0
Subtotal				**115**	**92.0**
Double mutation					
429 + 445	CAG + CAC	to CAC + CGC	Gln + His to His + Arg	1	0.8
445 + 450	CAC + TCG	to TAC + TTG	His + Ser to Tyr + Leu	1	0.8
445 + 460	CAC + GAG	to TAC + GGG	His + Glu to Tyr + Gly	1	0.8
Total				**118**	**94.4**
No mutation				7	5.6

Abbreviation: RIF, Rifampicin; MDR, multidrug resistance; Ins, Insertion; RRDR; RIF, resistance-determining region (RRDR); Phe, Phenylalanine; Gln, Glutamine; Lys, Lysine; Ser, Serine; Asn, Asparagine; Asp, Aspartate; Trp, Tryptophan; Tyr, Tyrosine; Val, Valine**;** His, Histidine; Arg,Arginine; Asp, Aspartic acid; Pro, Proline; Gly, Glycine; Leu, Leucine; Trp, Tryptophan; Glu, Glutamic acid

Table 8. Comparison of mutation rates in *katG* gene and *inhA* regulator in INHresistant MTB isolates in Thailand and other Asian Countries

Genes	Mutated Codon/Locus	% mutations[a] in different geographic regions										
		Thailand			Nepal	India (South)	Myanmar	Singapore	Vietnam	China	Korea	Range
		n=168 This study	n=160 ref 6	n=50 ref 3	n=113 ref 49	N=34 ref 41	N=178 ref 1	N=160 ref 30	N=82 ref 39	N=128 ref 86	N=52 ref 85	
katG	315	76.8	80.0	84	82.3	76.5	61.2	25.6	76.8	70.3	23.1	23.1 - 76.5
	Deletion	1.2	0.0	0.0	0.0	2.9	0.0	10.6	0.0	0.0	0.0	1.2– 10.6
inhA	-8	1.2	0.6	0.0	0.9	5.9	0.1	0.6	NA	4.7	1.9	0.1 – 5.9
	-15	15.5	10.6	14	10.6	14.7	10.7	26.9	NA	8.6	38.5	10.6 –38.5
Others[b]		3.0	2.5	0.0	7.1	8.8	1.1	2.5	9.8	4.6	28.8	1.1 - 28.8
None[c]		3.0	0.0	4	6.2	0.0	27.5	35.6	13.4	10.9	23.1	3.0 - 35.6

[a]The values include isolates with mutations at multiple codons

[b]Others: mutations in other gene

[c]None: no mutation in sequence region

NA: Not application

Ref: reference

Table 9. Comparison of mutation rates in RRDR of *rpoB* in RIF-resistant *M. tuberculosis* isolates in Thailand and other Asian countries

Mutated codon	% mutations[a] in different geographic regions											
	Thailand			Nepal	India (South)	Myanmar	Vietnam	China	Korea	Cambodia	Singapore	Rang
	n = 125	n = 33 ref 6	n = 153 ref 50	n = 109 ref 49	n = 74 ref 41	n = 128 ref 1	n = 74 ref 39	n = 128 ref 86	n=41 Ref 85	n = 101 ref 70	n = 51 ref 31	
430	0.0	0.0	0.7	0.9	0.0	0.0	0.0	3.1	2.4	0.0	0.0	0.7 - 3.1
432	0.8	0.0	2.6	4.6	0.0	1.1	0.0	0.0	2.4	0.0	2.0	0.8 - 4.6
435	9.6	0.0	2.6	15.6	12.5	2.8	9.5	10.2	9.8	17.8	11.8	2.6 - 17.8
437	0.0	0.0	0.7	0.0	0.0	1.1	0.0	0.0	2.4	2.0	2.0	0.7 - 2.4
441	0.0	3.0	0.7	0.0	4.2	0.0	1.4	0.8	2.4	0.0	2.0	0.7- 4.2
445	20.8	39.4	26.8	15.6	25.0	14.04	23.0	22.7	7.3	21.8	23.5	7.3 - 39.4
450	59.2	39.4	56.9	58.7	50.0	48.3	39.2	54.7	46.3	46.5	54.9	39.2 - 59.2
452	4.0	9.1	2.0	2.8	8.3	4.4	1.4	4.7	0.0	1.0	0.0	1.0 - 9.1
Others[b]	1.6	9.0	4.6	1.8	12.5	1.7	21.6	6.3	4.9	9.9	2.0	1.6 - 21.6
None[c]	4.0	3.0	0.0	2.8	0.0	28.7	4.1	6.25	22.0	5.9	2.0	2.0 – 28.7

[a]The values include isolates with mutations at multiple codons.

[b]Other point in RRDR or outside

[c] No mutation in the region

Ref: reference

CHAPTER II

Drug-resistant *Mycobacterium tuberculosis* and its genotypes isolated from an outbreak in western Thailand

Introduction

Tuberculosis (TB), which is caused by the *Mycobacterium tuberculosis* complex is a serious health-deteriorating disease and one of the leading causes of deaths worldwide. Thailand has a high TB burden, with approximately 120,000 new cases in 2018, and it is ranked among the top 30 TB-burdened countries. TB infection in Thailand, which has a population of about 69 million people, was reported to have an incidence rate of 153 new cases per 100,000 inhabitants. Multidrug-resistant TB (MDR-TB) caused by *Mycobacterium tuberculosis* (MTB), and resistance to at least isoniazid (INH) and rifampicin (RIF), was detected at about 2.3% and 24% of both new cases and retreated patients, respectively (80). The Kanchanaburi province, located in western Thailand, is one of the areas that has a high incidence of MDR-TB. In particular, the Thamaka district in Kanchanaburi had been documented as a hotspot, with multiple community outbreaks of MDR-TB from 2002 to 2010 (25). An epidemiological investigation, proper strategies, and rapid responses, including effective treatments, are deemed necessary to control TB transmission in the epidemic area.

Occurrence of drug-resistant TB can be categorised into primary and secondary drug resistance. The primary drug resistance is caused by an infection with drug-resistant MTB, while secondary drug resistance is an acquired resistance developed at a later stage by drug-susceptible strains (53). Major mechanisms of drug resistance that primarily involve genetic mutations have been reported in a number of studies (42, 46). For example, mutations in the *katG* gene account for 30-75% of INH resistance and are associated with a high-level resistance to INH. The most frequent mutation in the *katG* gene is Ser315Thr. In addition, a point mutation in the regulatory region of *inhA* promoter, mainly at the positions -15 and -8, accounts for 6-30% of INH resistance, resulting in a low-level resistance to INH (3, 27). Meanwhile, 90-96% of RIF resistance is caused by mutations in the *rpoB* gene in which the 81 bp of the rifampicin resistance determining region (RRDR) between codons 507–533 are considered a mutation hotspot region (27). The increasing incidence of MDR and extensive drug-resistant tuberculosis (XDR-TB; MDR plus resistance to fluoroquinolone and aminoglycoside) have become major hurdles to treating TB effectively (48). To date,

molecular epidemiological data from many countries have demonstrated that MDR and XDR-TB have been detected around the world and hence area serious threat to proper TB control. To overcome this epidemiological problem, efficient epidemiological strategies are needed (10, 29, 34).

For the purpose of molecular epidemiology, a number of genotyping methods have been reported and used to differentiate MTB strains. For example, spoligotyping (28), analysis of MIRU-VNTR (69), single nucleotide polymorphisms (SNP) (14, 37, 68), region of differences (RDs), large sequence polymorphisms (LSP) (14) and restriction fragment length polymorphism (RFLP) (14) are among the methods extensively used for typing MTB strains. Recently, whole genome sequencing (WGS) has been used to genetically characterise MTB (19). This method can provide comprehensive genetic information and help deepen our understanding of the factors affecting TB transmission. The genotyping of MTB is a molecular tool that can be used to investigate possible epidemiological links between TB patients, detect suspected outbreaks and distinguish exogenous reinfection from endogenous reactivation in relapse cases (69). For example, spoligotyping is a simple and rapid genotyping method that focuses on the detection of polymorphisms based on the presence or absence of 43 unique spacer sequences in the direct repeat (DR) regions (28). Spoligotyping has been used to categorise MTB strains into several families, such as Beijing, East African Indian (EAI), Central Asian (CAS), Latin American Mediterranean (LAM), unclassified (U) and T families (4). Analysis of MIRU-VNTR is another genotyping method that has been widely used to monitor the global transmission of TB (69). The combination of spoligotyping and MIRU-VNTR typing methods has become a traditional and practical tool providing sufficiently high resolution for the clear identification of MTB strains (35). Generally, MIRU-VNTR analysis has been used after spoligotyping to further subtype major MTB groups including the Beijing family, which commonly causes several major outbreaks worldwide, is over-represented in drug-resistant isolates and seems to have unique virulence properties (14).

Outbreaks of TB along the western border of Thailand with Myanmar are still a problem as genotyping is not routinely conducted during these events. The current study was conducted to gain an insight into the spread of TB in the Thamaka district of Kanchanaburi province, Thailand and neighbouring regions. Genotyping was conducted by spoliogotyping, MIRU-VNTR, and SNP analysis. Mutation characteristics in targeted genes associated with INH-, RIF-, FLQ- and aminoglycoside resistance in MTB isolates were also analysed.

Materials and Methods

Processing of MTB strains and samples

A total of 72 MTB isolates from Thai pulmonary-TB patients (registered for TB treatment during 2013-2014 MDR-TB outbreak in Thamaka District, Kanchanaburi Province) and strains isolated from Prachuap Khiri Khan Province in western Thailand and the surrounding area, Samut Sakhon Province, were collected and analysed. The locations of the sample collection sites are showed in Figure 9. These isolates were obtained from decontaminated sputum samples that were inoculated into BACTEC MGIT liquid media (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). Drug susceptibility testing (DST) for first-line drugs including INH, RIF, EMB, and STM was conducted using BACTEC MGIT 960 SIRE Kit, as per the manufacturer's instructions. A second-line drug susceptibility test was only carried out on some isolates (data not shown).

DNA was extracted from 500 μl of MTB cell culture suspension recovered from the BACTEC 960 MGIT media using magnetic bead-based nucleic acid extraction (SEEPREP 12, Seegene, Inc., Korea) as per the manufacturer's protocol. The extracted DNA with a final volume of 50 μl was kept at -20^{O}C until further molecular analyses.

Genotyping

MTB isolates were genotyped by spoligotyping as previously described. Briefly, DR region in chromosomal DNA of MTB was amplified by PCR (primer sequence in APPENDIX A). The hybridization of PCR products to specific DNA probes was then screened for spoligo-patterns. The obtained spoligo-patterns were converted to an octal code and compared with those in the international spoligotyping database. SpolDB4 database was used to identify spoligo-international types (SITs) (4). To subtype isolates, the 15 selected MIRU-VNTR (424, 580, 802, 960, 1644, 1955, 2163b, 2165, 2401, 2996, 3192, 3690, 4052, 4156, and 4348) were analyzed according to a method previously described (34). After amplification, the copy numbers of tandem repeats were determined from the sizes of amplicons (APPENDIX C) by agarose gel electrophoresis using 50 bp DNA ladders (New England, BiolabsInc., USA) as markers (69). A numerical value profile was assigned to each strain according to the number of variable repeats in each VNTR allele.

The allelic diversity among the MIRU-VNTR isolates was determined. The Hunter-Gaston discriminatory Indexes (HGDIs) (21) were calculated by the following

equation: HGDI = 1- $(1/N(N-1))\sum_j^S nj(nj-1)$), where N is the total number of isolates in the sample population, S is the total number of allelic, and n_j is the frequency of the allele in the locus. The discriminatory power was considered high if the HGDI value was >0.6, moderate if HGDI was between 0.3 to 0.6 and low if HGDI was <0.3, according to the definition by Sola *et al*., 2003 (66).

Analysis of clustering

Spoligotyping patterns and 15-locus MIRU-VNTR typing data were analyzed by MIRU-VNTR*plus* (http://www.miru-vntrplus.org/MIRU/index.faces) to establish the lineages and sub-lineages (76). Dendrograms were generated by the unweighted pair group method with arithmetic means (UPGMA). In the current study, the genotypes of Beijing isolates were compared with those of a previously published study (10). A minimum spanning tree (MST) was constructed based on the data of 15-locus MIRU-VNTR analysis using BioNumerics software version 6.6 (Applied Maths, Kortrijk, Belgium). Clusters were defined as two or more isolates sharing an identical 15-locus MIRU-VNTR pattern. The clustering rate was calculated using the following formula: the number of clustered isolates/ total number of isolates.

SNP analysis and gene mutation detectionby sequencing

SNP at the 1477596 locus was examined to subdivide modern and ancestral Beijing genotypes (14). DNA samples of Beijing isolates were amplified by PCR. The PCR mixture contained 1x Go*Taq* buffer (Promega Co., Madison, WI, USA), 0.26 mM of each dNTP, 0.3 μM of each primer (data sequence in APPENDIX A), 0.5 M Betaine, 0.5 units of Go*Taq* DNA polymerase and DNA of MTB. DNA sequences were aligned through pairwise alignment of the *M.tuberculosis* H37Rv using Bio-Edit software version 7.09 (20).

