MOLECULAR EPIDEMIOLOGY AND GENETIC ANALYSIS OF MULTIDRUG-/PRE-EXTENSIVELY DRUG/EXTENSIVELY DRUG-RESISTANT TUBERCULOSIS IN THAILAND DURING 2014-2017

MR. DITTHAWAT NONGHANPHITHAK

A THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

KHON KEAN UNIVERSITY

2021

MOLECULAR EPIDEMIOLOGY AND GENETIC ANALYSIS OF MULTIDRUG-/PRE-EXTENSIVELY DRUG/EXTENSIVELY DRUG-RESISTANT TUBERCULOSIS IN THAILAND DURING 2014-2017

MR. DITTHAWAT NONGHANPHITHAK

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN MEDICAL MICROBIOLOGY

GRADUATE SCHOOL KHON KAEN UNIVERSITY

2021

THESIS APPROVAL

KHON KAEN UNIVERSITY

FOR

DOCTOR OF PHILOSOPHY

IN MEDICAL MICROBIOLOGY

Thesis Title: Molecular epidemiology and genetic analysis of multidrug-/ pre-extensively drug/extensively drug-resistant tuberculosis in Thailand during 2014-2017

Author: Mr. Ditthawat Nonghanphithak

Thesis Examination Committee:

Assoc. Prof. Dr. Angkana Chaiprasert	Chairperson
Prof. Wipa Reechaipichitkul	Member
Assist. Prof. Dr. Wises Namwat	Member
Assoc. Prof. Dr. Kiatichai Faksri	Member
Dr. Kanin Salao	Member

Thesis Advisor: Kiatichai Faksri Digitally signed by เกียรติชัย ฟักศรี Date: 2021.06.12 22:34:06 +07'00'

..... Advisor

(Assoc. Prof. Dr. Kiatichai Faksri)

.....

(Prof. Dr. Surasakdi Wongratanacheewin)

Dean, Graduate School

.....

(Assoc. Prof. Apichat Jiravuttipong)

Dean, Faculty of Medicine

ดิษฐวัฒน์ หนองหารพิทักษ์. 2564. **ระบาดวิทยาเชิงโมเลกุลและการวิเคราะห์ทางพันธุกรรมของเชื้อวัณโรคดื้อยาหลายขนาน หลายขนานชนิดรุนแรง และหลายขนานชนิดรุนแรงมากในประเทศไทย ในช่วงปี 2557-2560.** วิทยานิพนธ์ปริญญาปรัชญาดุษฎีบัณฑิต สาขาวิชาจุลชีววิทยาทางการแพทย์ บัณฑิตวิทยาลัย มหาวิทยาลัยขอนแก่น.
อาจารย์ที่ปรึกษาวิทยานิพนธ์: รศ.ดร. เกียรติไชย ฟักศรี

บทคัดย่อ

วัณโรคที่มีสาเหตุการติดเชื้อจากแบคทีเรีย มัยโคแบคทีเรียม ทูเบอร์คูโลซิส เป็นปัญหาหลักของสาธารณสุขที่สำคัญอย่างยิ่งในประเทศไทย สายพันธุ์ของเชื้อวัณโรคดื้อยาหลายขนานชนิดรุนแรง (MDR) และหลายขนานชนิดรุนแรงมาก (pre-XDR และ XDR) เป็นสาเหตุที่ทำให้การควบคุมวัณโรคมีความซับซ้อนมากยิ่งขึ้น การตรวจวิเคราะห์ด้วยการตรวจหาลำดับนิวคลิโอไทด์ทั้งจีโนม (whole-genome sequencing (WGS)) ของเชื้อวัณโรคดื้อยาที่กระจายอยู่ในชุมชนจะทำให้สามารถเห็นรูปแบบการแพร่กระจายของเชื้อดังกล่าวพร้อมแนวทางในการบริหารจัดการทางคลินิกได้อย่างเหมาะสม นอกจากนี้ เทคนิค WGS สามารถใช้เพื่อทดสอบความไวต่อยาแบบจีโนไทป์ที่มีความละเอียดและครอบคลุมสูงกว่าเทคนิคอณูชีววิทยาอื่น ๆ อย่างไรก็ตาม ในประเทศไทย ข้อมูลทางระบาดวิทยาเชิงโมเลกุลของเชื้อวัณโรคดื้อยาในระดับประเทศยังไม่เป็นที่ทราบแน่ชัด นอกจากนี้ความรู้ในเรื่องความสัมพันธ์ระหว่างการกลายพันธุ์ที่เกี่ยวข้องกับการดื้อยาและผลการทดสอบความไวต่อยาแบบฟีโนทัยป์ยังพบว่ามีอยู่จำกัด โดยเฉพาะข้อมูลของยาต้านวัณโรคสูตรสอง

เพื่อทราบถึงสถานการณ์การระบาดของวัณโรคดื้อยาในประเทศไทย คณะผู้วิจัยได้วิเคราะห์ข้อมูล WGS ของเชื้อวัณโรคดื้อยาที่ถูกเก็บมาแบบสุ่มจำนวน 579 ตัวอย่าง โดยเป็นเชื้อที่เพาะแยกเชื้อจากผู้ป่วยวัณโรคดื้อยา (วินิจฉัยด้วยการทดสอบความไวต่อยาแบบฟีโนทัยป์ด้วยวิธี LJ proportion test) ที่รักษาอยู่ในโรงพยาบาล 230 แห่ง ครอบคลุมมากถึง 71 จังหวัด ทั่วประเทศ ระหว่างปี พ.ศ. 2557-2560 จากข้อมูลพบว่า เชื้อวัณโรคดื้อยาส่วนใหญ่เป็นชนิด MDR 466 (80.5%) ตัวอย่าง ตามด้วยชนิด pre-XDR 81 (14.0%) ตัวอย่าง และชนิด XDR 32 (5.5%) ตัวอย่าง โดยเชื้อวัณโรคดื้อยาส่วนใหญ่เป็นสายพันธุ์ East-Asian lineage (Linage2) จำนวน 482 (83.2%) ตัวอย่าง จากการวิเคราะห์กลุ่มของการระบาดพบว่ากลุ่มที่เคยมีการระบาด (Clade) ของเชื้อวัณโรคดื้อยามีจำนวน 13 clades ซึ่งมีความสัมพันธ์อย่างมีนัยยะสำคัญกับภูมิศาสตร์ ($p < 0.001$) ทั้งนี้ เมื่อวิเคราะห์โดยใช้เกณฑ์แบบละเอียด (cut off ≤11 SNPs) พบว่ากลุ่มที่มีการระบาด (Cluster) ของเชื้อวัณโรคดื้อยามีจำนวนทั้งสิ้น 89 clusters (n = 281; 48.5%) ทั่วประเทศ ประกอบไปด้วยเชื้อ

วัณโรคดื้อยาที่ถูกวินิจฉัยด้วยจีโนทัยป์ ชนิด MDR จำนวน 205 ตัวอย่าง ชนิด pre-XDR 46 ตัวอย่าง ชนิด XDR 19 ตัวอย่าง และชนิด poly-drug resistant 11 ตัวอย่าง โดยเชื้อวัณโรคดื้อยาส่วนใหญ่ใน clusters จะพบรูปแบบการกลายพันธุ์ในยีนที่สัมพันธ์กับการดื้อยาที่คล้ายคลึงกัน ซึ่งเป็นไปได้ว่ามีวัณโรคดื้อยาชนิดที่เกิดขึ้นในผู้ป่วยที่ไม่เคยได้รับการรักษามาก่อน (possible primary drug resistance) แบ่งเป็นวัณโรคชนิด MDR (n = 176/205; 85.9%) pre-XDR (n = 29/46; 63.0%) และ XDR (n = 14/19; 73.7%) นอกจากนี้ ปัจจัยเสี่ยงที่สัมพันธ์กับการระบาดของเชื้อวัณโรคดื้อยา ได้แก่ การพำนักอาศัยในจังหวัดของภาคตะวันตกของประเทศ (OR 2.4, 95%CI 1.5-3.9) การติดเชื้อวัณโรคสายพันธุ์ lineage2.2.1 (OR 3., 95%CI 2.4-5.3) เป็นต้น

เพื่อเพิ่มขีดความสามารถของวิธี WGS ในการทดสอบความไวต่อยาสำหรับวินิจฉัยวัณโรคดื้อยา คณะผู้วิจัยทำการเปรียบเทียบผลการทดสอบความไวต่อยาด้วยวิธีต่าง ๆ (แบบฟีโนทัยป์ด้วยวิธี agar proportion test และ minimum inhibitory concentration (MIC) test และแบบจีโนทัยป์ด้วยวิธี WGS) ของเชื้อวัณโรคดื้อยาจำนวน 60 สายพันธุ์ ประกอบไปด้วย วัณโรคดื้อยาชนิด poly-drug resistant 1 ตัวอย่าง ชนิด MDR 34 ตัวอย่าง และชนิด XDR 25 ตัวอย่าง จากข้อมูลพบค่า agreement ที่สูงระหว่างวิธี WGS และวิธี MIC ของยาต้านวัณโรคเป็นส่วนใหญ่ ยกเว้นยา ethambutol (65%) และยา ethionamide (62%) เมื่อวิเคราะห์ผลการทดสอบความไวต่อยาด้วยวิธี MIC กับตำแหน่งการกลายพันธุ์ของเชื้อวัณโรคดื้อยาพบว่า เชื้อวัณโรคดื้อยาที่พบการกลายพันธุ์ตำแหน่ง -15 c/t *inhA* มีระดับ MIC ของยา isoniazid อยู่ในระดับต่ำอย่างมีนัยยะสำคัญเมื่อเทียบกับเชื้อวัณโรคดื้อยาที่พบการกลายพันธุ์ตำแหน่ง *katG* Ser315Thr (p <0.001) นอกจากนี้รูปแบบของระดับ MIC ดังกล่าว ยังสามารถพบได้ในยา ethambutol (ตำแหน่ง *embB* Gly406Asp เทียบกับตำแหน่ง *embB* Met306Ile) ยา streptomycin (ตำแหน่ง *gid* Gly73Ala เทียบกับตำแหน่ง *rpsL* Lys43Arg) ยา moxifloxacin (ตำแหน่ง *gyrA* Ala90Val เทียบกับตำแหน่ง *gyrA* Asp94Gly) และยา rifabutin (ตำแหน่ง *rpoB* Asp435Phe/Tyr/Val เทียบกับตำแหน่ง *rpoB* Ser450Leu) นอกจากนี้ เมื่อวิเคราะห์ระดับ MIC ของยาต้านวัณโรคในกลุ่มเชื้อวัณโรคดื้อยาแบบ heteroresistance พบว่า เชื้อวัณโรคดื้อยาที่มีสัดส่วน mapped read ของตำแหน่งที่พบการกลายพันธุ์ในระดับต่ำ จะมีระดับ MIC ต่ำกว่าเชื้อที่มีสัดส่วน mapped read ที่สูงกว่า

สรุป ความชุกของเชื้อวัณโรคดื้อยาในประเทศไทยเกิดจากการที่มีอยู่ของกลุ่มการระบาดจำนวนหลายกลุ่ม การตัดลูกโซ่ของการระบาดเป็นสิ่งที่ต้องดำเนินการอย่างเร่งด่วนเพื่อช่วยลดความชุกของวัณโรคดื้อยา ซึ่งจะส่งผลทำให้การควบคุมวัณโรคในประเทศไทยมีประสิทธิภาพมากยิ่งขึ้น นอกจากนี้ วิธี WGS สามารถใช้ตรวจวินิจฉัยวัณโรคดื้อยา และการกลายพันธุ์ที่เกี่ยวข้องกับการดื้อยาที่ตรวจด้วยวิธีดังกล่าวมีความสัมพันธ์กับระดับ MIC ของยาต้านวัณโรค

Ditthawat Nonghanphithk. 2021. **Molecular Epidemiology and Genetic Analysis of Multidrug-/pre-extensively Drug/extensively Drug-resistant Tuberculosis in Thailand during 2014-2017**. Doctor of Philosophy Thesis in Medical Microbiology, Graduate School, Khon Kaen University

Thesis Advisor: Assoc. Prof. Dr. Kiatichai Faksri

ABSTRACT

Tuberculosis (TB), caused by *Mycobacterium tuberculosis* (*Mtb*), is a major public health problem in Thailand. Drug-resistant (DR) TB that are multi-drug resistant (MDR), pre-extensively drug-resistant (pre-XDR) and extensively drug-resistant (XDR) are complicating disease control. Whole-genome sequencing (WGS) of circulating *Mtb* can provide insights into the transmission of DR *Mtb* strains and inform clinical management. WGS also provides the highest genetic resolution for genotypic drug-susceptibility test (DST). However, the molecular epidemiology of DR-TB at nationwide scale in Thailand is unknown. In addition, there is limited information about the association between drug-resistance mutations and the phenotypic DST, especially for second-line drugs.