DNA samples from drug-resistant strains were subjected to an analysis of mutations in genes associated with drug resistance, targeting the following: the *katG* coding sequence and *inhA* promoter region for INH; RRDR in *rpoB* for RIF; QRDR in *gyrA* for fluoroquinolone; and 16S ribosomal RNA gene in *rrs* for kanamycin or aminoglycosides (48, 49). The PCR components used were the same as those described above. The amplified DNA fragments were subjected to sequence analysis using BigDye terminator V3.1 (Life Technologies Co., CA) reagents and a 3130 genetic analyser (Life Technologies Co., CA), as per the manufacturer's protocol. Sequences were compared with those of wild-type reference H37Rv and analyzed using Bio-Edit software.

Results

MTB phenotyping by DST

Based on the DST, phenotypes resistant to four first-line anti-TB drugs were identified in all 72 MTB isolates. Of them, 62.5% (n=45) were susceptible to INH and RIF, while the rest (n=27) were drug-resistant MTB, including RIF mono-, INH mono-resistant, and MDR at 5.6% (n=4), 1.4 % (n=1) and 30.6% (n=22), respectively (Table 10).

Spoligotyping

The spoligotyping results showed clustering among 59 isolates and 13 different non-clustered spoligotypes (Table 11). Overall, there were 20 spoligotype patterns and the clustering rate was 82%. The largest cluster belonged to the Beijing SIT1 family (n=38; 53%). Non-Beijing (n=34; 47.2%) included EAI2_NTB (n=6; 8.3%), EAI5 (n=8; 11.2%), EAI6_BGD1 (n=4; 5.6%), H3 (n=1; 1.4%), T1 (n=2; 2.8%) U (n=6; 8.4%) and new spoligotypes (n=7; 9.8%) (Table 11). There were 6 small clusters of T1 SIT53 (n=2), EAI2_NTB SIT89 (n=6), EAI5 SIT236 (n=3), EAI5 SIT256 (n=4), U SIT523 (n=4), and EAI6_BGD1 SIT591 (n=2) among 21 non-Beijing, as shown in Table 11. The rest of the non-Beijing group (n=13, 18%) was non-clustered (Table 11). Of 38 Beijing (SIT1) genotypes, 19 (50%) were drug-resistant and 19 (50%) susceptible to both INH and RIF**.**

MIRU-VNTR analysis

According to the 15-locus MIRU-VNTR analysis, 5 different subtypes were generated among 25 Beijing isolates consisting of Bj-I (n=16), Bj-II (n=2), Bj-III (n=2), Bj-IV (n=2) and Bj-V (n=3). The remaining of the 13 Beijing isolates were single or non-clustered Beijing subtypes (Figure 10A). The largest cluster was a group of Bj-I subtypes that contained 16 isolates (42%, 16 of 38 Beijing), including 14 MDR-TB isolates, 1 mono RIF- resistant and 1 susceptible isolate. All members of Bj-I cluster were from the Kanchanaburi Province. In addition, Beijing drug-susceptible strains were found in clusters of Bj-II, Bj-III, and Bj-IV. The other three Beijing isolates were categorised into the Bj-V cluster. These three isolated were from the Samut Sakorn Province. Two of these isolates were MDR and one was INH mono-resistant. The remaining Beijing isolate was from the Prachuap Khiri Khan Province and non-clustered (Figure 10A). In non-Beijing groups, the U SIT523 was found to have an identical MIRU-VNTR pattern in two of four isolates (Figure 10B). The MIRU-VNTR

pattern in the matching U strains was 443 824 652 641 113. Overall, 59 clustered isolates that shared spoligotypes were subtyped by the MIRU-VNTR analysis, generating 38 different MIRU-VNTR patterns (Figure 10A). Of these, 6 patterns formed major clusters containing 27 isolates, while 32 unique individual MIRU-VNTR patterns were identified. The dendrogram resulting from clustering analysis is shown in Figure 10.

Analysis of discriminatory power and diversity

A discriminatory power of the selected 15-locus MIRU-VNTR for differentiating the Beijing genotype was observed (Table 12). The comparison results showed that the HGDI value of 19 Beijing drug-resistant isolates (MDR, INH mono-resistance and RIF mono-resistance) was 0.38 lower than that of drug-susceptible Beijing isolates, with a value of 0.98. The difference in discriminatory power between the two groups resulted in the difference in generating the cluster of Beijing isolates, as shown in Figure 10C, 10D. In the drug-susceptible Beijing group, allelic diversity was highly discriminated in locus 2163b and moderate in loci 0424, 0802, 1955, and 2996. Thus, 3 clusters consisting of 6 isolates and 13 single isolates were differtiated among drug-susceptible Beijing strains. Nonetheless, although this selected 15-locus MIRU-VNTR had a high power to distinguish Beijing isolates susceptible to INH and RIF, it could not differentiate the MDR Beijing family that included 14 of 16 isolates in the Bj I cluster and 2 of 3 isolates in the Bj-V cluster.

SNP for identification of modern and ancestral Beijing

SNP data from locus 1477596 indicated that 37 of 38 Beijing isolates were either a SNP-T type or a modern Beijing sub-lineage (97%). Only one isolate was observed to have a SNP-C type (ancestral Beijing subtype) (Figure 10A).

Identical MTB genotypes

MIRU-VNTR data from Beijing isolates that were collected from past TB outbreaks occurring during the periods 2006-2012 and 2013-2014 were retrieved from a previous report (10) and compared with those from the current study (Figure 11). The results showed a major clonal expansion of 62 MDR-MTB isolates originated from previous outbreaks in 2006 (n=3), 2007 (n=1), 2008 (n=5, including 1 XDR-TB isolate), 2009 (n=8), 2010 (n=5), 2011 (n=1), and 2012 (n=1), and isolates in the Bj- I cluster (Figure 11) from the current study (n=38).

Mutations associated with drug resistance

The sequencing results showed a single type of mutation in the *katG* gene at a codon Ser315Thr in all INH-resistant isolates. Meanwhile, a substitution of Ser531Leu

in *rpoB* was found in almost all RIF-resistant strains except for two isolates carrying His526Arg and Ser533Pro. Of those 22 MDR, 20 isolates, which consisted of 16 Beijing, 1 EAI6_BGD1 (SIT591) and 3 U (SIT523), had a common mutation of Ser315Thr in the *katG* gene and Ser531Leu in the *rpoB* gene for typical INH and RIF resistances, respectively (Figure 10). Only one Beijing MDR strain was found to carry Ser315Thr in the *katG* gene but had a different mutation of His526Arg in the *rpoB* gene. This isolate was from the Samut Sakorn Province. Another one had no data of DNA sequences in any of the targeted genes. As for MDR, a common mutation of Ser531Leu in the *rpoB* gene was found in three of four mono RIF-resistant isolates that were Beijing, EAI5 SIT526, and new spoligotypes, while Leu533Pro was carried by EAI6_EGD1 SIT591. Most of the MDR isolates had no mutations in the *rrs* and the *gyrA* genes, which conferred resistance to KAN and FLQ, respectively. The exceptions were two isolates of the U SIT523 strain, which were observed to have the additional mutations A1400G and Asp94Gly in the *rrs* and the *gryA* genes, respectively (Figure 10B). Therefore, these two identical U SIT523 isolates were considered to be strains that cause a transmission of XDR-TB.

Discussion

A number of MDR cases has been detected in the Thamaka district, in the Kanchanaburi province, Thailand, and epidemic outbreaks of MDR-TB have continuously been reported in this areas since then. Previously, the potential of a specific MDR-TB strain to cause a continuous outbreak of drug-resistant TB was reported, showing the genetic diversity of MTB strains in these region (25, 53). Currently, molecular detection is being conducted for rapid detection of drug-resistant cases, but no genotyping is routinely carried out. Therefore, the genetic information of clinical isolates currently available is still limited. A previous study using retrospective isolates from 2006 showed that most of the MDR strains belonged to the Beijing genotype, and that U SIT523 was the XDR-TB strain (10). The findings of the current work confirmed that primary transmission of drug-resistant TB was the major factor driving the increase of TB drug resistance in local outbreaks in Thailand and that a drug resistance mechanism related to INH and RIF was caused by common mutations in genes *katG* and *rpoB*.

Molecular epidemiology has helped to deepen understanding of the TB epidemic in the studied regions. In this study, most of the isolates were from the western outbreak. Three cases from Samut Sakhon were excluded from the calculation

for the outbreak information. Our spoligotyping results indicated that more than 50% (35/72) of outbreak strain were clustered and belonged to the Beijing family. Based on the DST results and spoligotyping, we showed that the Beijing family was prevalent and that the majority of the strains were drug-resistant. In addition, almost all Beijing members in the current study belonged to the modern type, which was determined by SNP analysis at the position of 1477596 (14, 16, 43, 68). The modern Beijing strain has become a great concern in epidemiology because it is spreading worldwide and causing outbreaks in many countries (9, 23, 26, 32). In the current study, only one isolate was the ancestral Beijing type, which is usually associated with low virulence of TB (14, 35, 43). This finding was in agreement with previous reports of TB incidence in many countries where the Beijing type was predominant (23, 26, 32, 75). A recently study reported that about 66% of Beijing strains were drug-resistant and were prevalent in Thai TB patients (6). No information regarding the ancient or modern type were reported. Beijing strains have become a global health threat because of their frequent association with a high mutation rate, hypervirulence, immune evasion, treatment failure, and drug resistance (47).

In the current study, the analysis of molecular data based on 15-locus MIRU VNTR showed diverse genetic backgrounds among drug-susceptible Beijing strains. By contrast, MDR Beijing strains had restricted genetic backgrounds. As a result, we proposed that the method provided a discrimination power sufficiently high to differentiate MTB strains in the epidemic Thai areas. We found that the largest cluster of Beijing isolates circulating in these areas was the Bj-I subtype. Compared with a previous study (Figure 11), the specific Bj-I genotype of MDR-MTB was also found to be dominant, forming the largest cluster among Beijing strains. As shown in Figure 10A, the successful expansion of a specific clone of the Bj-I sub-lineage resulted in restricted genetic backgrounds. From this evidence, we proposed that the specific clone of Beijing sublineage with multidrug resistance was highly transmissible and may be the major cause of MDR-TB outbreaks in the Thamaka district. Our minimum span diagram clearly showed that predominant cluster-sharing genotypes, identical to past strains, have been continuously spreading to date. Rapid identification of TB cases, patient isolation, and appropriate treatments are required for a future, more effective outbreak control.

Molecular drug resistance analysis by DNA sequencing showed that genetic alterations were highly associated with drug-resistant phenotypes and identified common mutations in the *katG* and the *rpoB* genes. Additional mutations conferring

resistance to a fluoroquinolones and injectable aminoglycosides were further identified in two of three U SIT523 isolates, suggesting the emergence of XDR-TB. The absence of second-line drug resistance among a large number of MDR-TB genotypes suggested that under a MDR treatment, acquisition of additional resistance to second-line drugs was limited. MDR was also found in non-Beijing strains, but in a low percentage, accounting for merely 12.5% (4 of 22) (Figure 10). The occurrence of MDR with individual unique genotypes in Beijing and non-Beijing groups could be explained by the fact that MTB developed a spontaneous mutation or acquired drug resistance resulting in the development of new drug-resistant strains. On the other hand, it is a concern that the strains found in Thailand might be from Myanmar because a long boder is shared by the two countries and Myanmar is on of the top 20 global MDR-TB burdened countries and Myanmar is one of the top 20 global MDR-TB burdened countries; 9,000 MDR-TB cases occur in Myanmar each year (58, 81). However, based on data analysis from a previous report, there was no MDR-TB outbreak strain similar to the one in Kanchanaburi (58) Hence, The TB outbreak in Kanchanaburi Province probably emerged from loc al strains. Concerning the treatment and management of MDR- and XDR-TB, TB national guidelines corresponding to WHO commendations were followed. Overall, primary resistance seems to be a major driving force of the outbreak rather than development of secondary drug resistance by susceptible strains.

It should be noted that the current study has several limitations. The sample size of the MTB population was limited by the short time of the sample collection. In addition, the demographic and clinical data obtained for each patient was limited. Nonetheless, even though the number of samples were small, we were able to find clusters of MTB among the isolates from the Kanchanaburi Province. It is also worth noting that in neighbouring provinces, no identical genotypes to those identified in Kanchanaburi were identified. These results may imply that there was no TB spread across the provinces, but the evidence remains inconclusive due to the small sample sizes from those other provinces.

Our data suggest that the specific Beijing genotype strain with MDR was the most prevent MTB strain causing continuous TB transmission in the areas of Thailand studied. Our results also showed that this specific MTB strain, which was dominant, virulent and resistant to INH and RIF, has not been eliminated from the region. In addition, it was demonstrated that XDR-TB concurrently emerged and caused transmission. Finally, in the current study, we showed that genotyping based on

spoligotyping and 15 MIRU-VNTR analysis was a useful tool for successfully discrimination MTB strains.