To determine the transmission scenario of DR-TB in Thailand, we analyzed WGS data of randomly selected 579 phenotypically DR *Mtb* strains isolated from TB patients treated at 230 hospitals across 71 provinces in Thailand during 2014–2017. The majority of *Mtb* were MDR-TB (n = 466; 80.5%), with 81 (14.0%) pre-XDR-TB and 32 (5.5%) XDR-TB. Most of *Mtb* isolates were from the East-Asian lineage (Lineage2; n = 482; 83.2%). There were 13 major transmission clades, with significantly associated with geography ($p < 0.001$). Using a ≤11 SNP cut-off between isolates, 281/579 (48.5%) formed 89 clonal clusters, including 205 MDR-TB, 46 pre-XDR-TB, 19 XDR-TB, and 11 poly DR-TB isolates based on genotypic drug-resistance. Members of most clusters had the same subset of drug resistance-associated mutations, supporting potential primary resistance in MDR-TB (n = 176/205; 85.9%), pre-XDR-TB (n = 29/46; 63.0%), and XDR-TB (n = 14/19; 73.7%). The western region of Thailand (OR 2.4, 95%CI: 1.5-3.9) and infection with lineage Lineage2.2.1 (OR 3.6, 95%CI: 2.4-5.3) increased the risk of DR-TB transmission.

To extend reliability and applicability of WGS analysis for DST of DR-TB, we selected *Mtb* causing DR-TB detected by phenotypic DST (n = 60) using agar proportion method including one poly DR-TB, 34 MDR-TB and 25 XDR-TB. DST results from WGS were compared with agar proportion method and minimum inhibitory concentration (MIC) tests. Agreement between WGS-based DST and MIC tests was high for all drugs except ethambutol (65%) and ethionamide (62%). Isolates harboring the -15 c/t *inhA* promoter mutation had a significantly lower MIC for isoniazid than did isolates with the *katG* Ser315Thr mutation ($p < 0.001$). Similar patterns were observed for ethambutol (*embB* Gly406Asp vs. *embB* Met306Ile), streptomycin (*gid* Gly73Ala vs. *rpsL* Lys43Arg), moxifloxacin (*gyrA* Ala90Val vs. *gyrA* Asp94Gly) and rifabutin (*rpoB* Asp435Phe/Tyr/Val vs. *rpoB* Ser450Leu). For genotypic heteroresistance, isolates with lower proportion of mapped read tended to has lower MIC of anti-TB drugs than those with higher proportion.

In conclusion, high prevalence of DR-TB in Thailand is due to multi-clonal epidemics. Cutting of the transmission chains involving DR *Mtb* strains is urgently needed to control TB. In addition, WGS can be used for determination of DR-TB and the association of drug-resistance mutations that associated with MIC levels.

I’m grateful to dedicate all goodness portion of this thesis to all tuberculosis patients

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Assoc. Prof. Dr. Kiatichai Faksri who introduce me to the research world, for his patience, motivation, and kindly support. His guidance and encouragement always support me throughout my challenging journey of graduate studies.

Besides my advisor, I would like to thank all collaborators and thesis committee; Assoc. Prof. Dr. Angkana Chaiprasert, Ms. Saijai Smithtikarn, Dr. Phalin Kamolwat, Dr. Petchawan Pungrassami, Prof. Virasakdi Chongsuvivatwong, Dr. Surakameth Mahasirimongkol, Prof. Wipa Reechaipichitkul, Asst. Prof. Dr. Chaniya Leepiyasakulchai, Asst. Prof. Dr. Wises Namwat and Dr. Kanin Salao: as well as, the oversea collaborators; Prof. Dr. Taane G. Clark, Asst. Prof. Dr. Jody E. Phelan and Assoc. Prof. Dr. David Blair for their kindly supports, constructive comments and suggestions that help me a lot to overcome all obstacles of my thesis work..

My sincere thanks also go to the National Research Council of Thailand and the Health System Research Institute for providing research fund to support my thesis work. I would like to thank Research and Diagnostic Center for Emerging Infectious Diseases and Faculty of Medicine, Khon Kaen University, for research assistant scholarship throughout my Ph.D. program. In addition, I would like to thank the National Tuberculosis Reference Laboratory, Division of Tuberculosis, Department of Disease Control of Thailand, that provide all access to the research facilities during my Ph.D. study.

In particular, I thank to Ms. Pimjai Ananta, Ms. Orawee Kaewprasert and Ms. Pratchakan Chaiyachat for their kindness, helpfulness and technical assistance for bioinformatics analysis and laboratory works. I would like to thank all of department and faculty members for their helps both direct and indirect ways and also to whom that may be connected to my research, but are not mention here.

Last but not least, I am grateful to thank my family, especially my parent for their love, devoting, understanding and supporting throughout my life.

Ditthawat Nonghanphithak

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Cont.)

LIST OF ABBREVIATIONS

A	adenine
ADP	adenosine diphosphate
Ala	alanine
AMK	amikacin
Arg	arginine
Asn	asparagine
Asp	aspartate
ATP	adenosine triphosphate
BDQ	bedaquiline
bp	base pair
C	cytosine
CC	critical concentration
CAP	capreomycin
CFZ	clofazamine
CI	confidence interval
CRyPTIC	Comprehensive Resistance Prediction for Tuberculosis: an International Consortium
CTAB	cetyl-trimethyl-ammonium bromide-sodium chloride
DCS	D-cycloserine
del	deletion
DLM	delamanid
DNA	deoxyribonucleic acid
DR	drug-resistant
DST	drug susceptibility test
EMB	ethambutol
ETO	ethionamide
FQ	fluoroquinolone
G	guanine
GAT	gatifloxacin

LIST OF ABBREVIATIONS (Cont.)

Gb	gigabyte
Gly	glycine
HIV	human immunodeficiency virus
IC50	half-maximal inhibitory concentration
Ile	isoleucine
INH	isoniazid
IPM-CLN	imipenem-cilastatin
IR-TB	isoniazid-resistant tuberculosis
KAN	kanamycin
Leu	leucine
LFX	levofloxacin
Lys	lysine
LZD	linezolid
MDR	multidrug-resistant
Met	methionine
MFX	moxifloxacin
MIC	minimum inhibitory concentration
MIRU-VNTR	mycobacterial interspersed repetitive unit–variable number of tandem repeats
MLST	multi locus sequencing typing
MPM	meropenem
Mtb	*Mycobacterium tuberculosis*
NGS	next generation sequencing
NTRL	National Tuberculosis Reference Laboratory
OADC	oleic albumin dextrose catalase
OFX	ofloxacin
OR	odds ratio
PAS	*para*-aminosalicylic acid

LIST OF ABBREVIATIONS (Cont.)

PE	proline-glutamate
PEE	proline-proline-glutamate
PCR	polymerase chain reaction
Phe	phenylalanine
Pro	Proline
PTO	prothionamide
PZA	pyrazinamide
ReSeqTB	The Relational Sequencing TB Data Platform
RFB	rifabutin
RFLP	restriction fragment length polymorphism
RFP	rifapentine
RIF	rifampicin
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RR-TB	rifampicin-resistant tuberculosis
SD	standard deviation
Ser	serine
SLD	second-line drug
SLID	second-line injectable drug
SNP	single nucleotide polymorphism
STR	streptomycin
T	thymine
TB	tuberculosis
Thr	threonine
TRD	terizidone
tRNA	transfer ribonucleic acid
Tyr	tyrosine
Val	valine

LIST OF ABBREVIATIONS (Cont.)

WGS	whole-genome sequencing
WHO	World Health Organization
XDR	extensively drug-resistant

CHAPTER I
INTRODUCTION

1. Background and rationale

Tuberculosis (TB) remains the major public health problem worldwide. The ending TB project has proposed by World Health Organization (WHO) in order to end TB by 2035. According to WHO ranking in 2019, Thailand is ranked in the top 14 of high burden countries with TB incidence (108,000 cases), TB/HIV coinfection (11,000 cases) and multidrug-resistant (MDR) TB prevalence (3,900 cases) [1]. Even though, there are small numbers of extensively drug-resistant (XDR) TB cases in Thailand [1], treatment of XDR-TB is more difficult, require long term hospitalization and expensive medications.

Effective TB control depends on several strategies such as rapid identification of new TB case, appropriate treatment of TB patient and prevention of TB transmission [2]. Next generation sequencing (NGS), whole genome sequencing (WGS), provides the highest resolution of genetic information of *Mycobacterium tuberculosis* (*Mtb*) and many applications for example clustering the transmission and prediction of drug-resistant (DR) TB [3]. Previously, the conventional molecular assays including IS*6110* restriction fragment length polymorphism (IS*6110* RFLP) [4], spoligotyping [5] and mycobacterial interspersed repetitive unit–variable number of tandem repeats (MIRU-VNTR) [6] are used for investigation of TB transmission. However, few parts of *Mtb* genome were analyzed that limit in genetic resolution and unable to differentiate among closely related strains [2].

In Thailand, there was a large cluster of MDR-TB transmission (148 cases) in Kanchanaburi during 2002-2010 [7]. However, only four isolates were further analyzed and subjected to perform WGS [8] and other unidentified clusters of DR-TB might be spreading to the neighborhood. Molecular epidemiology of DR-TB in nation-wide scale using WGS is limited among several countries [9-14] including Thailand [15-18]. Therefore, study of molecular epidemiology of DR-TB (MDR/pre-XDR/XDR-TB)

using a nation-wide sample could provide comprehensive and significant information for effective TB control in Thailand.

DR-TB is diagnosed by phenotypic drug susceptibility testing (DST), gold standard test, using agar proportional method to infer DST profiles [19]. However, this phenotypic DST is laborious and time consuming which might lead to inappropriate prescription for TB treatment. In addition, several mutations in *Mtb* were found to be associated with drug-resistance [20, 21]. On the other hands, the genotypic DSTs, GeneXpert MTB/RIF [22], GenoTypeMTBDRplus [23] and GenoTypeMTDRsl [24] are available and approved for rapid identification of DR-TB. However, these methods are limited to only key mutations associated with drug-resistance. WGS provides comprehensive mutations and can be used for guidance at the early stage of TB treatment before phenotypic DST profiles are reportable [25].