Summary

Multidrug-resistant tuberculosis outbreaks have occurred in the Thamaka district, Kanchanaburi province in Thailand. This study aimed to characterise this pathogen and its genotypes. Seventy-two *Mycobacterium tuberculosis* isolates were collected and genotyped by spoligotyping 15-locus mycobacterial interspersed repetitive unit-variable-number tandem repeat (MIRU-VNTR) and single nucleotide polymorphism genotyping, and their drug resistance was analysed. The spoligotyping results showed that Beijing SIT1 was the predominant sub-lineage (n=38, 52.8%) while the remaining were non-Beijing sub-lineages (n=34). The MIRU-VNTR analysis showed thatthe most Beijing isolates (n=37) belonged to the modern type, forming 5 clusters and 13 individual patterns. In *katG*, only mutation Ser315Thr was identified. In *rpoB,* Ser531Leu was the predominant except His526Arg and Leu533Pro found in two isolates. A cluster of 14 Beijing strains contained these common mutations and shared the MIRU-VNTR genotype with isolates spreading previously in the Thamaka district. Two U SIT523 isolates contained mutations A1400G in *rrs* and Asp94Gly in *gyr*A genes, indicating a spread of extensively drug-resistant tuberculosis (XDR-TB). Most mutations were associated with drug resistance, and the specific MDR Beijing and XDR-TB in U SIT523 isolates are remaining. This genotyping was likely a key tool for tracking TB transmission in the Thamaka district, Thailand.

Keyword:MIRU-VNTR, Mutation, Outbreak, Spoligotyping, Thailand, Tuberculosi

Table 10. Phenotypic characteristics of MTB isolates according to drug susceptibility tests with first-line anti-TB drugs

Strain	Drug resistant patterns				Frequency		Proportion (%)	
	INH	**RIF**	**EMB**	**STM**				
Susceptible to INH and RIF					**45**		**62.5**	
	S	S	S	S	42		58.3	
	S	S	S	R	3		4.2	
Resistant to INH and RIF					**27**		**37.5**	
RIF mono-resistance	S	R	S	S	3	4	4.2	5.6
	S	R	S	R	1		1.4	
INH mono-resistance	R	S	S	R	1		1.4	
MDR and XDR	R	R	S	R	13		18.1	
	R	R	S	S	2	22	2.8	30.6
	R	R	R	R	7		9.7	
					72		**100**	

Abbreviations: MTB, *Mycobacterium tuberculosis*; INH, isoniazid; RIF, rifampicin; EMB, ethambutol; STM, streptomycin; S, susceptible; R, resistance; MDR, multi-drug resistance; XDR, extensively-drug resistance.

Table 11. Classification of MTB strains based on spoligotyping

No.	Clade	SIT	Octal Code	Frequency	%		Clustering rate (%)
				59			82
1	BEIJING	1	000000000003771	38		52.8	53
2	EAI2_NTB	89	674000003413771	6	8.3		8
3	EAI5	236	777777777413771	3	4.2		4
4	EAI5	256	777777777413671	4	5.6	11.2	6
5	EAI5	1395	474377767413771	1	1.4		
6	EAI6_BGD1	292	777777757413371	1	1.4		
7		591	777777757413771	2	2.8	5.6	3
8	EAI6_BGD1	1414	777757757413371	1	1.4		
9	H3	50	777777777720771	1	1.4		
10	T1	53	777777777760771	2	2.8		3
11	U	523	777777777777771	4	5.6		6
12	U	1189	677777477403771	1	1.4	8.4	
13	U	1391	777777700003371	1	1.4		
14	NEW1	ND	737777747413771	1	1.4		
15	NEW2	ND	12004377763771	1	1.4		
16	NEW3	ND	674000003412771	1	1.4		
17	NEW4	ND	367777670020731	1	1.4	9.8	
18	NEW5	ND	774177774000071	1	1.4		
19	NEW6	ND	000000007777711	1	1.4		
20	NEW7	ND	674000003413731	1	1.4		
	Total			**72**	**100**		

Abbreviations: SIT, spoligo-international type number; ND, no data

Table 12. Allelic distribution and discrimination power of MIRU-VNTR in Beijing genotypes resistant and susceptible to isoniazid and rifampicin

locus	DST result	N	MIRU-VNTR Copy number											Discreminatory power	
			1	2	3	4	5	6	7	8	9	10	ND	HGDI	Level
424	DR & MDR	19				19								0.00	Low
	Susceptible to INH&RIF	19		1	4	13	1							0.51	Moderate
580	DR & MDR	19		16									3	0.28	Low
	Susceptible to INH&RIF	19		18									1	0.11	Low
802	DR & MDR	19			16		3							0.28	Low
	susceptible to INH&RIF	19	1	3	15									0.37	Moderate
960	DR & MDR	19			19									0.00	Low
	Susceptible to INH&RIF	19		1	18									0.11	Low
1644	DR & MDR	19			19									0.00	Low
	Susceptible to INH&RIF	19			19									0.00	Low
1955	DR & MDR	19					19							0.00	Low
	Susceptible to INH&RIF	19		1	1	1	14	2						0.46	Moderate
2163b	DR & MDR	19				16	3							0.28	Low
	Susceptible to INH&RIF	19		1		3	6	9						0.68	High
2165	DR & MDR	19				19								0.00	Low
	Susceptible to INH&RIF	19				19								0.00	Low
2401	DR & MDR	19				19								0.00	Low
	Susceptible to INH&RIF	19				17	2							0.20	Low
2996	DR & MDR	19					1		18					0.11	Low
	Susceptible to INH&RIF	19				1	1		15	2				0.38	Moderate
3192	DR & MDR	19					19							0.00	Low
	Susceptible to INH&RIF	19					19							0.00	Low
3690	DR & MDR	19			19									0.00	Low
	Susceptible to INH&RIF	19			18	1								0.11	Low
4052	DR & MDR	19								19				0.00	Low
	Susceptible to INH&RIF	19							1	17		1		0.20	Low
4156	DR & MDR	19		19										0.00	Low
	Susceptible to INH&RIF	19		18			1							0.11	Low
4348	DR & MDR	19			19									0.00	Low
	Susceptible to INH&RIF	19			18	1								0.11	Low

Abbreviations:DR, drug resistance; HGDI, Hunter-Gaston discriminatory index; INH, isoniazid; MIRU-VNTR, mycobacterial interspersed repetitive unit-variable-number tandem repeat; RIF, rifampicin.

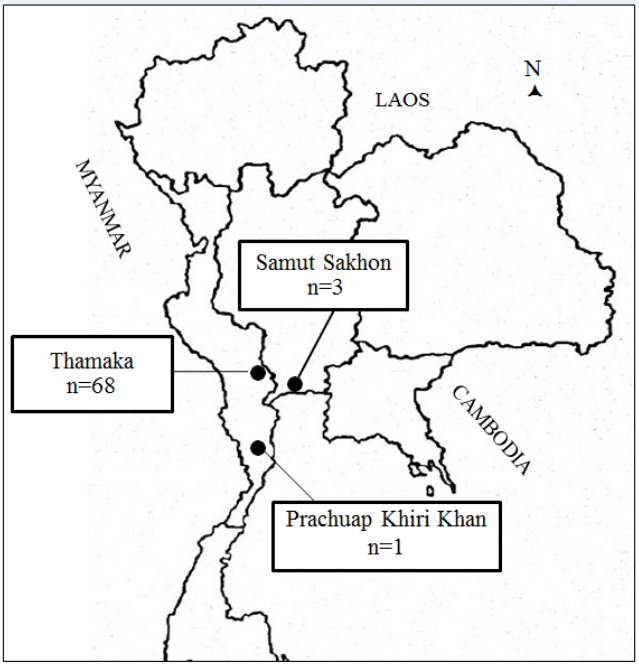


Figure 9. Map of sample collection sites and the number (n) of isolates from the Thamaka District, Kanchanaburi Province and surrounding areas in western Thailand, and Samut Sakhon and Prachuap Khiri Khan Provinces, Thailand.

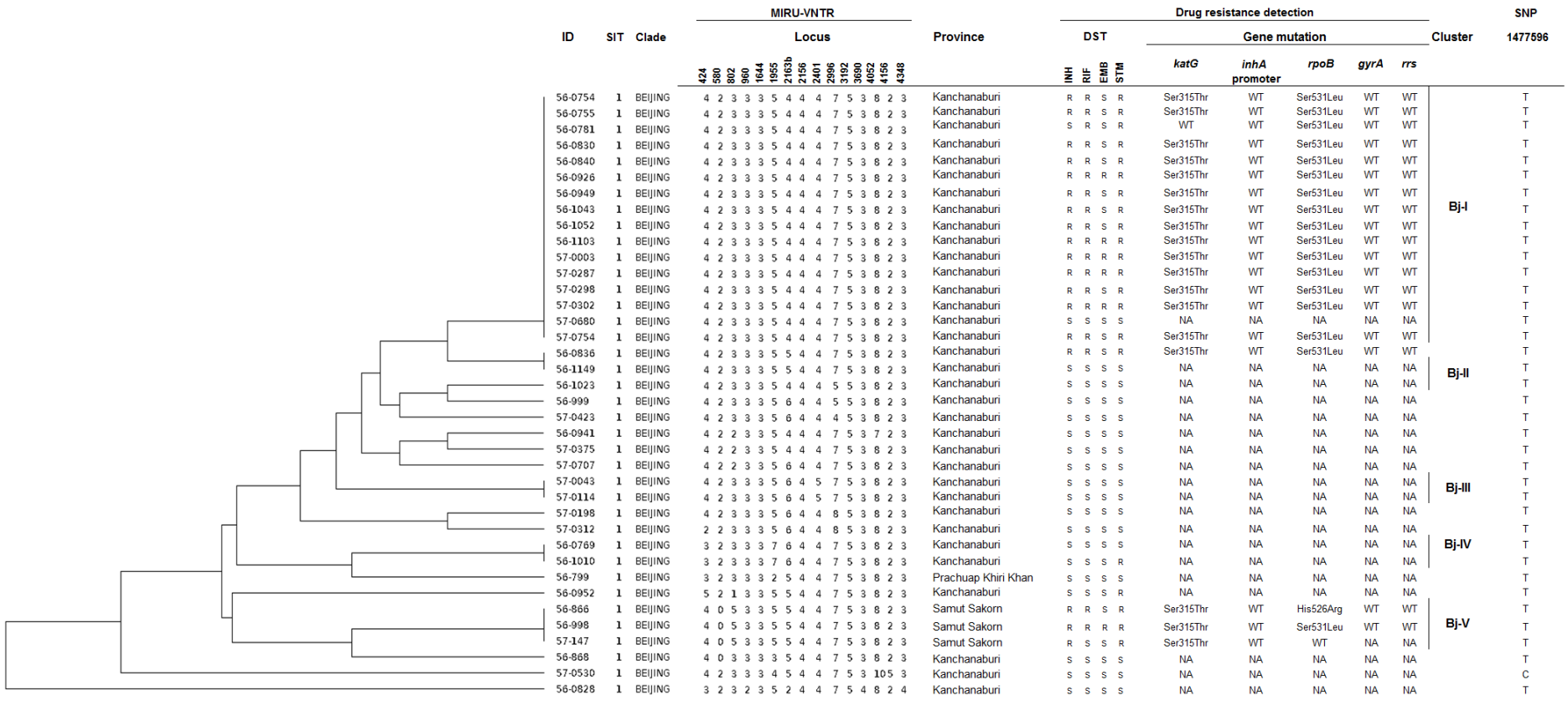
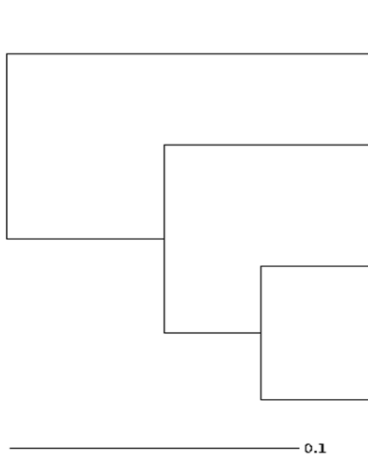


Figure 10A. Clustering analysis based on the 15 loci-MIRU-VNTR analysis. Spoligotypes, phenotypic and genotypic characteristics of drug resistance were incorporated. A-D, dendrogram of genetic relationships, A, 38 Beijing isolates