The minimal inhibitory concentration (MIC) test, alternative phenotypic DST, provides quantitation and level of drug-resistance for adjusting therapeutic regimens for TB treatment [26, 27]. Several studies reported the correlation between drug-resistance mutations with phenotypic DST [21, 28]. Also the association between drug-resistance mutations with certain ranges of MIC values have been reported [29-35]. However, such information is limited, especially for second-line dugs (SLDs) [21, 36].

Heteroresistance, defined as the mixture of susceptible and resistant of *Mtb* strains in particular [37], was reported to influence on quantitative DST [38, 39]. However, none of study performed direct comparisons between MIC levels and genotypic heteroresistance, based on variant frequencies, using WGS data.

Taken together, research in molecular epidemiology and genetic analysis of DR-TB (MDR/pre-XDR/XDR-TB) using nation-wide samples would provide more insight information for effective TB control, subsequently provide a great benefit to economy and society in Thailand which promote the End TB strategy by the year 2035 proposed by WHO.

2. Hypothesis

2.1 There should be DR-TB (MDR/pre-XDR/XDR-TB) transmission in Thailand. WGS analysis will reveal several clusters of DR-TB which can be associated with geographical regions.

2.2 The mutations associated with drug-resistance can be identified by WGS provide comprehensive genotypic DST profiles and good performance for prediction of DR-TB.

2.3 The mutations associated with drug-resistance can be identified by WGS might be able to correlate to certain MIC values from low to high which can indicate the level of drug-resistance.

3. Objectives

3.1 To study molecular epidemiology of DR-TB (MDR/pre-XDR/XDR-TB) by WGS using nationwide samples.

3.2 To determine the diagnostic performance and agreement between phenotypic and WGS-based genotypic DST.

3.3 To analyze the correlation between MIC values and WGS-based genotypic DST.

4. Scope and limitation of research

4.1 Data of DR-TB during 2014-2017 were retrieved from the National TB Reference Laboratory (NTRL), Division of TB, Ministry of Public Health and Drug Resistance Tuberculosis Fund Laboratory, Faculty of Medicine Siriraj Hospital, Mahidol University.

4.2 DR-TB cases (MDR-/pre-XDR/XDR-TB) were identified using phenotypic DST, agar and/or Lowenstein-Jensen (LJ) media proportion method. MDR- and pre-XDR-TB were randomly selected for 10-20% from the record. In addition, all cuturable XDR-TB cases were collected.

4.3 The molecular epidemiology of DR-TB in Thailand was investigated by WGS and applicable bioinformatics tools.

4.4 Diagnostic performance and agreement between genotypic and phenotypic DSTs were investigated by comparison between phenotypic DSTs (agar proportion and MIC test) with WGS-based genotypic DST.

4.5 Correlation between MIC level using Sensititre assay and drug-resistance mutations identified by WGS were analyzed.

5. Anticipated outcomes

This study will provide new insight knowledge regarding the molecular epidemiology of DR-TB at nationwide scale which will facilitate the establishment of DR-TB surveillance using WGS in Thailand and worldwide. The output from this project should be able to reveal multi-clusters of DR-TB and several factors associated with DR-TB transmission. DR-TB cases should be reduced when the public health organization implement an action into particular area where the hot spots were identified. Also, an information of diagnostic performance of WGS for prediction of phenotypic DST and correlation between MIC value with drug-resistance mutations will promote the development of rapid identification of DR-TB for better management of DR-TB patients. This study will contribute to the Ending TB project and promote an eradication of TB in Thailand.

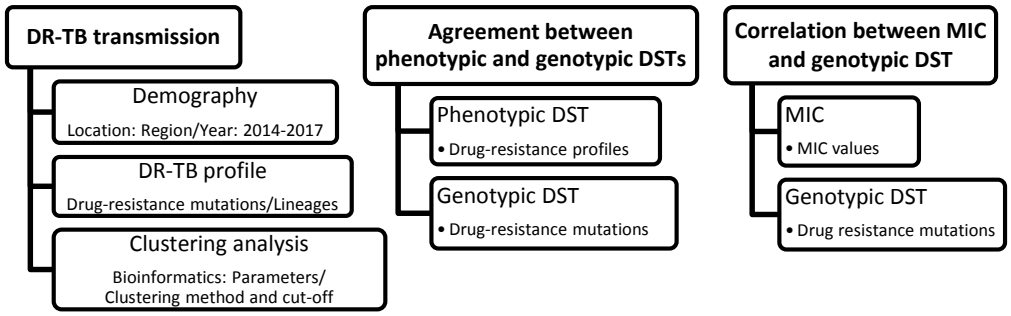

Figure 1 Conceptual framework of this study. There will be three objectives in this study. First, Molecular epidemiology of DR-TB in Thailand during 2014-2017. Several factors are contributed with DR-TB transmission including demographic characteristic of DR-TB patients, lineage and drug-resistance profiles of isolated *Mtb*. Selected parameters throughout bioinformatics pipeline affect stringency of WGS data used in downstream analysis. In addition, clustering method and cut-off for identification of DR-TB clusters influence on resolution in order to differentiate among closely related isolates within identified clusters. For the second and third objectives, drug-resistance profiles using phenotypic DST/MIC-test and identified drug-resistance mutations might have some influence on diagnostic performance and correlation of WGS-based DST for prediction of phenotypic DST and drug-resistance level.

CHAPTER II
LITERATURE REVIEWS

1. General characteristic of *Mtb*

Mtb was firstly discovered in 1882 by Robert Koch [40]. *Mtb* and other mycobacteria that cause disease in human and animals are categorized the *Mtb* complex *including M. africanum, M. bovis, M. canettii, M, caprae, M. microti, M. pinnipedii* and *M. tuberculosis* [41]. *Mtb* is an aerobe, acid-fastness, non-motility, non-encapsulated, non-spore forming and obligate intracellular bacteria. In general, *Mtb* is likely to grow in high oxygen tissues, the lungs. The lipid-rich cell wall, mycolic acids, of *Mtb* resistant to acidified organic solvents as same as *Nocardia* spp. The multiplication rate of *Mtb* is very slow, one single cell derives approximately every 15-20 hours. Treatment of *Mtb* infection is prolong due to ability of *Mtb* persistence inside the host. *Mtb* is categorized into six major lineages, lineage 1 (L1: Indo-Oceanic lineage), lineage 2 (L2: Beijing lineage), lineage 3 (L3: Central Asian lineage), lineage 4 (L4: Euro-American lineage), lineage 5 (L5: *M. africanum* West African type I) and lineage 6 (L6: *M. africanum* West African type II). In addition, lineage 7 was newly discovered lineage. Lineages were associated with certain geographical regions and diverse in virulence and ability to cause TB pathogenesis and transmission [42].

2. Genome of *Mtb*

The complete draft of *Mtb*, H37Rv (Figure 2), genome was successfully characterized in 1998 [43]. *Mtb* genome contains 4,411,529 bp and carrying approximately 4,000 genes with G and C rich content (65.6%). Compared to other bacteria, *Mtb* contains very large portion of coding region that producing about 250 different enzymes that involved in enzyme production in lipogenesis and lipolysis, and repetitive glycine-rich proteins (the proline-glutamate **(**PE) and proline-proline-glutamate (PEE) families) involved in antigenic variation.

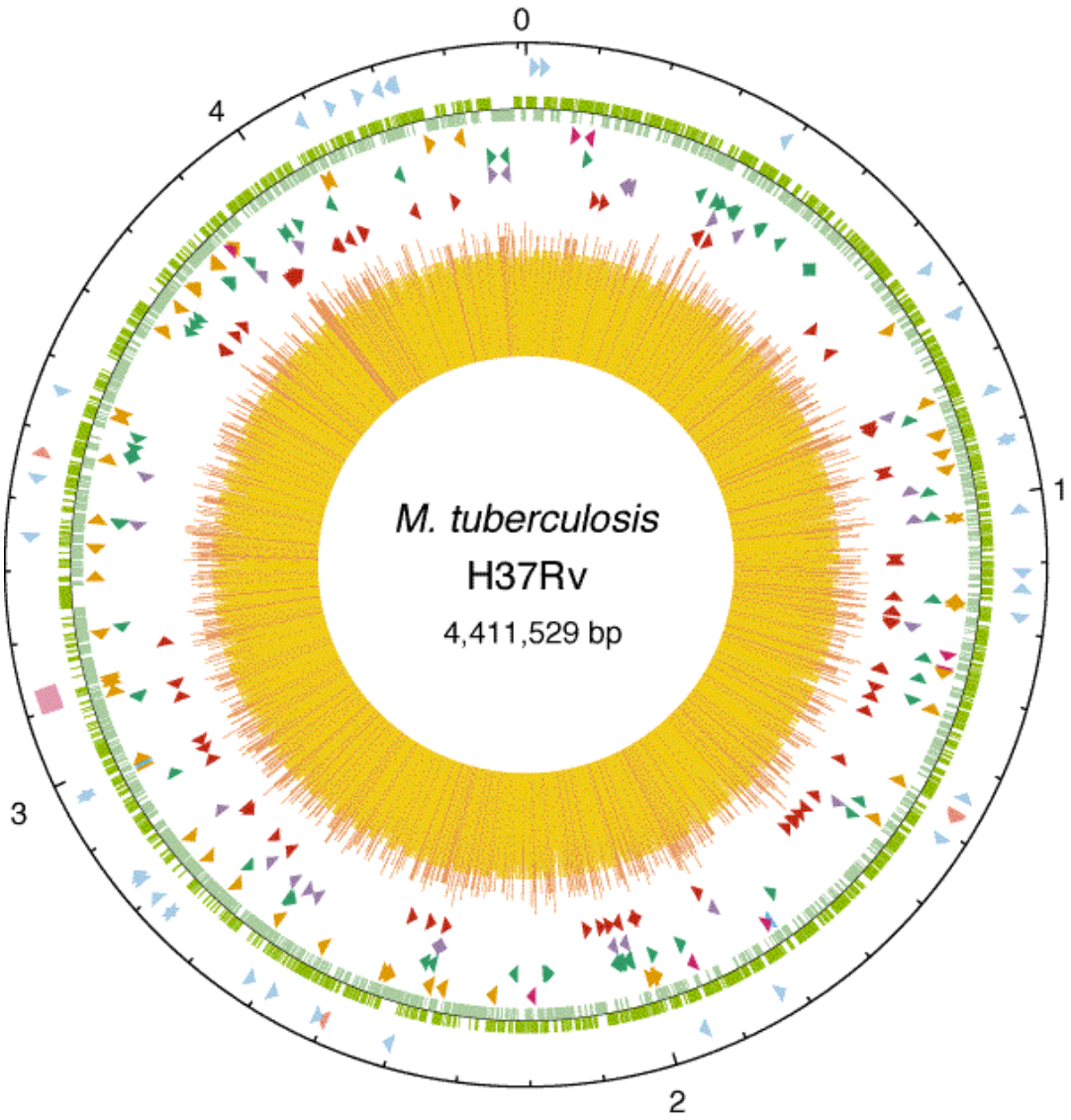

Figure 2 Genome map of *Mtb* H37Rv. The external circle line displays the scale of genome size (million bases) and 0 denoting the replication origin. The next line represents the direct repeat region in pink box, genes for stable type of RNA, tRNAs in blue and others in pink. The inner green lines show the coding rgions, clockwise in deep green and anti-clockwise in pale green. The next ring displays repetitive DNA including insertion sequences (yellow), 13E12 REP family (deep pink) and prophage (blue). The next green rig depicts the PPE families. The inner purple ring displays the PE families except polymorphic repetitive sequences. The next red ring displays the polymorphic GC-rich repetitive sequence. The center bar chart represents G and C contents including <65% (yellow) and >65% (red) [43].

3. Pathogenesis of TB

Mtb infection take places when the airborne droplet nuclei containing *Mtb* enters the lungs via inhalation, the droplets reach into alveolar space and ingested by alveolar macrophages. Effective innate immunity facilitates *Mtb* eradication when host infected with small number of *Mtb* cells. If host immunity fails to stop the infection, *Mtb* invades lung interstitial tissue to the lung parenchyma, by either transcytosis across alveolar epithelium or transmigration via infected alveolar macrophages. Adaptive immunity is involved when the infected antigen presenting cells, dendritic cells or monocytes, transport the *Mtb* to nearby lymph nodes and present *Mtb* antigen to T cell. After this step, the immune cells are recruited to lung parenchyma and form a granuloma to control the infection. The granuloma persistent for long term by effective immunity. Poor immunity promotes the multiplication *Mtb* and failure of granuloma formation. Thus, *Mtb* can disseminate to other tissues and organs including apex of the lung, regional lymph nodes, brain, bone and kidneys through lymphatic vessels or bloodstream which representing a symptomatic of active TB disease [44].

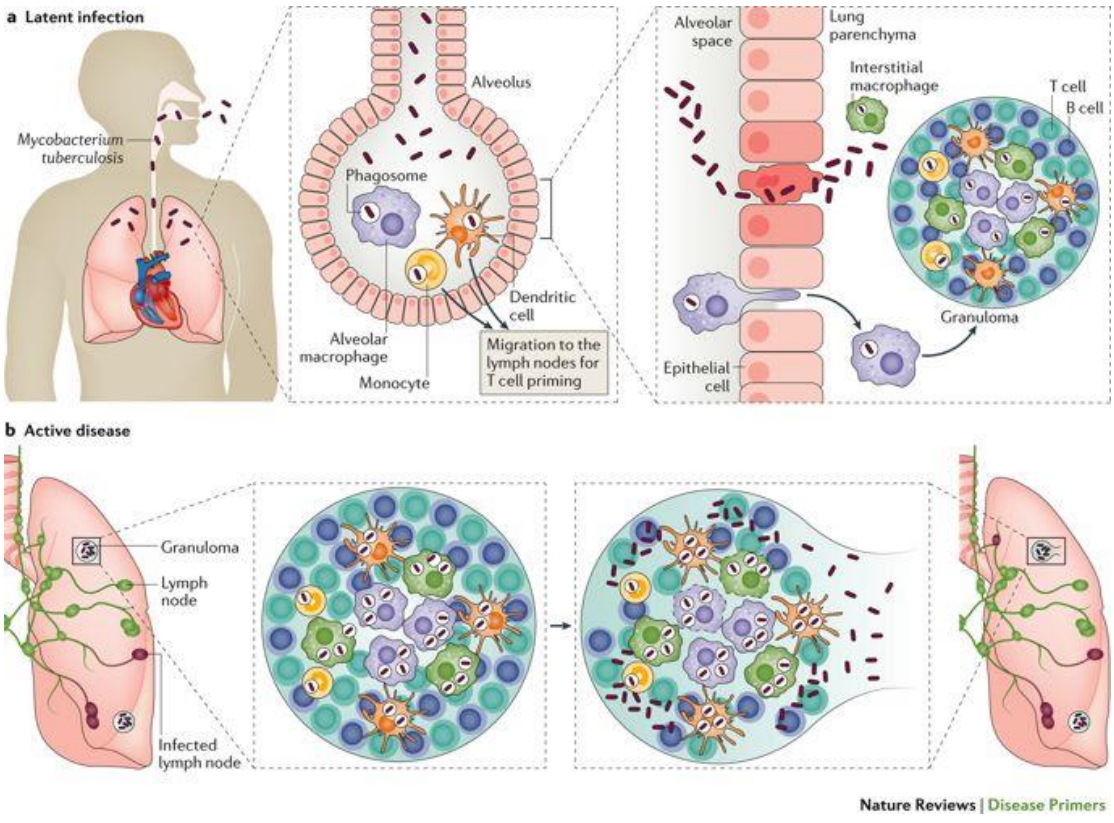

Figure 3 Pathogenesis of TB [44].

4. Pathogenesis of DR-TB

In general, DR-TB develops through fragmented treatment (A). Although, TB patient received an excellent treatment, acquired resistance can be developed by several factors, including efflux pumps (B), pharmacokinetic (C), different in drug penetration due to extensive immunopathology of the lung (D). Cough aerosol is the major cause of primary DR-TB transmission (E). Pathogen related factors and fitness cost also involve in DR-TB transmission (F). Compensatory mutations in drug-resistance *Mtb* lead to structural and physiological changes of bacterial pathways subsequently disrupt the protective immune responses and promote the progressive disease (G).

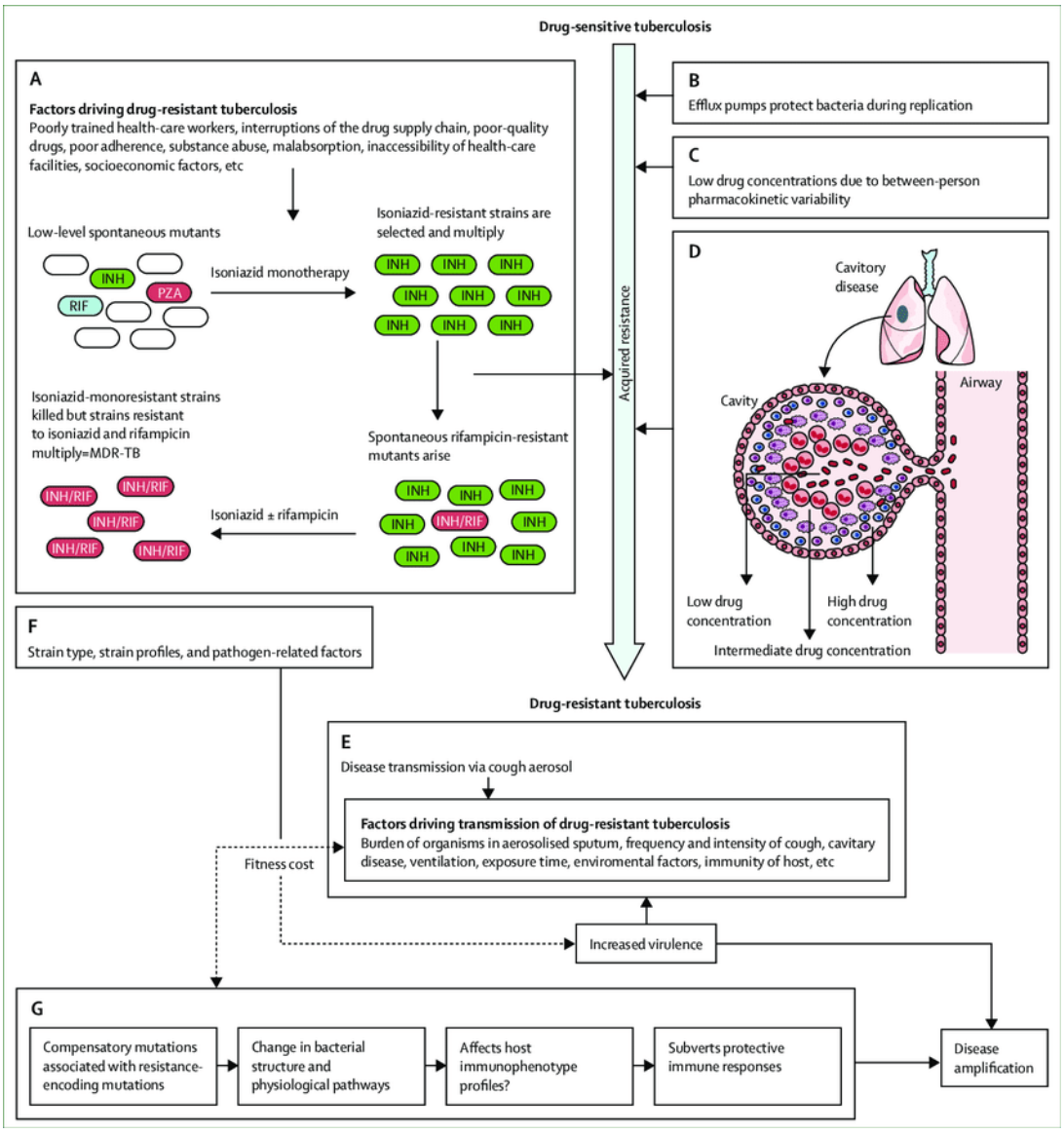

Figure 4 Pathogenesis of DR-TB [45].

5. Anti-TB drugs (Previous grouping definition)

Based on WHO classification, the anti-TB drugs are categorized into 5 categories according to its effectiveness and most common used to rarely used or unclear effectiveness. Anti-TB drugs and mechanisms of action were listed in Table 1 [46-48].

Table 1 Anti-TB drugs and mechanisms of action

Group	Anti-TB drug	Mechanism of action
Group 1: First-line oral drugs	Isoniazid (INH)	Inhibition of mycolic acid synthesis
	RIF, Rifabutin (RFB) and Rifapentine (RFP)	Inhibition of RNA synthesis
	Ethambutol (EMB)	Inhibition of arabinogalactan synthesis
	Pyrazinamide (PZA)	Inhibition of trans-translation, pantothenate, CoA synthesis and reduction of membrane energy
Group 2: Second line injectable drugs (SLIDs)	Steptomycin (STR), Kanamycin (KAN), Amikacin (AMK) and Capreomycin (CAP)	Inhibition of protein synthesis
Group 3: Fluoroquinolones (FQs)	Levofloxacin (LFX), Ofloxacin (OFX), Moxifloxacin (MFX) and Gatifloxacin (GAT)	Inhibition of DNA synthesis by interference of mycobacterial topoisomerase
Group 4: Oral bacteriostatic second-line anti-TB drugs	Ethionamide (ETO) and Prothionamide (PTO)	Inhibition of mycolic synthesis
	D-cycloserine (DCS) and Terizidone (TRD)	Inhibition of peptidoglycan synthesis
	Para-aminosalicylic acid (PAS)	Inhibition of folic acid synthesis and thymine nucleotide metabolism
Group 5: Anti-TB drugs with limited data on efficacy and/or long term safety in the treatment of DR-TB (new anti-TB drugs were included)	Linezolid (LZD)	Inhibition of protein synthesis
	Clofazamine (CFZ)	Production of reactive oxygen species and membrane destruction
	Amoxicillin/clavulanate and Meropenem (MPM)/clavulanate	Inhibition of protein synthesis
	Thioacetazone	Unknown
	Bedaquiline (BDQ)	Inhibition of ATP synthesis
	Delamanid (DLM) and Pretomanid	Inhibition of mycolic acid synthesis and production of reactive nitrogen species

6. Anti-TB drugs (Recent grouping definition)

WHO launches a new grouping of anti-TB drugs into four main groups (First-line, Group A, Group B and Group C) based on the current knowledge, effectiveness and safety of drug used in TB treatment. SILDs are no longer recommended in the shorter regimen. However, SLIDs are deprioritized in Group C which only two of SLIDs are retained, AMK and STR. Treatment outcomes of FQs resistance is poor. Currently, only LFX and MFX are recommended in shorter and longer regimens. BDQ and LZD become as a part of the current regimen and have widely been standardized. Resistance to BDQ and LZD are rare. The anti-TB drugs listed in Group A-C are suggested for longer regimens of DR-TB treatment. The drugs listed in each group can be changed in the future. In conclusion, this newly groping of anti-TB drugs would engage the reader to understand the current definition of pre-XDR/XDR-TB and guideline for treatment of DR-TB. New grouping of anti-TB drugs is classified in Table 2.