ID	Spoligotype		MIRU-VNTR profile														Province	DST				Gene mutation				
			Locus																							
	SIT	Clade	424	580	802	960	1644	1955	2163b	2156	2401	2996	3192	3690	4052	4156	4348		INH	RIF	EMB	STM	*katG*	*InhA promoter*	*rpoB*	*rrs*	*gyrA*
56-0760	89	EAI2_NTB	1	9	2	4	2	10	1	4	2	2	3	2	5	1	2	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-0854	89	EAI2_NTB	1	5	2	4	3	5	1	4	2	2	4	2	8	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
57-0169	89	EAI2_NTB	1	5	2	4	3	5	1	4	2	2	4	2	7	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
57-0708	89	EAI2_NTB	1	5	2	4	3	10	1	4	2	2	4	2	8	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
57-0309	89	EAI2_NTB	1	5	3	4	3	10	1	4	2	2	4	2	7	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-1070	89	EAI2_NTB	1	5	2	4	2	6	1	4	2	2	5	2	8	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-0763	53	T1	2	0	2	3	3	7	5	3	2	5	3	1	7	1	2	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
57-0197	53	T1	2	0	5	3	2	10	2	3	2	5	3	3	8	1	2	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-0764	236	EAI5	2	0	3	4	3	5	2	7	2	2	5	7	6	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-0986	236	EAI5	2	9	3	4	3	5	2	7	2	2	5	6	6	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-0947	256	EAI5	2	5	3	4	2	5	2	5	2	2	5	5	6	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
57-0115	256	EAI5	2	5	3	4	2	5	2	5	2	2	5	4	6	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-1029	256	EAI5	2	5	3	4	3	5	2	5	2	2	5	4	6	1	3	Kanchanaburi	S	S	S	R	NA	NA	NA	NA	NA
56-940	256	EAI5	2	6	3	4	2	5	2	7	2	2	5	5	5	1	3	Kanchanaburi	S	R	S	S	WT	WT	Ser531Leu	NA	NA
56-0797	236	EAI5	2	3	3	4	2	4	1	7	2	2	5	4	6	1	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-1027	591	EAI6_BGD1	2	7	3	4	4	6	9	7	2	2	5	4	6	1	3	Kanchanaburi	R	R	S	S	Ser315Thr	WT	Ser531Leu	WT	WT
56-1047	591	EAI6_BGD1	2	7	3	4	4	5	9	7	2	2	5	4	6	1	3	Kanchanaburi	S	R	S	S	WT	WT	Leu533Pro	WT	WT
56-0851	523	U	4	4	3	8	2	4	6	5	2	6	4	1	1	1	3	Kanchanaburi	R	R	R	R	Ser315Thr	WT	Ser531Leu	A1400G	Asp94Gly
56-1146	523	U	4	4	3	8	2	4	6	5	2	6	4	1	1	1	3	Kanchanaburi	R	R	S	S	Ser315Thr	WT	Ser531Leu	A1400G	Asp94Gly
56-1086	523	U	4	2	3	9	2	4	6	5	2	2	7	4	1	2	3	Kanchanaburi	S	S	S	S	NA	NA	NA	NA	NA
56-1028	523	U	4	4	3	9	3	5	6	4	2	6	5	3	3	2	3	Kanchanaburi	R	R	S	R	Ser315Thr	WT	Ser531Leu	WT	WT

Figure 10B. Clustering analysis based on the 15 loci-MIRU-VNTR analysis. Spoligotypes, phenotypic, and genotypic characteristics of drug resistance were incorporated. A-D, dendrogram of genetic relationships, B, of 21 non-Beijing isolates.

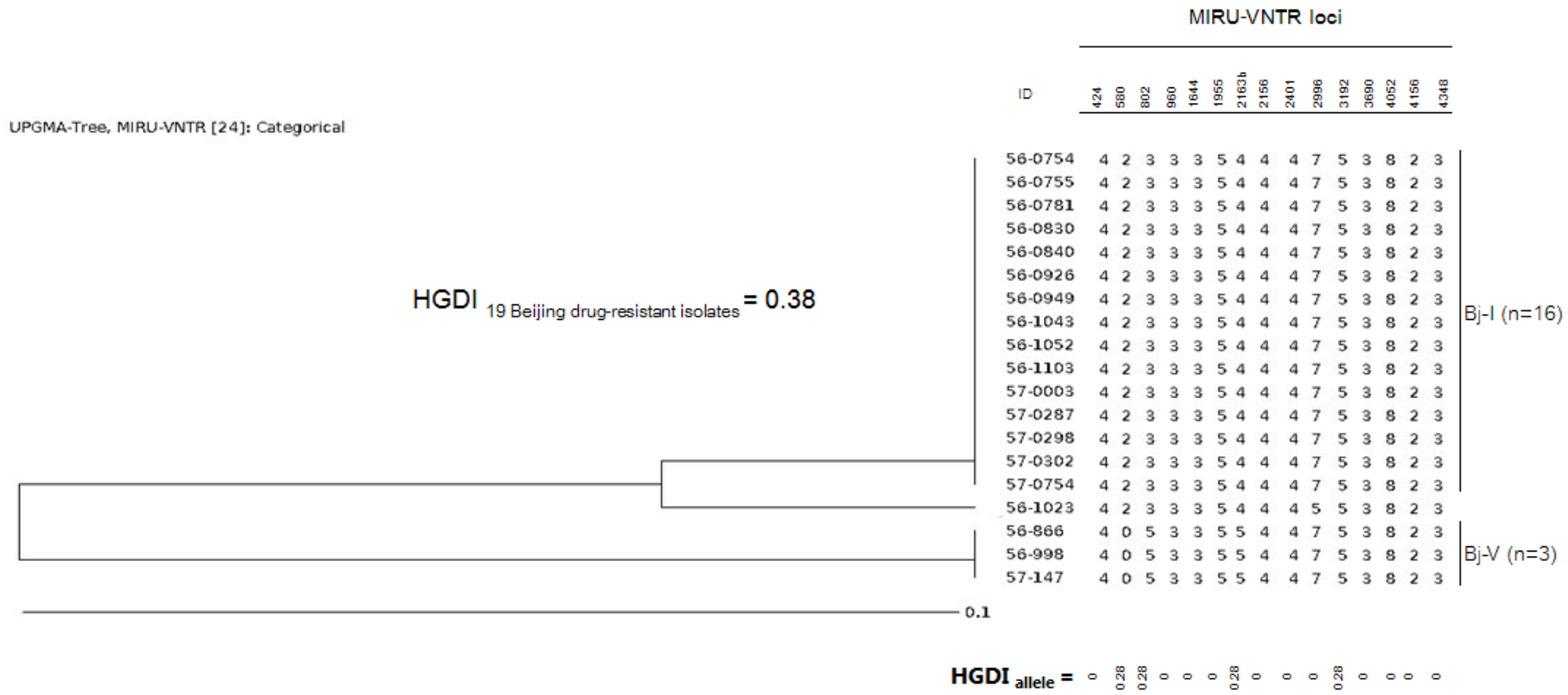


Figure 10C.

Clustering analysis based on the 15 loci-MIRU-VNTR analysis. Spoligotypes, phenotypic and genotypic characteristics of drug resistance were incorporated. A-D, dendrogram of genetic relationships, C, of drug-resistant Beijing strains;

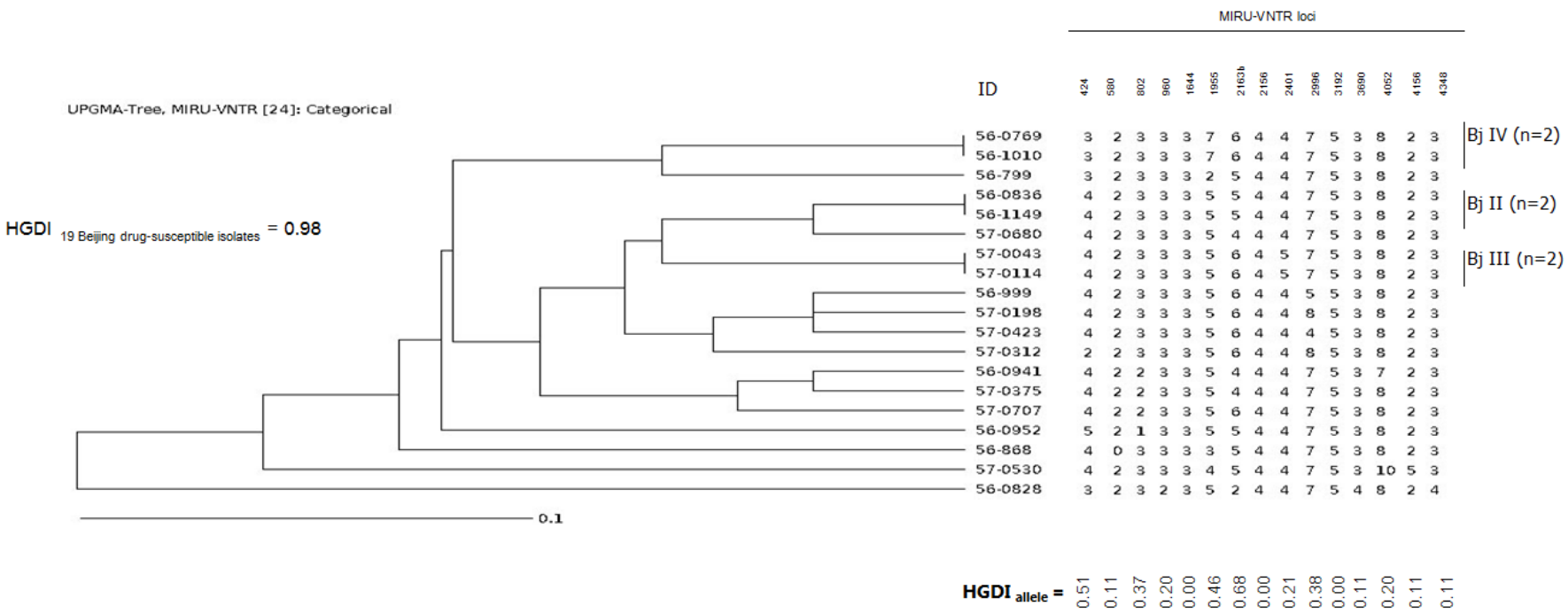


Figure 10D. Clustering analysis based on the 15 loci-MIRU-VNTR analysis. Spoligotypes, phenotypic and genotypic characteristics of drug resistance were incorporated. A-D, dendrogram of genetic relationships, D, of drug-susceptible Beijing strains.

Abbreviations: HGDI, Hunter-Gaston Discriminatory Indexes; ID, identification; SIT, spoligo-international types; DST, drug susceptibility testing; MIRU-VNTR, mycobacterial interspersed repetitive unit-variable number tandem repeat; SNP, single nucleotide polymorphism; INH, Isoniazid; RIF, Rifampicin; Kan, Kanamycin; S, Susceptible; R, Resistance; WT, wildtype; Ser, Serine; Thr, Threonine; His, Histidine; Arg, Arginine; Leu, Leucine; Asp, Asparagine; Gly, Glycine; A, Adenine; G, Guanine; NA, Not applicable.

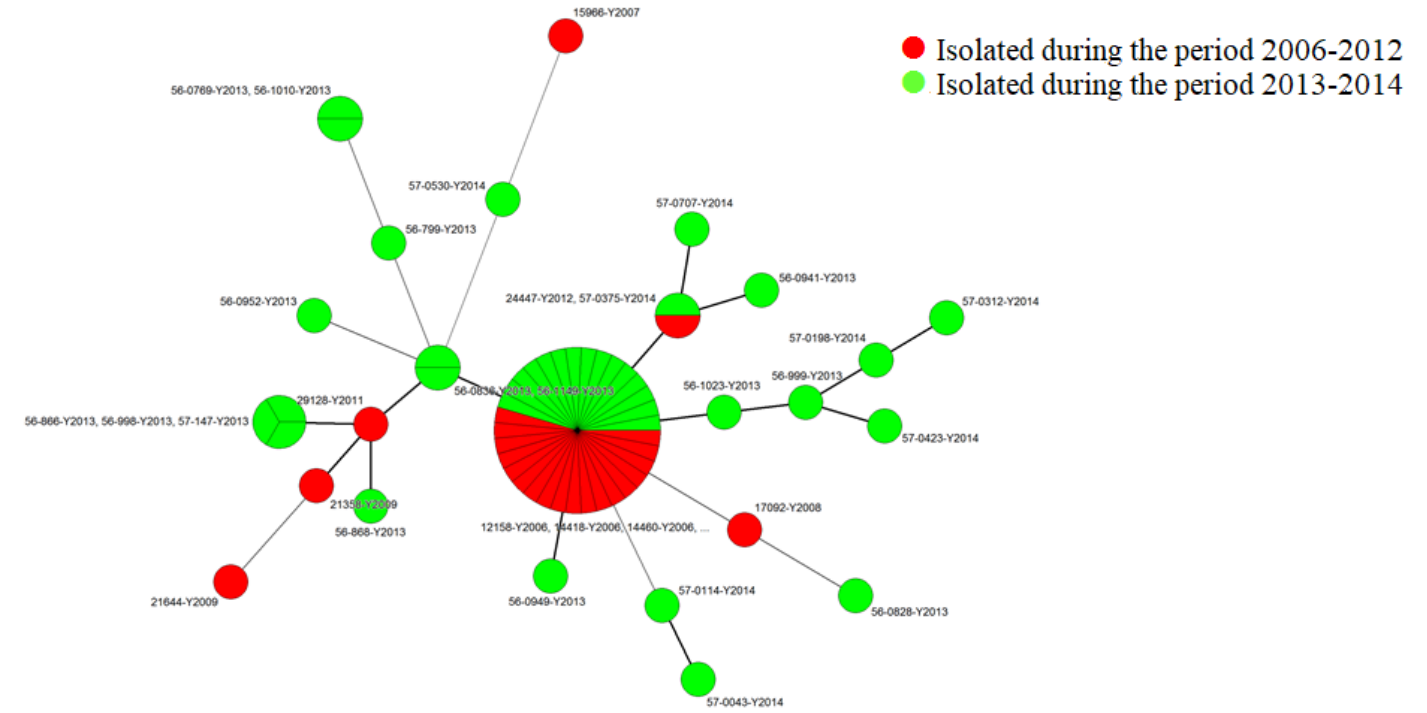


Figure 11. Comparison of genotypes amongst Beijing isolates in western outbreak areas in previous (during the year 2006-2012) and recent (during the year 2013-2014) studies, based on MIRU-VNTR; a minimum spanning tree was constructed based on 15 loci of MIRU-VNTR genotyping of the Beijing strains (n=24) isolated during the period 2006-2013, as reported in a previous study and the present work (n=38). Circles denote different types discriminated by 15 locus-MIRU-VNTR genotypes. The origins of each isolate are represented by different colours.