Table 2 Current grouping of anti-TB drugs with spectrum of phenotypic and genotypic DSTs endorsed (X) or in endorsement plan (O) by WHO

Group	Anti-TB drug	Phenotypic DST	MIC based (Microtitre plates)	Cartridge based (Xpert)	Line probe assay	WGS
Frist line	INH	X	O	O	X	O
	RIF	X	O	X	X	O
	EMB	X	O			O
	PZA	X	O		O	O
Group A	LFX	X	O	O	X	O
	MFX	X	O	O	X	O
	BDQ	X	O			O
	LZD	X	O			O
Group B	CFZ	X	O			O
	DCS					
Group C	DLM	X	O			O
	MPM					
	Imipenem-cilastatin (IPM-CLN)					
	AMK	X	O	O	X	O
	STR	X	O			O
	ETO	X				O
	PTO	X				O
	PAS					

7. Current definition of pre-XDR-TB and XDR-TB

According the meeting in October 2020, WHO reports new definition of pre-XDR/XDR-TB are defined in Table 3.

Table 3 Current definition of pre-XDR/XDR-TB

Pattern of DR-TB	Definition
Pre-XDR-TB	MDR-TB or RIF-resistant TB (RR-TB) and additional resistance to any FQ (LFX or MFX)
XDR-TB	MDR-TB or RR-TB and additional resistance to any FQ and at least one drug from Group A

8. Treatment of DR-TB

The most recent guidance for DR-TB treatment was updated and published in 2020 by WHO especially for treatment of MDR-TB/RR-TB [49]. The current regimens were recommended for DR-TB treatment are showed in Table 4.

Table 4 The current recommendation for DR-TB treatment by WHO

Type of DR-TB	Treatment	Remark
RIF-susceptible and INH-resistant TB (IR-TB)	RIF, EMB, PZA and LVX (6 months)	STR or other SLIDs are not recommend
MDR-TB or RR-TB (Shorter regimens)	Oral BDQ-containing regimen (9-12 months)	Patient have not been received the treatment with SLDs >1 month and not resistance to FQs.
Longer regimens for MDR-TB or RR-TB were classified into three groups. **Group A: (3 drugs are included)**: LFX or MFX, BDQ and LZD **Group B: (1 or 2 drugs are included)**: CFZ and DCS or TRD **Group C: (added to complete the regimen, group A and B cannot be used)**: EMB, DLM, PZA, IPM-CLN or MPM, AMK or STR, ETO or PTO, and PAS		
MDR-TB or RR-TB (Longer regimens) Treatment duration = 18-20 months Culture conversion take 15-17 months	3 drugs (group A) and 1 drug (group B) or 1 or 2 drugs (group A) and 1 drug (group B), then, another	KAN and CAP are not added. LFX or MFX should be added. BDQ should be added for ≥18 years old (may be added for 6-17 years old). LZD should be added. CFZ and DCS or TRD may be added. EMB may be added. DLM may be added for ≥3 years old. PZA may be added. IPM-CLN or MPM may be added. AMK (or STR) may be added for ≥18

Table 4 The current recommendation for DR-TB treatment by WHO (Cont.)

Type of DR-TB	Treatment	Remark
Intensive phase for 6-7 months was applied when AMK (or STR) is used	drugs in group C are included	years old with confirmed susceptible result. ETO or PTO may be added if BDQ, LZD, CFZ or DLM are not used. PAS may be added if BDQ, LZD, CFZ or DLM are not used. Clavulanic acid should not be added.

9. Estimated DR-TB incidence

Based on the global TB report 2018, estimated MDR/RR-TB cases were 3.5% of incidence cases and 18% from treated cases worldwide. Globally, there were estimated 558,000 incident MDR/RR-TB cases and 82% of them have MDR-TB. The top three prevalence of MDR/RR-TB cases were found in India, China and Russian (Figure 5). In addition, estimated XDR-TB cases were 8.5% of MDR-TB cases. Globally TB and MDR/RR-TB cases are showed in Table 5. Thailand is the one of the high MDR-TB burden. During the past five years, TB and MDR/RR-TB cases in Thailand have dramatically increased (Table 6).

Table 5 Globally TB cases report and estimates MDR/RR-TB in 2017 [50]

	TB case	MDR/RR-TB	Laboratory-confirmed	
WHO Region	(n)	(n)	MDR/RR-TB (n)	XDR-TB (n)
African	1,323,450	39,000	26,845	867
Americas	243,064	8,100	4,084	121
Eastern Mediterranean	536,185	21,000	4,969	168
European	264,563	76,000	48,299	6,758
South-east Asia	2,965,311	99,000	51,788	2,755
Western Pacific	1,375550	87,000	24,699	131
Global	**6,708,123**	**330,100**	**160,684**	**10,800**

Table 6 TB cases and MDR/RR-TB in Thailand during 2013-2017 [50-54]

	TB case	MDR/RR-TB	Laboratory-confirmed	
Year	(n)	estimation (n)	MDR/RR-TB (n)	XDR-TB (n)
2013	66,415	1,880	230	NA
2014	71,618	2,200	506	NA
2015	66,179	2,500	466	5
2016	72,014	2,700	955	13
2017	82,008	2,700	1,339	7

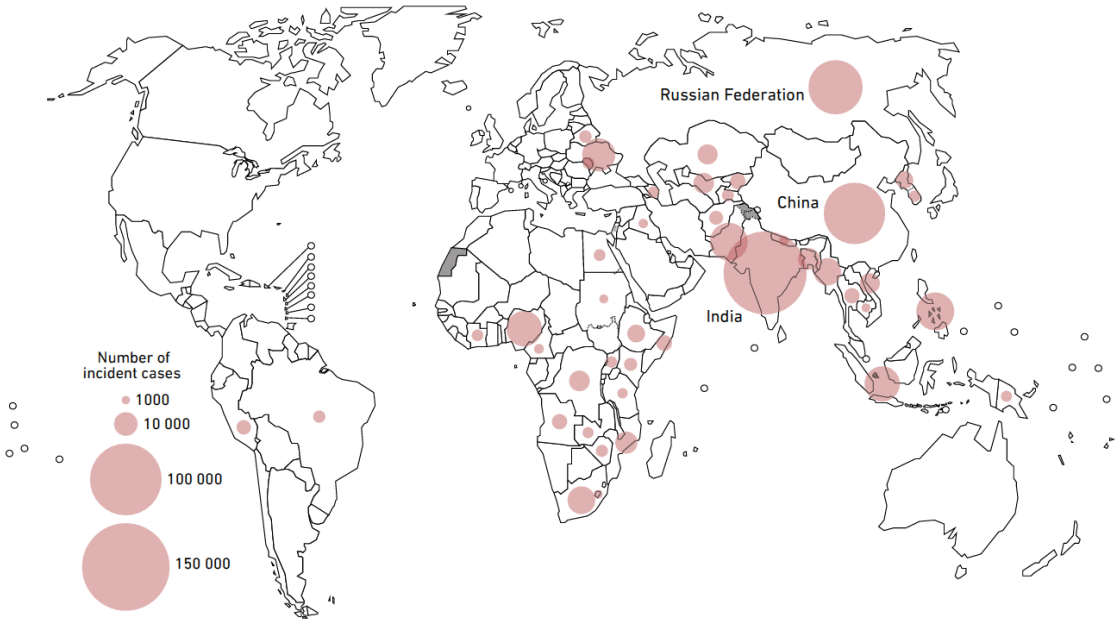

Figure 5 Estimation of MDR/RR-TB incidence in 2017 for countries with at least 1,000 incidence [50].

10. Mechanisms of DR-TB

DR-TB can be caused by intrinsic and acquired resistance (Figure 6) [55].

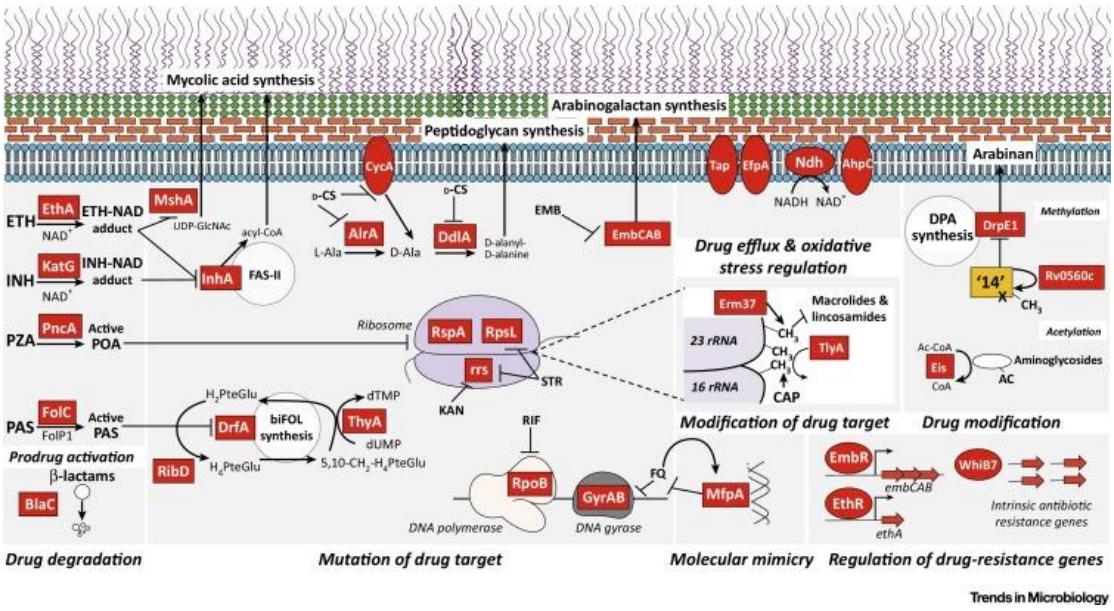

Figure 6 Overview of intrinsic and acquired drug-resistance mechanisms of *Mtb* [56].

10.1 Intrinsic resistance

10.1.1 Cell wall impermeability

The cell wall of mycobacteria composts of three main components: mycolic acid, arabinogalactan and peptidoglycan. These barrier protects the bacteria from environmental stress and certain antibiotics. Impairment of the call wall promotes the sensitivity to many anti-TB drugs. Several enzymes involved in *Mtb* cell wall, for example GlmU, MurX and Alr, and these proteins can be used as targeted anti-TB drugs [55].

10.1.2 Dormancy and latency

Dormancy state defined as non-replication with low-absent metabolic activity of mycobacteria. The dormant cells represent asymptomatic infection without active disease, or persistence state, long term survival of mycobacteria in the present of antibiotics although they are normally susceptible to the drugs. The mycobacteria decrease their metabolisms including respiration and transcription rates, energy metabolism, synthesis of lipid and cell division, that affect to the production of antibiotic targeted proteins and promote *Mtb* tolerance to antibiotic agents [55]. The Ltds enzyme and ClpB/DnaK chaperones play role during this state. An information about the factors associated with dormancy would provide the candidate of drug-targets in order to inhibit the dormancy of *Mtb*.

10.1.3 Porin channels

The lipid rich membrane of mycobacteria is permeable for certain agents except the hydrophilic agents. In general, the hydrophobic agents use the channel proteins in order to pass through the outer membrane. Porin-like protein, MspA, plays role in this step and can be found in *Mtb* and *M. smegmatis.* Heterologous expression of MSpA in *M. smegmatis* increases susceptibility to anti-TB drugs such as INH, STR and EMB. Better understanding of the influx system would facilitate an invention of novel agents that can cross the cell wall of *Mtb* [55].

10.1.4 Efflux pumps

Not only the influx system, the efflux system also contributes to low permeability of the mycobacterial cell wall by transporting drug molecules out of the cell. This efflux system participates in the intrinsic resistance to anti-TB drugs [55]. Over expression of the efflux pumps also promote drug-resistance, especially for those

Mtb isolates without mutations associated with drug-resistance. Regulation of the efflux system by prevention an overexpression of efflux pump using the inhibitors provide shorten TB treatment.

10.1.5 Modification of antibiotic-targets

Modification of antibiotic-targets also consider as intrinsic resistance [55]. Alteration of specific drug binding site of 23S rRNA by methylation, product of *erm*, prevents the macrolide-binding site. The expression of *erm* can cause resistance to clarithromycin, erythromycin and ketolide in non-tuberculous mycobacterium species. In addition, the pentapeptide, produced by *mfpA*, act as DNA mimicry which bind to DNA gyrase and lead to prevention of the drug-binding site for several FQs. Loss of methylated rRNA by deactivation of methyltransferase cause resistance to STR, CAP and viomycin. In addition, the RNA polymerase binding protein A can interact with RNA polymerase and inhibit the binding of RIF.

10.1.6 Degradation and modification of antibiotics

Degradation of antibiotics involve several enzymes that cleave several classes of drugs that consist of β-lactams, aminoglycosides and macrolides [55]. The β-lactams inhibit the transpeptidase in the final step of peptidoglycan cross-link. The production of β-lactamase (class A), encoded by the *blaC* in mycobacteria, can cause resistance to β-lactam. This enzyme hydrolyzes the β-lactam ring of the drug. In addition, the β-lactamase has broad hydrolysis spectrum to the carbapenems. Modification of antibiotics by adding the chemical groups to specific sites using the enzyme leads to prevention of targeted drug binding site of the drug. For example, the N-cetyltransferase acetylates the aminoglycosides bearing 2′ amino group and leads to aminoglycosides resistance. In addition, ADP-ribosyltransferase can transfer the ADP-ribose unit to RIF's hydroxyl residue leading to RIF resistance in *M. smegmatis*.

10.1.7 Activation of transcription factors

Transcriptional regulator, whiB7, participates in intrinsic resistance by activation of drug-resistance genes [55]. The expression of whiB7 gene is induced by stress conditions and sub-inhibitory concentration of anti-TB drugs. This regulator controls the expression of *eis* that contributes in intracellular survival of *Mtb*. Other regulators, for example dosR also mediate the survival of *Mtb* in granulomatous lesions.

10.2 Acquired resistance

Acquired DR in *Mtb* mainly cause by spontaneous mutations, including point mutation, small/large insertion or deletion of nucleotide, modification of the drug target, inhibition of enzyme involving pro-drugs activation or increase the targeted drug [46]. *Mtb* carrying such mutations are selected, increase their population and replace the drug-susceptible population [57] during improper regimens, patient non-adherence, differences in pharmacodynamics and pharmacogenomics, drug quality and kinetic of drug administration [58]. Also, compensatory mechanisms could facilitate the transmission of drug-resistance isolates [59]. This acquired secondary mutation restore the fitness cost of *Mtb* harboring drug-resistance mutations [60] and modify its phenotype [61]. The major genes of *Mtb* that involved in DR-TB were illustrated in Figure 7 and listed in Table 7.

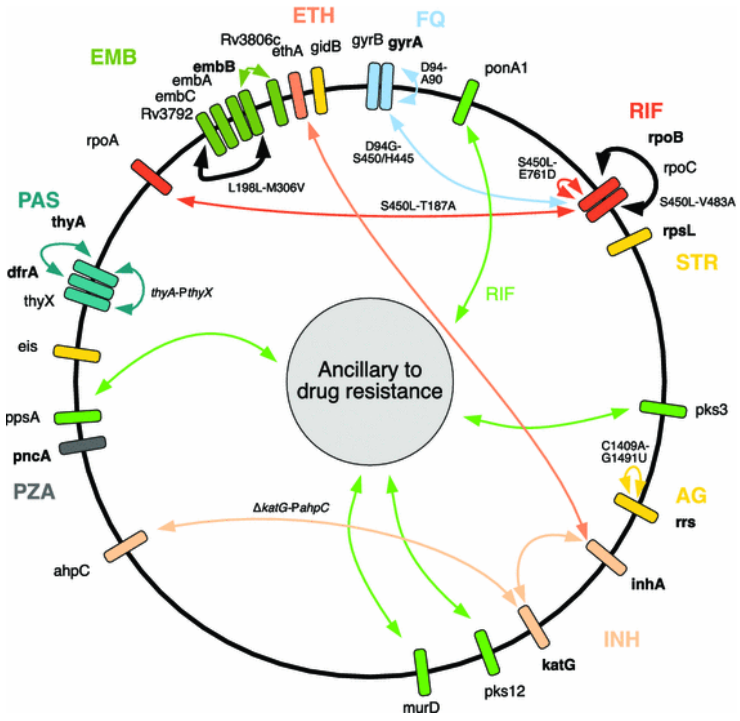

Figure 7 Key drug-resistance conferring genes located in *Mtb* genome. Several genes that directly involved in drug-resistance are showed in bold. Lines indicate the putative interactions between genes involving the same or different drugs and also denote genes involving board/indirect, ancillary, mechanisms of drug-resistance. Bold lines denote putative compensatory mechanisms for example *rpoB* to *rpoC* and *embB* to *Rv3972* [62].

Table 7 Molecular mechanisms of DR-TB

Drug	**Gene**	**Function**
INH	*katG*	Catalase/peroxidate
	inhA	Enoyl reductase
RIF, RFB and RFP	*rpoB*	RNA polymerase (β-subunit)
EMB	*embA,* *embB*	Arabinosyl transferase
PZA	*pncA*	Pyrazinamidase/Nicotinamidase
	rpsA	Ribosomal protein (S1)
	panD	Aspartate decarboxylase
STR	*rpsL*	Ribosomal protein (S12)
	gidB	Guanosine methyltransferase
KAN	*rrs*	rRNA (16S)
	eis	Aminoglycoside acetyltransferase
AMK	*rrs*	rRNA (16S)
	eis	Aminoglycoside acetyltransferase
CAP	*rrs*	rRNA (16S)
	tlyA	Methyltransferase
OFX, LFX, MFX and GAT	*gryA*	DNA gyrase (subunit A)
	gryB	DNA gyrase (subunit B)
ETO and PTO	*ethA*	Flavin monooxygenase
	inhA	Enoyl reductase
DCS and TRD	*alr*	Alanine racemase
	ddl	D-alanine-D-alanine ligase
	cycA	D-serine proton symporter
PAS	*thyA*	Thymidylate synthase A
	folC	Dihydrofolate synthase
	dfrA	Dihydrofolate reductase
	ribD	Enzyme involved in riboflavin biosynthesis
LZD	*rplC*	Ribosomal proteins L3
	rrl	rRNA (23S)
CFZ	Rv 0678	Transcription regulator for MmpS5-MmpL5 efflux pump
Amoxicillin/clavulanate and MPM/clavulanate	Unknown	Unknown
Thioacetazone	Unknown	Unknown
BDQ	*atpE*	ATP synthase (subunit c)
	Rv 0678	Transcription regulator for MmpS5-MmpL5 efflux pump
DLM and Pretomanid	*ddn*	Deazaflavin-dependent nitroreductase
	fdg1	F420-dependent glucose-6-phosphate dehydrogenase

11. Concept of WGS applications in TB research

WGS is an excellent tool that can be used to promote an effective TB control. This tool provides an accurate and rapid detection of DR-TB, which helping clinician for appropriate prescription of DR-TB. In addition, it used to investigate TB transmission by providing the highest resolution of genetic data and discriminatory power to differentiate among closely related isolates when compared to traditional assays. The applications of WGS in TB research were illustrated in Figure 8 and 9.

Targeted NGS and WGS are mainly applied in TB research. Targeted NGS is used for rapid detection of specific sequence of interested loci, for example specific genes associated with drug-resistance. On the other hand, WGS provides the nearly complete genome of *Mtb*, offering higher depth (>20X) and coverage (>98%), which essential for combination of epidemiological and genomic information in order to track the transmission. Both methods rely upon the same basic workflows and can be run in the same instrument, however, the processing steps might different according to the application. The applications of WGS are consisting of TB surveillance (a), TB typing (b), genotypic DST profile (c) and outbreak investigation (d) (Figure 8).

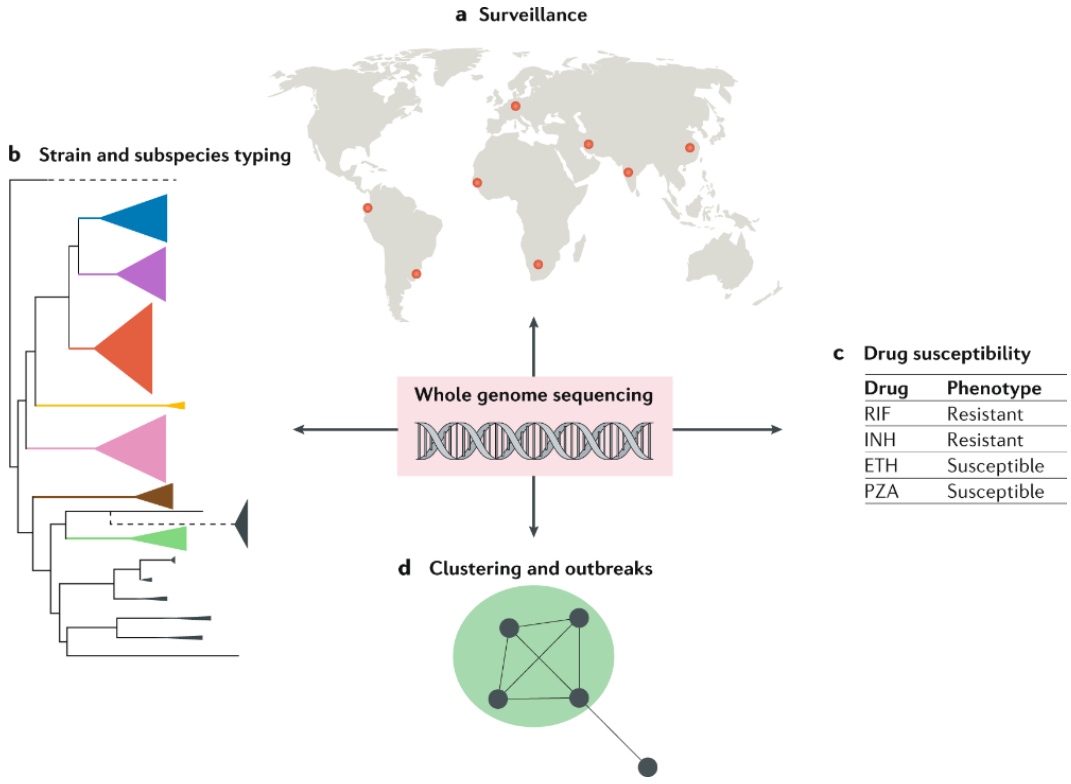

Figure 8 Applications of WGS for TB project in public health [3].