CONCLUSIONS

Because ofthe significant public health problem of TB and DR-TB in Thailand, the genetic characteristics of DR-MTBwere studied. A high rate of mutations conferring INH or RIF resistance was identifiedin this MTB population. The high percentage of mutations in the *katG*, *inhA* promoter, and the *rpoB* gene indicated that most INH- and RIF-resistant MTB in this setting could be detected by mutation analysis. Rare mutations raised concern for possible misdetection by routine molecular tests. The obtained data couldbenefit to adjustthe treatment regimen and controlthe transmission of drug-resistant TB in Thailand.

Spoligotyping and the analysis of specific 15-locus MIRU-VNTR revealed the genetic diversity and transmission links among MTB isolates. The Beijing genotype with MDR characteristics was the most prevalent MTB strain causing continuous TB transmission in the studied areas. The results also showed that thespecific MTB genotype which was dominant, virulent, and resistant to INH and RIF, has not been eliminated from the hotspot region. In addition, the concurrently emerged XDR-MTB genotype SIT523 additionally generated transmission of more threatening drug-resistant TB. MDR- and XDR-TB still continuously occur and may spread to neighboringand other areas in Thailand. The genotyping based on spoligotyping and 15-MIRU-VNTR analysis was proposed as a useful tool for discriminating local MTB strains.

The finding on molecular characteristics of DR-MTB can be applicable in developing new tools for detecting genetic determinants conferring resistance to anti-TB drugs, while most DR-MTB strains can be rapidly detected by molecular analysis. The epidemiological data on TB, and DR- TB is valuable to design appropriate interventions to control TB. The information generated from this study can serve as a basis to understand TB and DR-TB and help to plan effective strategies for TB control in Thailand.

ACKNOWLEDGEMENTS

This thesis was supported by the Japan Society for the Promotion of Science RONPAKU (Dissertation Ph.D.) Program (Grant No. NRCT−11141) to Ms. Janisara Rudeeaneksin, in part by a grant from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan, for the Joint Research Program of the Hokkaido University Research Program of the Hokkaido University International

Institute for Zoonosis Control to Professor Yasuhiko Suzuki, and in part by the Japan Agency for Medical Research and Development (AMED) [under grant number JP20jk021005, JP02jk021005, JP20jm011021 and JP20wrn0125008] to Professor Yasuhiko Suzuki, in the part by the budget of the National Institute of Health, Department of Medical Sciences, Ministry of Public Health, Thailand to Dr. Benjawan Phetsuksiri, Ph.D.

I would like to express my sincere gratitude and great appreciation to my supervisor, Professor Yasuhiko Suzuki, and Professor Chie Nakajima, for supervising me working on this interesting and challenging research and for their support and advice during my research and then writing up the thesis. I would like to thank my co-supervisor, Dr. Benjawan Phetsuksiri, for her guidance, support, and insightful comments. I am grateful to thanks Hokkaido University International Institute for Zoonosis Contro, Hokkaido University, Japan, for supporting research at Hokkaido University and for being helpful in the scientific analysis of molecular characteristics of MTB. I would like to express my deep gratitude to Professor Hideaki Higashi, Division of Infection and Immunity, Graduate School of Infectious Diseases,and Associate Professor Norikazu Isoda, Laboratory of Microbiology, Graduate School of Infectious diseases who gave valuable advice for correcting and improving the dissertation.

I am grateful to thank Ms. Haruka Suzuki and Ms. Yukari Fukushima, staffs of the Bioresource division, for their helpful technical assistance/advice about DNA sequencing and MIRU-VNTR typing and preparing academic documents for my work at Hokkaido University International Institute for Zoonosis Control and thank Dr. Yogendra Shah, for helpful suggestions on clustering analysis using the MIRU-VNTRplus program online. Additionally, I would like to thank all of the graduate students in the Bioresource division for sharing their knowledge during my research.

I am thankful to my colleague, Ms. Sopa Srisungngam, Ms. Supranee Bunchoo, and Mr. Wipat Klayut in the Mycobacteria group at the Sasakawa Research Building, the National Institute of Health, Thailand, for their help partially in DNA preparation, MIRU-VNTR analysis, PCR amplification, and DNA sequencing.

I am thankful to Dr. Nattakan Tipkrua, Mr. Krairerk Suthumfrom, and Dr. Anupong Sujariyakul from the Office of Disease Prevention and Control 5, Ratchaburi, Worasak Suthachai from the Office of Disease Prevention and Control Region 1, Chiang Mai, Mr. Nasron Jeklon from the Office of Disease Prevention and Control Region 12, Songkhla, for providing their valuable DNA of MTB. I am particularly thankful to Ms.

Sutudsanee Vimolsarte, a former senior medical technologist from the Regional Medical Sciences Center 7 KhonKaen, Thailand, for her valuable collaboration with the TB network in the northeast region and DNA samples.

Last but not least, I would also like to thank the numerous collaborators involved in our work, particularly various district hospitals in many provinces of Thailand, for the gift of their valuable collection of clinical strains.

REFERENCES

1) Aye KS, Nakajima C, Yamaguchi T, Win MM, Shwe MM, Win AA, Lwin T, Nyunt WW, Ti T, Suzuki Y.Genotypic characterization of multi-drugresistant *Mycobacterium tuberculosis* isolates in Myanmar. J Infect Chemother, 22, 174-179, 2016.
2) Anukool U, Phunpae P, Tharinjaroen CS, Butr-Indr B, Saikaew S, Netirat N, Intorasoot S, Suthachai V, Tragoolpua K, Chaiprasert A. Genotypic distribution and a potential diagnostic assay of multidrug-resistant tuberculosis in northern Thailand. Infect Drug Resist, 13, 3375-3382, 2020.
3) Boonaiam S, Chaiprasert A, Prammananan T, Leechawengwongs M. Genotypic analysis of genes associated with isoniazid and ethionamide resistance in MDR-TB isolates from Thailand. Clin Microbiol Infect, 16, 396-399, 2010.
4) Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, Allix C, Aristimuño L, Arora J, Baumanis V, Binder L, Cafrune P, Cataldi A, Cheong S, Diel R, Ellermeier C, Evans JT, Fauville-Dufaux M, Ferdinand S, Garcia de Viedma D, Garzelli C, Gazzola L, Gomes HM, Guttierez MC, Hawkey PM, van Helden PD, Kadival GV, Kreiswirth BN, Kremer K, Kubin M, Kulkarni SP, Liens B, Lillebaek T, Ho ML, Martin C, Martin C, Mokrousov I, Narvskaïa O, Ngeow YF, Naumann L, Niemann S, Parwati I, Rahim Z, Rasolofo-Razanamparany V, Rasolonavalona T, Rossetti ML, Rüsch-Gerdes S, Sajduda A, Samper S, Shemyakin IG, Singh UB, Somoskovi A, Skuce RA, van Soolingen D, Streicher EM, Suffys PN, Tortoli E, Tracevska T, Vincent V, Victor TC, Warren RM, Yap SF, Zaman K, Portaels F, Rastogi N, Sola C. *Mycobacterium tuberculosis* complex genetic diversity: Mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol 23, 1-17, 2006.

5) Charoenpak R, Santimaleeworagun W, Suwanpimolkul G, Manosuthi W, Kongsanan P, Petsong S, Puttilerpong C. Association between the phenotype and genotype of isoniazid resistance among *Mycobacterium tuberculosis* isolates in Thailand. Infect Drug Resist13, 627-634, 2020.

6) Cheunoy W, Haile M, Chaiprasert A, Prammananan T, Cristea-Fernström M, Vondracek M, Chryssanthou E, Hoffner S, Petrini B. Drug resistance and genotypic analysis of *Mycobacterium tuberculosis* strains from Thai tuberculosis patients. APMIS, 117, 286-290, 2009.

7) Cohen KA, Bishai WR, Pym AS. Molecular basis of drug resistance in *Mycobacterium tuberculosis*. Microbiol Spectr, 2, 1-15, 2014.

8) Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. PLOS ONE, 4, e7815, 2009.

9) Couvin D, Rastogi N. Tuberculosis a global emergency: tools and methods to monitor, understand, and control the epidemic with specific example of the Beijing lineage. Tuberculosis, 95, S177-S189, 2015.

10) Disratthakit A, Meada S, Prammananan T, Thaipisuttikul I, Doi N, Chaiprasert A. Genotypic diversity of multidrug-, quinolone- and extensively drug-resistant *Mycobacterium tuberculosis* isolates in Thailand. Infect Genet Evol, 2, 432-439, 2015.

11) Dooley KE, Lahlou O, Ghali I, Knudsen J, Elmessaoudi MD, Cherkaoui I, El Aouad R. Risk factors for tuberculosis treatment failure, default, or relapse and outcomes of retreatment in Morocco. BMC Public Health, 11, 140, 2011.

12) Eldholm V, Matee M, Mfinanga SG, Heun M, Dahle UR. A first insight into the genetic diversity of *Mycobacterium tuberculosis* in Dares Salaam, Tanzania, assessed by spoligotyping. BMC Microbiol, 6, 76, 2006.

13) Fàbrega A, Madurga S, Giralt E, Vila J. Mechanism of action and resistance to quinolones. Microb Biotechnol, 2, 40–61, 2009.

14) Faksri K, Drobniewski F, Nikolayevskyy V, Brown T, Prammananan T, Palittapongarnpim P, Prayoonwiwat N, Chaiprasert A. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on IS6110, SNP, LSP and VNTR profiles from Thailand. Infect Genet Evol, 11, 1142-1149, 2011.

15) Frothingham R, Meeker-O'Connell WA. Genetic diversity in the Mycobacterium tuberculosis complex based on variable numbers of tandem DNA repeats. Microbiology, 144, 1189-1196, 1998.

16) Filliol I, Motiwala AS, Cavatore M, Qi W, Hazbón MH, Bobadilla del Valle M, Fyfe J, García-García L, Rastogi N, Sola C, Zozio T, Guerrero MI, León CI, Crabtree J, Angiuoli S, Eisenach KD, Durmaz R, Joloba ML, Rendón A, Sifuentes-Osornio J, Ponce de León A, Cave MD, Fleischmann R, Whittam TS, Alland D. 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, 2006.

17) Gandhi NR, Brust JCM, Shah NS. A new era fortreatment of drug-resistant tuberculosis. Eur Respir J, 52, 1801350, 2018.

18) Grandjean L, Moore DA. Tuberculosis in the developing world: recent advances in diagnosis with special consideration of extensively drug-resistant tuberculosis. Curr Opin Infect Dis, 21, 454-461, 2008.

19) Guerra-Assunção JA, Crampin AC, Houben RM, Mzembe T, Mallard K, Coll F, Khan P, Banda L, Chiwaya A, Pereira RP, McNerney R, Fine PE, Parkhill J, Clark TG, Glynn JR. Large-scale whole genome sequencing of *M. tuberculosis* provides insights into transmission in a high prevalence area. eLife, 3, e05166, 2015.

20) Hall A. A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser (Oxford). BioEdit: 1999.

21) Hunter PR, Gaston MA. Numerical index of the discriminatory ability of typing system: an application of Simpson's index of diversity. J Clin Microbiol, 26, 2465-2466, 1988.

22) Iem V, Somphavong S, Buisson Y, Steenkeste N, Breysse F, Chomarat M, Sylavanh P, Nanthavong P, Rajoharison A, Berland JL, Paboriboune P. Resistance of *Mycobacterium tuberculosis* to antibiotics in Lao PDR: first multicentric study conducted in 3 hospitals. BMC Infect Dis, 13, 275, 2013.

23) Iwamoto T, Grandjean L, Arikawa K, Nakanishi N, Caviedes L, Coronel J, Sheen P, Wada T, Taype CA, Shaw MA, Moore DA, Gilman RH. Genetic diversity and transmission characteristics of Beijing family strains of *Mycobacterium tuberculosis* in Peru. PLOS ONE, 7, e49651, 2012.

24) O'Neill J. Tackling drug-resistant infections globally: Final report and recommendation. In the review on antimicrobial resistance. Wellcome Trust, London. pp. 1-88, 2016.

25) Jiraphongsa C, Wangteeraprasert T, Henpraserttae N, Sanguanwongse N. Panya, L. Sukkasitvanichkul J, Pittayawonganon C. Community outbreak of multidrug resistance tuberculosis, Kanchanaburi Province, Thailand on 2002-June 2010. J Prev Med Assoc Thail, 3, 261-271, 2011.

26) Johnson R, Warren RM, van der Spuy GD, Gey van Pittius NC, Theron D, Streicher EM, Bosman M, Coetzee GJ, van Helden PD, Victor TC. Drug-resistant tuberculosis epidemic in the Western Cape driven by a virulent Beijing genotype strain. Int J Tuberc Lung Dis, 14, 119-121, 2010.