In addition, WGS is also used for four main reasons to characterize *Mtb* isolates (Figure 9) [63] including transmission chain analysis weather individuals infected with closely related isolates (a), diversity of mycobacterial within host that caused by either microevolution or mixed infection (b), microevolution over duration of treatment due to either relapse or re-infection with new *Mtb* isolates (c) and identification of resistance mechanisms weather individuals infected with drug-resistance strain (primary resistance) or development of drug-resistance mutations after having TB treatment (acquired resistance) (d).

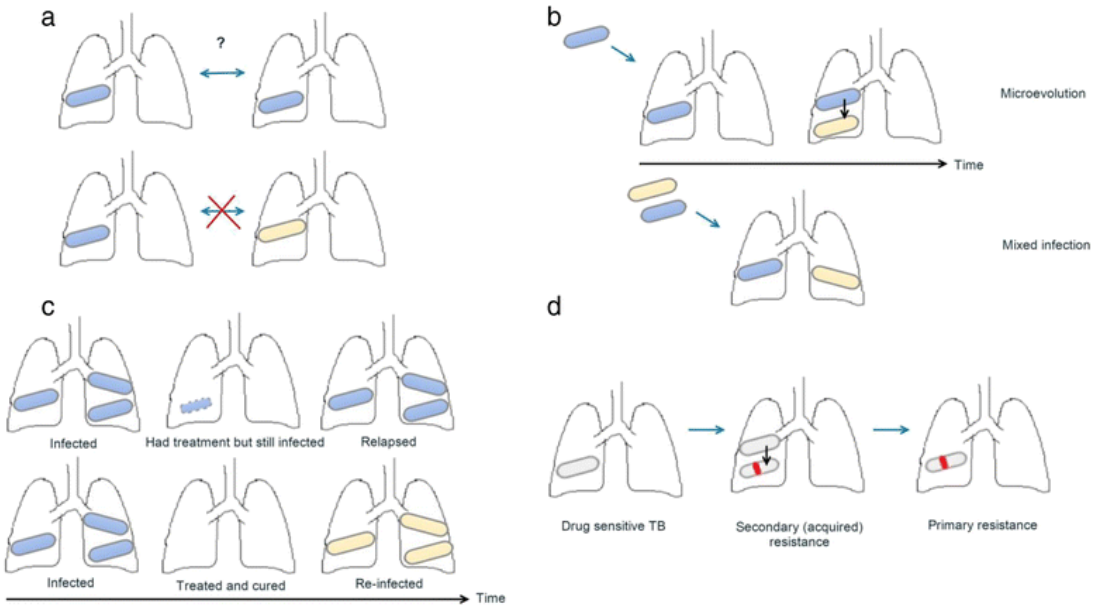

Figure 9 Four mains implication of WGS to characterize *Mtb* isolates [63].

12. Workflow of WGS analysis

In general, WGS workflow consisting of four main steps: (1) DNA extraction and purification, (2) library preparation of the extraction, (3) sequencing of the sample and (4) bioinformatics analysis of sequenced data. WGS workflow for TB research is illustrated in Figure 10.

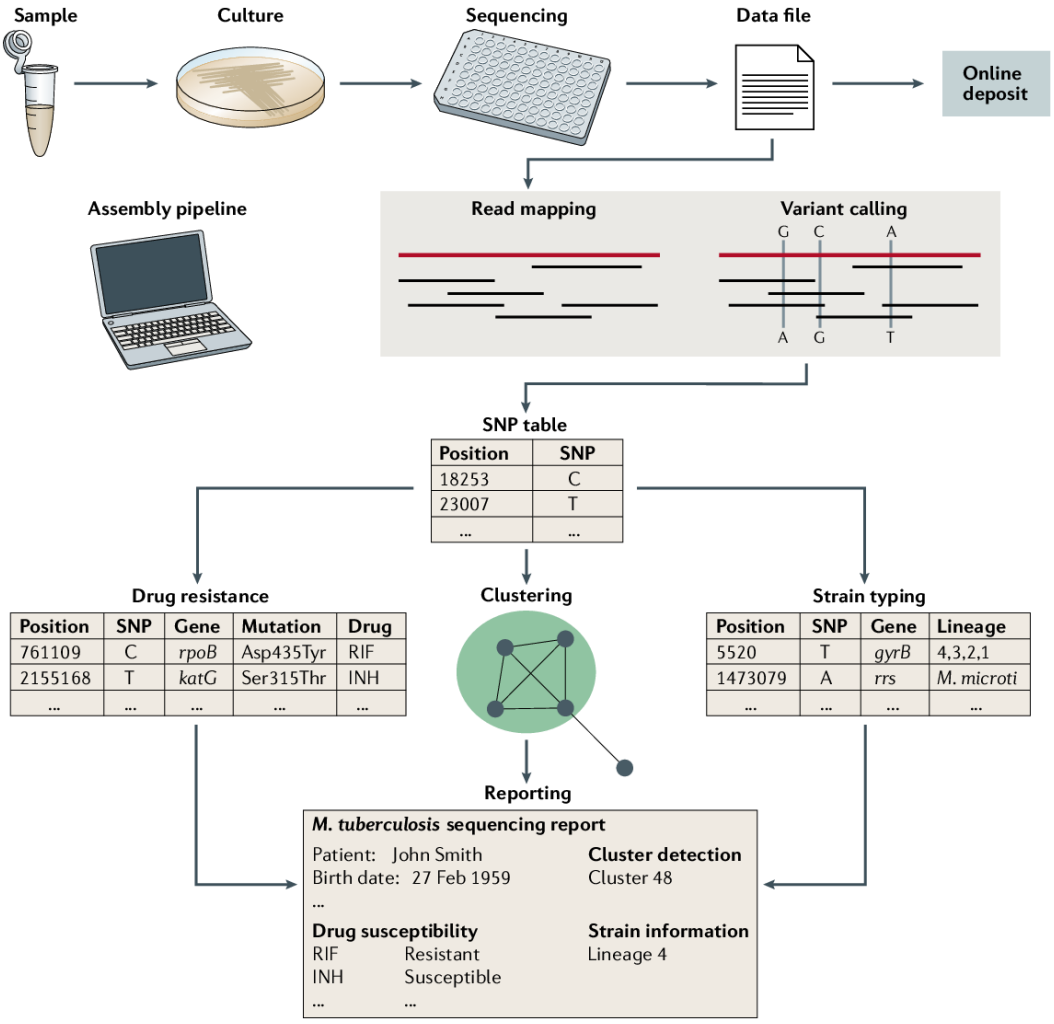

Figure 10 WGS workflow for TB research. This workflow begins with *Mtb* culture, follow by extraction of genomic DNA and sequencing. The sequenced data can be analyzed through *in silico* or in-house bioinformatics. First, raw fastq file is evaluated for QC check and trimmed to remove the adaptor. Then, the sequenced reads are aligned and mapped to reference genome (*Mtb* H37Rv) and variants are called. This data can be used for further analysis including prediction of drug-resistance, strain typing or clustering.

12.1 DNA extraction

DNA extraction is the crucial step to obtain a good quality of sequenced data. The extracted sample has to pass the minimum requirements for both of purity and quantity. The specific requirements of the extraction are vary based on selected library preparation kit and sequencing application (Table 8).

Table 8 Library preparation kits and sequencing applications [36]

Library preparation kit	NGS application	System compatibility	Input quantity (ng)
Nextera XT	WGS/Targeted	All Illumina	1
Nextera DNA Flex	WGS/Targeted	All Illumina	1-500
AmpliSeq	Targeted	All Illumina	1-100
Ion Xpress Plus Fragment	WGS/Targeted	PGM and S5 Ion Torrent	100
MuSeek	WGS/Targeted	PGM and Proton	100

12.2 Library preparation

Library-construction process generates a collection of specific fragmented DNA attached with adaptors. This step including fragmentation of DNA, end-repair of the fragments, phosphorylation, adenylation, ligation to adaptors, PCR amplification and library purification [64]. Library preparation kits are showed in Table 8.

12.3 Platforms of NGS

Several NGS platforms are available for TB research. However, there are two main platforms that widely used as following. Characteristic of NGS instruments are showed in Table 9 and Table 10.

12.3.1 Illumina: Sequencing by synthesis

This platform involves DNA polymerase sequencing of amplified fragmented library with reversible terminator nucleotides. Several sources of errors might due to phasing (when loss of fragments or increased additional base due to incomplete deblocking or blocking respectively) and fluorescent noise (incomplete cleavage of fluorescent label from previous cycles). The overall error rate is 0.5% (1 in 200 bases) [65].

12.3.2 Iron Torrent: pH-based sequencing

This platform involves detection of pH changing caused by hydrogen ions release during DNA synthesis. Compared to Illumina, this method provides a

shorter run and longer reads. However, the reads are single-stranded which provide less data. The error rate is higher due to the signal generated from the library fragments with repeat of the same nucleotide, homopolymers (in insertion or deletion regions). Read coverage in GC- or AT-rich regions is poor. The overall error rate is 1% (1 in 100 bases) [65].

Table 9 Characteristics of commonly used NGS instruments

Instrument	**Chemistry (Sequencing by Synthesis)**	**Detection**	**Data output (Gb)**	**Maximum read length(bp)**	**Sequencing time (hours)**
iSeq	Bridge PCR	Fluorescence	0.3-1.2	2 x 150	9-17.5
MiniSeq	Bridge PCR	Fluorescence	1.7-7.5	2 x 150	4-24
MiSeq	Bridge PCR	Fluorescence	0.3-15	2 x 300	4-55
NextSeq	Bridge PCR	Fluorescence	10-120	2 x 150	12-30
HiSeq (2500)	Bridge PCR	Fluorescence	10-1000	2 x 150	<36
Nova Seq (5000/6000)	Bridge PCR	Fluorescence	2000-6000	2 x 150	16-44
PGM	Emulsion PCR	Proton/ Semiconductor	0.08-2	400	3-10
S5	Emulsion PCR	Proton/ Semiconductor	0.6-15	400	up to 19
Proton	Emulsion PCR	Proton/ Semiconductor	10-15	200	4-24

Table 10 Strengths and weaknesses of commonly used NGS instruments

Instrument	**Manufacturer**	**Strengths**	**Weaknesses**
iSeq	Illumina	Low initial cost, short running time	Lower read length, low throughput
MiniSeq	Illumina	Low initial cost, short running time	Lower read length, low throughput
MiSeq	Illumina	Higher read length	Long running time
NextSeq	Illumina	Throughput	Long running time
HiSeq (2500)	Illumina	Throughput, accuracy of read	Long running time, high initial cost
Nova Seq (5000/6000)	Illumina	Throughput, accuracy of read	Long run time, high initial cost
PGM	Thermo fisher	Short running time, higher read length	Low throughput, homopolymers
S5	Thermo fisher	Higher read length	Homopolymers
Proton	Thermo fisher	Short running time, higher read length	Homopolymers

12.4 Bioinformatics and data analysis

Data analysis of sequenced data requires computational resources and bioinformatics skills in order to manage sophisticated and high volume of data. Bioinformatics analysis of *Mtb* requires command-line skill for customizing in-house scripts and pipelines for variants detection, prediction of drug-resistance and clustering the transmission. MTBseq [66] and UVP-ReSeqTB [67] pipelines are available for sequence from Illumina. In general, some of sequenced data (10%) are lost due to Illumina is unable to capture the long regions in highly variable certain regions such as insertion and PE-PPE genes [68, 69]. In addition, some sequenced data can be lost due to parameters (stringency) such as base quality, base alignment quality, mapping quality, read depth, coverage and frequency of the variants.