27) Kalokhe AS, Shafiq M, Lee JC, Ray SM, Wang YF, Metchock B, Anderson AM, Nguyen ML. Multidrug-resistant tuberculosis drug susceptibility and molecular diagnostic testing. Am J Med Sci, 345, 143-148, 2013.

28) Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, Bunschoten A, Molhuizen H, Shaw R, Goyal M, van Embden J. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol, 35, 907-914, 1997.

29) Klopper M, Warren RM, Hayes C, Gey van Pittius NC, Streicher EM, Müller B, Sirgel FA, Chabula-Nxiweni M, Hoosain E, Coetzee G, David van Helden P, Victor TC, Trollip AP. Emergence and spread of extensively and totally drug-resistant tuberculosis, South Africa. Emerg Infect Dis, 19, 449-455, 2013.

30) Lee AS, Lim IH, Tang LL, Telenti A, Wong SY. Contribution of *kasA* analysis to detection of isoniazid-resistant *Mycobacterium tuberculosis* in Singapore. Antimicrob Agents Chemother, 43, 2087-2089, 1999.

31) Lee AS, Lim IH, Tang LL, Wong SY. High frequency of mutations in the *rpoB* gene in rifampin-resistant clinical isolates of *Mycobacterium tuberculosis* from Singapore. J Clin Microbiol, 43, 2026-2027, 2005.

32) Li D, Dong CB, Cui JY, Nakajima C, Zhang CL, Pan XL, Sun GX, Dai EY, Suzuki Y, Zhuang M, Ling H. Dominant modern sublineages and a new modern sublineage of *Mycobacterium tuberculosis* Beijing family clinical isolates in Heilongjiang Province, China. Infect Genet Evol, 27, 294-299, 2014.

33) Li G, Zhang J, Guo Q, Wei J, Jiang Y, Zhao X, Zhao LL, Liu Z, Lu J, Wan K. Study of efflux pump gene expression in RIF-mono resistant *Mycobacterium tuberculosis* clinical isolates. J Antibiot, 68, 431–435, 2015.

34) Li Y, Cao X, Li S, Wang H, Wei J, Liu P, Wang J, Zhang Z, Gao H, Li M, Wan K, Dai E. Characterization of *Mycobacterium tuberculosis* isolates from Hebei, China: Genotypes and drug susceptibility phenotypes. BMC Infect Dis, 16, 107, 2016.

35) Li Y, Pang Y, Zhang T, Xian X, Yang J, Wang R, Wang P, Zhang M, Chen W. Genotypes of *Mycobacterium tuberculosis* isolates circulating in Shaanxi Province, China. PLOS ONE, 15, e0242971, 2020.

36) Lu J, Liu M, Wang Y, Pang Y, Zhao Z. Mechanisms of fluoroquinolone monoresistance in *Mycobacterium tuberculosis*. FEMS Microbiol Lett, 353, 40–48, 2014.

37) Kato-Maeda M, Metcalfe JZ, FloresL. Genotyping of *Mycobacterium tuberculosis*: application in epidemiologic studies. Future Microbiol, 6, 203–216, 2011.

38) Meaza A, Kebede A, Yaregal Z, Dagne Z, Moga S, Yenew B, Diriba G, Molalign H, Tadesse M, Adisse D, Getahun M, Desta K. Evaluation of genotype MTBDR*plus* VER 2.0 line probe assay for the detection of MDR-TB in smear positive and negative sputum samples. BMC Infect Dis, 17, 280, 2017.

39) Minh NN, Van Bac N, Son NT, Lien VT, Ha CH, Cuong NH, Mai CT, Le TH. Molecular characteristics of rifampin-and isoniazid-resistant *Mycobacterium tuberculosis* strains isolated in Vietnam**.** J Clin Microbiol, 50, 598-601, 2012.

40) Müller B, Borrell S, Rose G, Gagneux S. The heterogeneous evolution of multidrug-resistant *Mycobacterium tuberculosis*. Trends Genet, 29, 160-169, 2013.

41) Munir A, Kumar N, Ramalingam SB, Tamilzhalagan S, Shanmugam SK, Palaniappan AN, Nair D, Priyadarshini P, Natarajan M, Tripathy S, Ranganathan UD, Peacock SJ, Parkhill J, Blundell TL, Malhotra S. Identification and characterization of genetic determinants of isoniazid and rifampicin resistance in *Mycobacterium tuberculosis* in southern India. Sci Rep, 9, 10283, 2019.

42) Musser JM. Antimicrobial agent resistance in mycobacteria: molecular genetic insights. Clin Microbiol Rev, 8, 496-514, 1995.

43) Nakajima C, Tamaru A, Rahim Z, Poudel A, Maharjan B, Khin Saw Aye, Ling H, Hattori T, Iwamoto T, Fukushima Y, Suzuki H, Suzuki Y, Matsuba T. Simple multiplex PCR assay for identification of Beijing family *Mycobacterium tuberculosis* isolates with a lineage-specific mutation in Rv0679c. J Clin Microbiol, 51, 2025-2032, 2013.

44) Nonghanphithak D, Chaiprasert A, Smithtikarn S, Kamolwat P, Pungrassami P, Chongsuvivatwong V, Mahasirimongkol S, Reechaipichitkul W, Leepiyasakulchai C, Phelan JE, Blair D, Clark TG, Faksri K. Clusters of drug-resistant *Mycobacterium tuberculosis* detected by whole-genome sequence analysis of nationwide sample, Thailand, 2014–2017. Emerg Infect Dis, 27, 813-822, 2021.

45) Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, Ginsberg A, Swaminathan S, Spigelman M, Getahun H, Menzies D, Raviglione M. Tuberculosis. Nat Rev Dis Primers, 2, 16076, 2016.

46) Palomino J, Martin A. Drug resistance mechanisms in *Mycobacterium tuberculosis*. Antibiotics, 3, 317-340, 2014.

47) Parwati I, Alisjahbana B, Apriani L, Soetikno RD, Ottenhoff TH, van der Zanden AG, van der Meer J, van Soolingen D, van Crevel R. *Mycobacterium tuberculosis* Beijing genotype is an independent risk factor for tuberculosis treatment failure in Indonesia. J Infect Dis, 201, 553-557, 2010.

48) Poudel A, Maharjan B, Nakajima C, Fukushima Y, Pandey BD, Beneke A, Suzuki Y. Characterization of extensively drug-resistant *Mycobacterium tuberculosis* in Nepal. Tuberculosis, 93, 84-88, 2013.

49) Poudel A, Nakajima C, Fukushima Y, Suzuki H, Pandey BD, Maharjan B, Suzuki Y. Molecular characterization of multidrug-resistant *Mycobacterium tuberculosis* isolated in Nepal. Antimicrob Agents Chemother, 56, 2831-2836, 2012.

50) Prammananan T, Cheunoy W, Taechamahapun D, Yorsangsukkamol J, Phunpruch S, Phdarat P, Leechawengwong M, Chaiprasert A. Distribution of *rpoB* mutation among multidrug-resistant *Mycobacterium tuberculosis* (MDR-TB) strains from Thailand and development of a rapid method for mutation detection. Clin Microbiol Infect, 14, 446-453, 2018.

51) Ramaswamy S, Musser JM. Molecular genetic basis of antimicrobial agent resistance in *Mycobacterium tuberculosis*: 1998 update. Tuber Lung Dis, 79, 3–29, 1998.

52) Ramaswamy SV, Reich R, Dou SJ, Jasperse L, Pan X, Wanger A, Quitugua T, Graviss EA. Single nucleotide polymorphisms in genes associated with isoniazid resistance in *Mycobacterium tuberculosis*. Antimicrob Agents Chemother, 47, 1241-1250, 2003.

53) Regmi SM, Chaiprasert A, Kulawonganunchai S, Tongsima S, Coker OO, Prammananan T, Viratyosin W, Thaipisuttikul I. Whole genome sequence

analysis of multidrug-resistant *Mycobacterium tuberculosis* Beijing isolates from an outbreak in Thailand. Mol Genet Genomics, 290, 1933-1941, 2015.

54) Reilly K. Tuberculosis: Current treatments and investigational therapies. US Pharm, 39, 3-7, 2014.
55) Rasoahanitralisoa R, Rakotosamimanana N, Stucki D, Sola C, Gagneux S, Rasolofo Razanamparany V. Evaluation of spoligotyping SNPs and customised MIRU-VNTR combination for genotyping *Mycobacterium tuberculosis* clinical isolates in Madagascar. PLOS ONE, 12, e0186088, 2017.
56) Rudeeaneksin J, Klayut W, Srisungngam S, Bunchoo S, Toonkomdang S, Wongchai T, Chuenchom N, Phetsuksiri B. Putative extensive and pre-extensive drug resistant-tuberculosis associated with unusual genotypes on the Thailand-Myanmar border. Rev Inst Med Trop Sao Paulo, 63, e85, 2021.
57) Rudeeaneksin J, Phetsuksiri B, Nakajima C, Bunchoo S, Suthum K, Tipkrua N, Fukushima Y, Suzuki Y. Drug-resistant *Mycobacterium tuberculosis* and its genotypes isolated from an outbreak in western Thailand. Trans R Soc Trop Med Hyg, 115, 886-895, 2020.
58) San LL, Aye KS, Oo NAT, Shwe MM, Fukushima Y, Gordon SV, Suzuki Y, Nakajima C. Insight into multidrug-resistant Beijing genotype *Mycobacterium tuberculosis* isolates in Myanmar. Int J Infect Dis, 76, 109-119, 2018.
59) Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Tuberculosis drug resistance mutation database. PLOS Med, 6, e1000002, 2009.
60) Sarkar S, Ganguly A, Sunwoo HH. Current overview of anti-tuberculosis drugs: metabolism and toxicities. Mycobact Dis, 6, 2, 2018.
61) Sanchez-Padilla E, Merker M, Beckert P, Jochims F, Dlamini T, Kahn P, Bonnet M, Niemann S. Detection of drug-resistant tuberculosis by Xpert MTB/RIF in Swaziland. N Engl J Med, 372, 1181–1182, 2015.
62) Seung KJ, Keshavjee S, Rich ML. Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis. Cold Spring Harb Perspect Med, 5, a017863, 2015.
63) Singh A, Prasad R, Balasubramanian V, Gupta N. Drug-resistant tuberculosis and HIV Infection: current perspectives. HIV/AIDS, 12, 9–31, 2020.
64) Singh R, Dwivedi SP, Gaharwar US, Meena R, Rajamani P, Prasad T. Recent updates on drug resistance in *Mycobacterium tuberculosis*. J Appl Microbiol, 128, 1547-1567, 2020.

65) Siu GK, Zhang Y, Lau TC, Lau RW, Ho PL, Yew WW, Tsui SK, Cheng VC, Yuen KY, Yam WC. Mutations outside the rifampicin resistance-determining region associated with rifampicin resistance in *Mycobacterium tuberculosis*, J Antimicrob Chemothe 66, 730-733, 2011.

66) Sola C, Filliol I, Legrand E, Lesjean S, Locht C, Supply P, Rastogi N. Genotyping of the *Mycobacterium* tuberculosis complex using MIRUs: association with VNTR and spoligotyping for molecular epidemiology and evolutionary genetics. Infect Genet Evol, 3, 125-133, 2003.

67) Spinato J, Boivin É, Bélanger-Trudelle É, Fauchon H, Tremblay C, Soualhine H. Genotypic characterization of drug resistant *Mycobacterium tuberculosis* in Quebec, 2002-2012. BMC Microbiol, 16, 164, 2016.

68) Srilohasin P, Prammananan T, Faksri K, Phelan JE, Suriyaphol P, Kamolwat P, Smithtikarn S, Disratthakit A, Regmi SM, Leechawengwongs M, Twee-Hee Ong R, Teo YY, Tongsima S, Clark TG, Chaiprasert A. Genomic evidence supporting the clonal expansion of extensively drug-resistant tuberculosis bacteria belonging to rare proto-Beijing genotype. Emerg Microbes Infect, 9, 2632-2641, 2020.

69) Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rüsch-Gerdes S, Willery E, Savine E, de Haas P, van Deutekom H, Roring S, Bifani P, Kurepina N, Kreiswirth B, Sola C, Rastogi N, Vatin V, Gutierrez MC, Fauville M, Niemann S, Skuce R, Kremer K, Locht C, van Soolingen D. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis.* J Clin Microbiol, 44, 4498-4510, 2006.

70) Surcouf C, Heng S, Pierre-Audigier C, Cadet-Daniel V, Namouchi A, Murray A, Gicquel B, Guillard B. Molecular detection of fluoroquinolone-resistance in multi-drug resistant tuberculosis in Cambodia suggests low association with XDR phenotypes. BMC Infect Dis, 11, 255, 2011.

71) Suthum K, Samosornsuk W, Samosornsuk S. Characterization of *katG*, *inhA*, *rpoB* and *pncA* in *Mycobacterium tuberculosis* isolates from MDR-TB risk patients in Thailand. J Infect Dev Ctries, 14, 268-276, 2020.

72) Suzuki Y, Katsukawa C, Tamaru A, Abe C, Makino M, Mizuguchi Y, Taniguchi H. Detection of kanamycin-resistant *Mycobacterium tuberculosis* by identifying mutations in the 16S rRNA Gene. J Clin Microbiol, 36, 1220–1225, 1998.

73) The Operational Plan to End Tuberculosis 2017 -2021 (https://tbthailand.org /download/ Manual/Thailand) [accessed 5 October 2021]

74) Unissa AN, Subbian S, Hanna LE, Selvakumar N. Overview on mechanisms of isoniazid action and resistance in *Mycobacterium tuberculosis*. Infect Genet Evol, 45, 474-492, 2016.

75) Wang J, Liu Y, Zhang CL, Ji BY, Zhang LZ, Shao YZ, Jiang SL, Suzuki Y, Nakajima C, Fan CL, Ma YP, Tian GW, Hattori T, Ling H. Genotypes and characteristic of clustering and drug susceptibility of *Mycobacterium tuberculosis* isolates collected in Heilongjiang province, China. J Clin Microbiol, 49, 1354-1362, 2011.

76) Weniger T, Krawczyk J, Supply P Niemann S, Harmsen D. MIRU-VNTRplus: A web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. Nucleic Acids Res, 38, W326-W331, 2010.

77) World Health Organization. Global tuberculosis report 2016. Geneva. Switzerland: World Health Organization, 2016.

78) World Health Organization. Global tuberculosis report 2017. Geneva. Switzerland: World Health Organization, 2017.

79) World Health Organization. Global tuberculosis report 2018. Geneva. Switzerland: World Health Organization, 2018.

80) World Health Organization. Global tuberculosis report 2019. Geneva. Switzerland: World Health Organization, 2019.

81) World Health Organization. Global tuberculosis report 2020. Geneva. Switzerland: World Health Organization, 2020.

82) World Health Organization. Global tuberculosis report 2021. Geneva. Switzerland: World Health Organization, 2021.

83) World Health Organization. Global tuberculosis report 2022. Geneva. Switzerland: World Health Organization, 2022.

84) World Health Organization. Meeting report of the WHO expert consultation on the definition of extensively drug-resistant tuberculosis, 27-29 October 2020.

85) Yoon JH, Nam JS, Kim KJ, Choi Y, Lee H, Cho SN, Ro YT. Molecular characterization of drug-resistant and susceptible *Mycobacterium tuberculosis* isolated from patients with tuberculosis in Korea. Diagn Microbiol Infect Dis, 72, 52-56, 2012.

86) Zhao LL, Chen Y, Liu HC, Xia Q, Wu XC, Sun Q, Zhao XQ, Li GL, Liu ZG, Wan KL. Molecular characterization of multidrug-resistant *Mycobacterium tuberculosis* isolates from China. Antimicrob Agents Chemother, 58, 1997-2005, 2014.

APPENDIX A

Sequences of primer (5/ to 3/) used for spoligotyping, MIRU-VNTR, SNP typing and drug resistance

Objective	Gene /locus	Primer Sequence
Spoligotyping		
	Direct region	F:5'-GGTTTTGGGTCTGACGAC-3'
		R:5'-CCGAGAGGGGACGGAAAC-3'
MIRU-VNTR		
	MIRU 4 (ETR-D)	F: 5'-GCGCGAGAGCCCGAACTGC-3'
		R: 5'- GCGCGAGCAGAAACGCCAGC-3'
	MIRU10	F:5'-GTTCTTGACCAACTGCAGTCGTCC-3'
		R: 5'-GCCACCTTGGTGATCAGCTACCT-3'
	MIRU16	F:5'-TCGGTGATCGGGTCCAGTCCAAGT-3'
		R: 5'-CCCGTCGTGCAGCCCTGGTAC-3'
	MIRU26	F: 5'-TAGGTCTACCGTCGAAATCTGTGAC-3'
		R: 5'-CATAGGCGACCAGGCGAATAG -3'
	MIRU31(ETR-E)	F: 5'-ACTGATTGGCTTCATACGGCTTTA-3'
		R: 5'-GTGCCGACGTGGTCTTGAT-3'
	MIRU39	F: 5'-CGCATCGACAAACTGGAGCCAAAC-3'
		R: 5'-CGGAAACGTCTACGCCCCACACAT-3'
	MIRU 40	F: 5'-GGGTTGCTGGATGACAACGTGT-3'
		R: 5'-GGGTGATCTCGGCGAAATCAGATA-3'
	ETR-A (VN2165)	F: 5'-AAATCGGTCCCATCACCTTCTTAT-3'
		R: 5'-CGAAGCCTGGGGTGCCCGCGATTT-3'
	QUB1b (VN2163b)	F: 5'-CGTAAGGGGGATGCGGGAAATAGG-3'
		R: 5'-CGAAGTGAATGGTGGCAT-3'
	QUB26 (VN4052)	F: 5'-GGCCAGGTCCTTCCCGAT-3'
		R: 5'-AACGCTCAGCTGTCGGAT-3'
	VN424 (Mtub04)	F: 5'-CAGTCCAGGTTGCAAGAGATGG-3'
		R: 5'-CGGCATCCTCAACAACGGTAGC-3'
	VN1955 (Mtub21)	F: 5'-AGATCCCAGTTGTCGTCGTC-3'
		R: 5'-CAACATCGCCTGGTTCTGTA-3'
	VN2401(Mtub30)	F: 5'-CTTGAAGCCCCGGTCTCATCTGT-3'
		R:5'-ACTTGAACCCCCACGCCCATTAGTA 3'
	VN3690(Mtub39)	F: 5'-CGGTGGAGGCGATGAACGTCTTC-3'
		R:5'-TAGAGCGGCACGGGGGAAAGCTTAG-3'
	VN4156 (QUB4156)	F: 5'-TGACCACGGATTGCTCTAGT-3'
		R: 5'-GCCGGCGTCCATGTT-3'
SNP typing for identified ancestral		
	1477596	F:5'-GTCGACAGCGCCAGAAAATG -3'
		R:5'-GCTCCTATGCCACCCAGCAC -3'
Drug resistance		
RIF	*rpoB* (1276 -1356)	F:5'-CAGGACGTGGAGGCGATCAC-3'*
		R:5'-GAGCCGATCAGACCGATGTGG-3'
INH	*katG* (812 -834)	F:5'-ATGGCCATGAACGACGTCGAAAC-3'*
		R:5'-CGCAGCGAGAGGTCAGTGGCCAG-3'
	InhA (-50 - -1)	F:5'-TCACACCGACAAACGTCACGAGC-3'*
		R:5'-AGCCAGCCGCTGTGCGATCGCCA-3'
Fluoroquinolone	*gyrA* (QRDR,220-339)	F:5'-AGCGCAGCTACATCGACTATGCG -3'*
		R:5'-CTTCGGTGTACCTCATCGCCGCC -3'
Kanamycin	*rrs* (1350-1550)	F:5'-CGGATCGGGGTCTGCAACTCGAC-3'*
		R:5'-CAAGAACCCCTCACGGCCTACG -3'

* primer for sequencing

APPENDIX B

The reference sequences of target gene (*rpoB*, *inh*A promotor, *kat*G, *gyr*A, *rrs*) for nucleotide alignment

>*rpo*B <H37Rv>

ttggcagattcccgccagagcaaaacagccgctagtcctagtccgagtcgcccgcaaagt
tcctcgaataactccgtacccggagcgccaaaccgggtctccttcgctaagctgcgcgaa
ccacttgaggttccgggactccttgacgtccagaccgattcgttcgagtggctgatcggt
tcgccgcgctggcgcgaatccgccgccgagcggggtgatgtcaacccagtgggtggcctg
gaagaggtgctctacgagctgtctccgatcgaggacttctccgggtcgatgtcgttgtcg
ttctctgaccctcgtttcgacgatgtcaaggcacccgtcgacgagtgcaaagacaaggac
atgacgtacgcggctccactgttcgtcaccgccgagttcatcaacaacaacaccggtgag
atcaagagtcagacggtgttcatgggtgacttcccgatgatgaccgagaagggcacgttc
atcatcaacgggaccgagcgtgtggtggtcagccagctggtgcggtcgcccggggtgtac
ttcgacgagaccattgacaagtccaccgacaagacgctgcacagcgtcaaggtgatcccg
agccgcggcgcgtggctcgagtttgacgtcgacaagcgcgacaccgtcggcgtgcgcatc
gaccgcaaacgccggcaaccggtcaccgtgctgctcaaggcgctgggctggaccagcgag
cagattgtcgagcggttcgggttctccgagatcatgcgatcgacgctggagaaggacaac
accgtcggcaccgacgaggcgctgttggacatctaccgcaagctgcgtccgggcgagccc
ccgaccaaagagtcagcgcagacgctgttggaaaacttgttcttcaaggagaagcgctac
gacctggcccgcgtcggtcgctataaggtcaacaagaagctcgggctgcatgtcggcgag
cccatcacgtcgtcgacgctgaccgaagaagacgtcgtggccaccatcgaatatctggtc
cgcttgcacgagggtcagaccacgatgaccgttccgggcggcgtcgaggtgccggtggaa
accgacgacatcgaccacttcggcaaccgccgcctgcgtacggtcggcgagctgatccaa
aaccagatccgggtcggcatgtcgcggatggagcgggtggtccgggagcggatgaccacc
caggacgtggaggcgatcacaccgcagacgttgatcaacatccggccggtggtcgccgcg
atcaaggagttcttcggcaccagccagctgagccaattcatggaccagaacaacccgctg
tcggggttgacccacaagcgccgactgtcggcgctggggcccggcggtctgtcacgtgag
cgtgccgggctggaggtccgcgacgtgcacccgtcgcactacggccggatgtgcccgatc
gaaacccctgaggggcccaacatcggtctgatcggctcgctgtcggtgtacgcgcgggtc
aacccgttcgggttcatcgaaacgccgtaccgcaaggtggtcgacggcgtggttagcgac
gagatcgtgtacctgaccgccgacgaggaggaccgccacgtggtggcacaggccaattcg
ccgatcgatgcggacggtcgcttcgtcgagccgcgcgtgctggtccgccgcaaggcgggc
gaggtggagtacgtgccctcgtctgaggtggactacatggacgtctcgccccgccagatg
gtgtcggtggccaccgcgatgattcccttcctggagcacgacgacgccaaccgtgccctc
atgggggcaaacatgcagcgccaggcggtgccgctggtccgtagcgaggccccgctggtg
ggcaccgggatggagctgcgcgcggcgatcgacgccggcgacgtcgtcgtcgccgaagaa

agcggcgtcatcgaggaggtgtcggccgactacatcactgtgatgcacgacaacggcacc
cggcgtacctaccggatgcgcaagtttgcccggtccaaccacggcacttgcgccaaccag
tgccccatcgtggacgcgggcgaccgagtcgaggccggtcaggtgatcgccgacggtccc
tgtactgacgacggcgagatggcgctgggcaagaacctgctggtggccatcatgccgtgg
gagggccacaactacgaggacgcgatcatcctgtccaaccgcctggtcgaagaggacgtg
ctcacctcgatccacatcgaggagcatgagatcgatgctcgcgacaccaagctgggtgcg
gaggagatcacccgcgacatcccgaacatctccgacgaggtgctcgccgacctggatgag
cggggcatcgtgcgcatcggtgccgaggttcgcgacggggacatcctggtcggcaaggtc
accccgaagggtgagaccgagctgacgccggaggagcggctgctgcgtgccatcttcggt
gagaaggcccgcgaggtgcgcgacacttcgctgaaggtgccgcacggcgaatccggcaag
gtgatcggcattcgggtgttttcccgcgaggacgaggacgagttgccggccggtgtcaac
gagctggtgcgtgtgtatgtggctcagaaacgcaagatctccgacggtgacaagctggcc
ggccggcacggcaacaagggcgtgatcggcaagatcctgccggttgaggacatgccgttc
cttgccgacggcaccccggtggacattattttgaacacccacggcgtgccgcgacggatg
aacatcggccagattttggagacccacctgggttggtgtgcccacagcggctggaaggtc
gacgccgccaaggggttccggactgggccgccaggctgcccgacgaactgctcgaggcg
cagccgaacgccattgtgtcgacgccggtgttcgacggcgcccaggaggccgagctgcag
ggcctgttgtcgtgcacgctgcccaaccgcgacggtgacgtgctggtcgacgccgacggc
aaggccatgctcttcgacgggcgcagcggcgagccgttcccgtacccggtcacggttggc
tacatgtacatcatgaagctgcaccacctggtggacgacaagatccacgcccgctccacc
gggccgtactcgatgatcacccagcagccgctgggcggtaaggcgcagttcggtggccag
cggttcggggagatggagtgctgggccatgcaggcctacggtgctgcctacaccctgcag
gagctgttgaccatcaagtccgatgacaccgtcggccgcgtcaaggtgtacgaggcgatc
gtcaagggtgagaacatcccggagccgggcatccccgagtcgttcaaggtgctgctcaaa
gaactgcagtcgctgtgcctcaacgtcgaggtgctatcgagtgacggtgcggcgatcgaa
ctgcgcgaaggtgaggacgaggacctggagcgggccgcggccaacctgggaatcaatctg
tcccgcaacgaatccgcaagtgtcgaggatcttgcgtaa

inhA promoter region (-50 -1)

gtggacataccgatttcggcccggccgcggcgagacgataggttgtcggg

katG <H37Rv>

gtgcccgagcaacacccacccattacagaaaccaccaccggagccgctagcaacggctgt
cccgtcgtgggtcatatgaaataccccgtcgagggcggcggaaaccaggactggtggccc
aaccggctcaatctgaaggtactgcaccaaaacccggccgtcgctgacccgatgggtgcg
gcgttcgactatgccgcggaggtcgcgaccatcgacgttgacgccctgacgcgggacatc
gaggaagtgatgaccacctcgcagccgtggtggcccgccgactacggccactacgggccg
ctgtttatccggatggcgtggcacgctgccggcacctaccgcatccacgacggccgcggc
ggcgccgggggcggcatgcagcggttcgcgccgcttaacagctggcccgacaacgccagc
ttggacaaggcgcgccggctgctgtggccggtcaagaagaagtacggcaagaagctctca
tgggcggacctgattgttttcgccggcaactgcgcgctggaatcgatgggcttcaagacg
ttcgggttcggcttcggccgggtcgaccagtgggagcccgatgaggtctattggggcaag

gaagccacctggctcggcgatgagcgttacagcggtaagcgggatctggagaacccgctg
gccgcggtgcagatggggctgatctacgtgaacccggaggggccgaacggcaacccggac
cccatggccgcggcggtcgacattcgcgagacgtttcggcgcatggccatgaacgacgtc
gaaacagcggcgctgatcgtcggcggtcacactttcggtaagacccatggcgccggcccg
gccgatctggtcggccccgaacccgaggctgctccgctggagcagatgggcttgggctgg
aagagctcgtatggcaccggaaccggtaaggacgcgatcaccagcggcatcgaggtcgta
tggacgaacaccccgacgaaatgggacaacagtttcctcgagatcctgtacggctacgag
tgggagctgacgaagagccctgctggcgcttggcaatacaccgccaaggacggcgccggt
gccggcaccatcccggacccgttcggcgggccagggcgctccccgacgatgctggccact
gacctctcgctgcgggtggatccgatctatgagcggatcacgcgtcgctggctggaacac
cccgaggaattggccgacgagttcgccaaggcctggtacaagctgatccaccgagacatg
ggtcccgttgcgagataccttgggccgctggtccccaagcagaccctgctgtggcaggat
ccggtccctgcggtcagccacgacctcgtcggcgaagccgagattgccagccttaagagc
cagatccgggcatcgggattgactgtctcacagctagtttcgaccgcatgggcggcggcg
tcgtcgttccgtggtagcgacaagcgcggcggcgccaacggtggtcgcatccgcctgcag
ccacaagtcgggtgggaggtcaacgaccccgacggggatctgcgcaaggtcattcgcacc
ctggaagagatccaggagtcattcaactccgcggcgccggggaacatcaaagtgtccttc
gccgacctcgtcgtgctcggtggctgtgccgccatagagaaagcagcaaaggcggctggc
cacaacatcacggtgcccttcaccccgggccgcacggatgcgtcgcaggaacaaaccgac
gtggaatcctttgccgtgctggagcccaaggcagatggcttccgaaactacctcggaaag
ggcaacccgttgccggccgagtacatgctgctcgacaaggcgaacctgcttacgctcagt
gcccctgagatgacggtgctggtaggtggcctgcgcgtcctcggcgcaaactacaagcgc
ttaccgctgggcgtgttcaccgaggcctccgagtcactgaccaacgacttcttcgtgaac
ctgctcgacatgggtatcacctgggagccctcgccagcagatgacgggacctaccagggc
aaggatggcagtggcaaggtgaagtggaccggcagccgcgtggacctggtcttcgggtcc
aactcggagttgcgggcgcttgtcgaggtctatggcgccgatgacgcgcagccgaagttc
gtgcaggacttcgtcgctgcctgggacaaggtgatgaacctcgacaggttcgacgtgcgc
tga

>*gyrA*<CDC1551>

atgacagacacgacgttgccgcctgacgactcgctcgaccggatcgaaccggttgacatc
cagcaggagatgcagcgcagctacatcgactatgcgatgagcgtgatcgtcggccgcgcg
ctgccggaggtgcgcgacgggctcaagcccgtgcatcgccgggtgctctatgcaatgttc
gattccggcttccgcccggaccgcagccacgccaagtcggcccggtcggttgccgagacc

atgggcaactaccacccgcacggcgacgcgtcgatctacgacaccctggtgcgcatggcc
cagccctggtcgctgcgctacccgctggtggacggccagggcaacttcggctcgccaggc
aatgacccaccggcggcgatgaggtacaccgaagcccggctgaccccgttggcgatggag
atgctgagggaaatcgacgaggagacagtcgatttcatccctaactacgacggccgggtg
caagagccgacggtgctacccagccggttccccaacctgctggccaacgggtcaggcggc
atcgcggtcggcatggcaaccaatatcccgccgcacaacctgcgtgagctggccgacgcg
gtgttctgggcgctggagaatcacgacgccgacgaagaggagaccctggccgcggtcatg
gggcgggttaaaggcccggacttcccgaccgccggactgatcgtcggatcccagggcacc
gctgatgcctacaaaactggccgcggctccattcgaatgcgcggagttgttgaggtagaa
gaggattcccgcggtcgtacctcgctggtgatcaccgagttgccgtatcaggtcaaccac
gacaacttcatcacttcgatcgccgaacaggtccgagacggcaagctggccggcatttcc
aacattgaggaccagtctagcgatcgggtcggtttacgcatcgtcatcgagatcaagcgc
gatgcggtggccaaggtggtgatcaataacctttacaagcacacccagctgcagaccagc
tttggcgccaacatgctagcgatcgtcgacggggtgccgcgcacgctgcggctggaccag
ctgatccgctattacgttgaccaccaactcgacgtcattgtgcggcgcaccacctaccgg
ctgcgcaaggcaaacgagcgagcccacattctgcgcggcctggttaaagcgctcgacgcg
ctggacgaggtcattgcactgatccgggcgtcggagaccgtcgatatcgcccgggccgga
ctgatcgagctgctcgacatcgacgagatccaggcccaggcaatcctggacatgcagttg
cggcgcctggccgcactggaacgccagcgcatcatcgacgacctggccaaaatcgaggcc
gagatcgccgatctggaagacatcctggcaaaacccgagcggcagcgtgggatcgtgcgc
gacgaactcgccgaaatcgtggacaggcacggcgacgaccggcgtacccggatcatcgcg
gccgacggagacgtcagcgacgaggatttgatcgcccgcgaggacgtcgttgtcactatc
accgaaacgggatacgccaagcgcaccaagaccgatctgtatcgcagccagaaacgcggc
ggcaagggcgtgcagggtgcggggttgaagcaggacgacatcgtcgcgcacttcttcgtg
tgctccacccacgatttgatcctgttcttcaccacccagggacgggtttatcgggccaag
gcctacgacttgcccgaggcctcccggacggcgcgcgggcagcacgtggccaacctgtta
gccttccagcccgaggaacgcatcgcccaggtcatccagattcgcggctacaccgacgcc
ccgtacctggtgctggccactcgcaacgggctggtgaaaaagtccaagctgaccgacttc
gactccaatcgctcgggcggaatcgtggcggtcaacctgcgcgacaacgacgagctggtc
ggtgcggtgctgtgttcggccgacgacgacctgctgctggtctcggccaacgggcagtcc
atcaggttctcggcgaccgacgaggcgctgcggccaatgggtcgtgccacctcgggtgtg
cagggcatgcggttcaatatcgacgaccggctgctgtcgctgaacgtcgtgcgtgaaggc
acctatctgctggtggcgacgtcaggggctatgcgaaacgtaccgcgatcgaggaatac
ccggtacagggccgcggcggtaaaggtgtgctgacggtcatgtacgaccgccggcgcggc
aggttggttggggcgttgattgtcgacgacgacagcgagctgtatgccgtcacttccggc
ggtggcgtgatccgcaccgcggcacgccaggttcgcaaggcgggacggcagaccaagggt
gttcggttgatgaatctgggcgagggcgacacactgttggccatcgcgcgcaacgccgaa
gaaagtggcgacgataatgccgtggacgccaacggcgcagaccagacgggcaattaa

>*rrs*<CDC1551>

Tttgtttggagagtttgatcctggctcaggacgaacgctggcggcgtgcttaacacat
Gcaagtcgaacggaaaggtctcttcggagatactcgagtggcgaacgggtgagtaaca
Cgtgggtgatctgccctgcacttcgggataagcctgggaaactgggtctaataccgga
Taggaccacgggatgcatgtcttgtggtggaaagcgctttagcggtgtgggatgagcc
Cgcggcctatcagcttgttggtggggtgacggcctaccaaggcgacgacgggtagccg
Gcctgagagggtgtccggccacactgggactgagatacggcccagactcctacgggag
Gcagcagtggggaatattgcacaatgggcgcaagcctgatgcagcgacgccgcgtggg
Ggatgacggccttcgggttgtaaacctctttcaccatcgacgaaggtccgggttctct
Cggattgacggtaggtggagaagaagcaccggccaactacgtgccagcagccgcggta
Atacgtagggtgcgagcgttgtccggaattactgggcgtaaagagctcgtaggtggtt
Tgtcgcgttgttcgtgaaatctcacggcttaactgtgagcgtgcgggcgatacgggca
Gactagagtactgcaggggagactggaattcctggtgtagcggtggaatgcgcagata
Tcaggaggaacaccggtggcgaaggcgggtctctgggcagtaactgacgctgaggagc
Gaaagcgtggggagcgaacaggattagataccctggtagtccacgccgtaaacggtgg
Gtactaggtgtgggtttccttccttgggatccgtgccgtagctaacgcattaagtacc
Ccgcctggggagtacggccgcaaggctaaaactcaaaggaattgacgggggcccgcac
Aagcggcggagcatgtggattaattcgatgcaacgcgaagaaccttacctgggtttga
Catgcacaggacgcgtctagagataggcgttcccttgtggcctgtgtgcaggtggtgc
Atggctgtcgtcagctcgtgtcgtgagatgttgggttaagtcccgcaacgagcgcaac
Ccttgtctcatgttgccagcacgtaatggtggggactcgtgagagactgccggggtca
Actcggaggaaggtggggatgacgtcaagtcatcatgcccccttatgtccagggcttca
Cacatgctacaatggccggtacaaagggctgcgatgccgcgaggttaagcgaatcctt
Aaaagccggtctcagttcggatcggggtctgcaactcgaccccgtgaagtcggagtcg
Ctagtaatcgcagatcagcaacgctgcggtgaatacgttcccgggccttgtacacacc
Gcccgtcacgtcatgaaagtcggtaacacccgaagccagtggcctaaccctcgggagg
Gagctgtcgaaggtgggatcggcgattgggacgaagtcgtaacaaggtagccgtaccg
gaaggtgcggctggatcacctcctttct

APPENDIX C

PCR product size (bp) and corresponding tandem-repeat number (MIRU-VNTR)

No.	Locus name	unit (bp)	Tandem repeat number for MTB										
			0	1	2	3	4	5	6	7	8	9	10
1	MIRU10	53	482	537	590	643	696	749	802	855	908		
2	MIRU16	53	565	618	671	724	777	830	883	936	989		
3	MIRU26	51	461	512	563	614	665	716	767	818	869		
4	MIRU31(ETR-E)	53	492	545	598	651	704	757	810	863	916		
5	MIRU40	54	354	408	462	516	570	624	678	732	786		
6	VN424(Mtu04)	51	170	221	272	323	374	425	476	527	578	629	680
7	VN1955(Mtub21)	57		149	206	263	320	377	434	491	548	605	
8	VN3690(Mtub39)	58		330	388	446	504	562	620	678	736		
9	ETR-A(VN2165)	75	195	270	345	420	495	570	645	720	795	870	945
10	ETR-C(VN577)	58	44	102	160	218	276	334	392	450	508		
11	QUB11b(VN2163b)	69		136	205	274	343	412	481	550	619		
12	QUB26(VN4052)	111		264	375	486	597	708	819	930	1041		
13	MIRU04(ETR-D)	77		252	329	406	483	560	637				
14	VN2401(Mtub30)	58	247	305	363	421	479	537	595	653	711		
15	VN4156(QUB4156)	59	563	622	681	740	799	858	917	976			
16	MIRU02	53	402	455	508	561	614	667	720	773	826		
17	MIRU20	77		514	591	668							
18	MIRU23	53	63	116	169	222	275	328	381	434	487	540	593
19	MIRU24	54	395	447	501	555	609	663	717	771	825		
20	MIRU27	53		551	604	657	710	763	816	869	922		
21	MIRU39	53	536	591	646	701	756	811	866	921	976		
22	ETR-B	57	121	178	235	292	349	406	463	520	577		
23	ETR-F	79	239	318	397	476	555	634					
24	QUB11a	69			305	374	443	512	581	650	719	788	857
25	QUB3232	56	156	212	268	324	380	436	492	548	604	660	716
26	QUB3336	59		171	230	289	348	407	466	525	584	643	702