Transmission was analyzed using phylogeny which created from multiple alignment sequences to illustrated the evolution and relationship among population. These aligned sequences contain high confidence and high quality variants without indels and drug-resistance mutations due to selective pressure [70]. Transmission can be clustered by SNP cut-off based clustering method. The cut-off of 5 or 12 SNP are often used to identify the recent transmission [71]. In addition, the multi locus sequencing typing (MLST) method define the cluster using the defined core set of genes (core genome MLST [72] and whole genome MLST [73]). These methods provide high resolution of genetic typing and show similar clustering results [74].

Prediction of drug-resistance profiles were identified by comparison between identified variants with database. The Relational Sequencing TB Data Platform (ReSeqTB) was found in order to accumulate the large data set and the association between phenotypic and genotypic DSTs worldwide [75, 76]. Another project, Comprehensive Resistance Prediction for TB: an International Consortium (CRyPTIC), propose to identify the correlation between quantitative DST (MIC value) and drug-resistance mutations [77]. Several web-based tools and software are available for prediction of drug-resistance by upload the sequenced data including KvarQ, PhyresSE, TBProfiler, PhyTB, CASTB and TGS-TB [78].

13. Molecular epidemiology and genetic analysis of DR-TB using WGS

In Western regions, WGS is widely used to identify DR-TB transmission. This method is used to characterize genetic characteristics of transmitted isolates, tacking their descent and prediction of drug-resistance (Table 11).

Globally, WGS analysis of DR-TB using nationwide sample is limited. In European countries, there were few studies analyzed WGS of MDR-TB. In Portugal, 50% of the MDR-TB isolates (sample size = 77 cases) formed clusters using MIRU-VNTR patterns and 15% of them showed epidemiological link [14]. In low-incidence countries, 21.7% and 25% of MDR-TB formed clusters in Poland (sample size = 46) [11] and Switzerland (sample size = 49) [9] respectively using IS*6110* RFLP, spoligotyping and MIRU-VNTR typing. In United Kingdom 15% MDR-TB (sample size = 189) formed clusters using MIRU-VNTR typing [10]. In high incidence setting, 31.2% of MDR-TB cases in China were grouped into 10 clusters using spoligotyping [12]. In Saudi Arabia, 48 isolates (67.6%) formed 14 clusters of MDR-TB (sample size = 71) using WGS, however, this study unable to confirm epidemiological link [13].

Research in DR-TB transmission using WGS and nationwide sample is limited due to limited in discriminatory power of typing methods, convenient sample size and few studies mentioned fully spectrum of DR-TB (MDR/pre-XDR/XDR-TB). WGS was recently applied for TB research in Thailand (Table 12). However, WGS-based epidemiology of nationwide MDR/pre-XDR/XDR-TB sample in Thailand have never been studied yet. Therefore, this study would facilitate DR-TB surveillance and promote the usefulness of WGS for prediction of DR-TB in our country.

Table 11 Review articles in molecular epidemiology of DR-TB using WGS

Year	Finding	Region	Ref
2009	One susceptible isolate and another MDR isolate had the same genetic pattern using IS*6110* RFLP and MIRU-VNTR. However, both Beijing isolates were differentiated by WGS.	Uzbekistan	[79]
2010	Phylogenetic tree of 14 DR-TB cases showed that each cluster shared differences mutation associated with INH resistance. Multiple DR-TB have arisen independently in each cluster and seem to be acquired. The isolates identified as Beijing had low fitness cost but promoted the transmission	South Africa	[80]
2012	WGS of 59 isolates, Samaran and global isolates, showed that the identified Bejing isolates with drug-resistance formed a monophyletic group. Genotypic DST showed strongly association to phenotypic DST. Mutation in *rpoC* was commonly found in the isolates carrying *rpoB* variants.	Russia	[81]
2013	MIRU-VNTR and WGS revealed identical genotype between two TB cases, support the possibility of transmission among Asian students. These students were infected with drug-resistance isolates from one patient or two patients but unable to identify sources.	United Kingdom	[82]
2013	The phylogenetic tree of 66 MDR-TB revealed that half of them was LAM9-c1 and associated with high mortality rate among male. The isolates shared identical drug-resistance mutations, supporting the primary resistance. LAM9-c1 is closely related to KwaZulu-Natal XDR-TB.	Panama	[83]
2013	Almost 70% of MDR-TB cases within four years was selected and perfumed the phylogenetic tree. Eight cases formed three clusters which having identical drug-resistance mutations to RIF and INH resistance.	Uganda	[84]
2014	The cladogram of 1,000 isolates revealed that two-third belonged to Beijing and 50% of them was MDR-TB. Mutations in *rpoB* associated with compensatory mutations in *rpoA* or *rpoC*. Mutations in *eis* promote their virulence. Combination of drug-resistance and compensatory mutations recover their fitness cost and enhance transmissibility.	Russia	[85]
2015	All available isolates over 15 years were sequenced. Of 66% of cases linked to at least other case. Transmission events were found and decreased over time. The highly transmission was driven by the Beijing/East Asian follow by the Indo Oceanic lineages.	Malawi	[86]

Table 11 Review articles in molecular epidemiology of DR-TB using WGS (Cont.)

Year	Finding	Region	Ref
2015	There was the first report about molecular epidemiology of *Mtb* (11 cases) which identified as Central American Beijing lineage. WGS was performed for five isolates and two of them carry mutations associated with FQs resistance.	Guatemala	[87]
2016	The researcher characterized the genome of *M. africanum* (lineage 5 and 6). They found that these lineages associated with geography.	Mali	[88]
2016	DR-TB isolates were performed WGS (90 cases). Phylogenetic tree revealed scatter MDR-TB. One-third of the cases had MDR-TB and 4 of them carry mutations associated with SLIDs resistance. Compensatory mutations in *rpoC* were observed in two cases.	Uganda	[89]
2016	The researcher applied WGS to investigate the transmission of 2 active XDR-TB patients and 33 people who contact the patients during 2 years.	United Kingdom	[90]
2016	WGS can reveal the direction of DR-TB transmission (16 cases) from 344 isolates.	United Kingdom	[91]
2017	The phylogeny of 138 *Mtb* isolates revealed the transmission of MDR-TB. In addition, XDR-TB transmission was caused by either infection of primary or acquired resistance isolates	Belarus	[92]
2017	WGS revealed the pattern of MDR-TB transmission among 324 patients from total cases (7,982 cases). According to the analysis, 87% of the clusters showed additional drug-resistance mutations through either emergence or fixation of mutations.	China	[93]
2017	WGS revealed the transmission of 386 XDR-TB patients and 84% of them formed 31 clusters. Of 212 patients formed the largest cluster while the other clusters contained clustering isolates for 2-14 cases.	South Africa	[94]
2017	WGS identified that 19% of 90 XDR-TB patients had a few SNP distance, 5 or fewer SNPs, suggesting the community-based transmission of XDR-TB.	South Africa	[95]
2017	Genomic analyses showed epidemiological link of TB infection in the same treatment centers weather they infected with the same or different isolates.	India	[96]
2017	Six clinical isolates were performed WGS to investigate the factors associated with transmission of Beijing-like isolates. Comparative genome analyses reveled that they shared the genetic variants associated with high EMB resistance, granuloma formation and virulence.	Columbia	[97]

Table 11 Review articles in molecular epidemiology of DR-TB using WGS (Cont.)

Year	Finding	Region	Ref
2017	Molecular epidemiology revealed the community and inter-patient transmission of MDR-TB. The genetic of mycobacterial had been changed during the treatment which promote SLDs resistance.	Nigeria	[98]
2017	WGS-based phylogenetic tree showed similar prevalence of East Asian and Central Asian isolates. Of 7 isolates from the clonal group (9 cases) had similar health post and geography, suggesting the epidemiological link of DR-TB transmission.	India	[99]
2018	WGS revealed the local transmission of XDR-TB isolates.	Papua New Guinea	[100]
2018	The researcher performed WGS of MDR-TB (4 cases) from the Horn of Africa and Sudan in order to investigate the outbreak from the refugees in 7 European countries.	European countries	[101]
2018	According to WGS-based spoligotyping, six DR-TB isolates had differences in spoligotype patterns, suggesting limited linking of DR-TB transmission using traditional method.	Tanzania	[102]
2018	*Mtb* isolates from initial and recurrent, at least 12 months, were selected to perform WGS in order to identify the recurrence of TB in treated patients, weather the recurrence caused by reactivation of the same isolate or reinfection with the new isolate. Mostly, the recurrences were due to reactivation.	Australia	[103]
2018	WGS was used to the genetic signatures associated with virulence or transmission for 82 isolates (Beijing, Manila and out group families) in Hawaii	USA	[104]
2018	The researchers characterized and compared the heterogeneous of *Mtb* clusters using VNTR and WGS. The heterogeneous VNTR clusters showed false clustering when compared to WGS.	Netherlands	[105]
2018	WGS revealed concurrent silent transmission of MDR-TB and emergence of *rpoB* Ile491Phe-bearing linage.	South Africa	[106]
2018	The researcher used WGS and Bayesian statistics to investigate the emergence, evolutionary and transmission of MDR-TB. Compensatory mutations was associated with higher drug-resistance rates and transmission.	Uzbekistan	[107]
2018	WGS could identify five additional clusters that was unable to recognize by 24-locus MIRU method. The clustering was mainly observed among drug-resistance isolates.	Vietnam	[108]

Table 11 Review articles in molecular epidemiology of DR-TB using WGS (Cont.)

Year	Finding	Region	Ref
2018	WGS was used to confirm the similarity of spoligotyping detected by MIRU-VNTR to track the TB transmission among household contacts.	Brazil	[109]
2018	WGS revealed the transmission of MDR-TB, genotypic DST profiles and compensatory mutations in *rpoC* and *rpoB*.	Saudi Arabia	[13]
2018	WGS was used to classify the lineages of *Mtb* and identify drug-resistance conferring mutations to the first-line drugs and SLDs.	Lebanon	[110]
2019	The researchers performed molecular epidemiology and genotypic DST using WGS to identify the recent transmission and additional drug-resistance.	Iran	[111]
2019	The researcher reviewed the genetic diversity of TB with mutations associated with drug-resistance across Latin America.	Latin America	[112]
2019	WGS revealed the transmission of 46 MDR-TB cases in Tunisia during 2012-2016. There were three main clusters, Harlem was predominant clone.	Tunisia	[113]
2019	WGS could identify the outbreak of 103 *Mtb* complex isolates in 2008-2014. The phylogeny displayed the outbreak of lineage 4.2.2.1 (37 isolates). In addition, almost 95% of them (35 isolates) shared identical drug-resistance mutations.	Serbia	[114]
2019	The researcher from Australia used WGS to investigate the transmission of MDR-TB in Papua New Guinea during 2010-2015. Half of MDR-TB cases (2 cases) from Australia citizens, who had epidemiological linked to the Torres Strait Protected Zone (TSPZ), showed cross-border transmission events from 2 plausible independent episodes of DR-TB transmission in the TSPZ.	Australia and Papua New Guinea	[115]
2019	The researcher integrated WGS, transcriptome and methylome across 22 *Mtb* isolates from Malawi. The results found that each lineage had different patterns of gene expression. Methylation played role in transcription of 50 genes. Expression of drug-resistance genes, *Rv2994*, *iniA* and *iniB* were different between ancient (L1 isolate) and modern (L2 and L4 isolates) isolates.	United Kingdom	[116]
2019	WGS identified four lineages (L1, L2, L3 and L4) from 81 *Mtb* isolates, the most predominance was L4 (90%). There were 6 clusters of transmission, L4.1.1 was predominant with TC6.	Mexico	[117]

Table 11 Review articles in molecular epidemiology of DR-TB using WGS (Cont.)

Year	Finding	Region	Ref
2019	WGS revealed the transmission of 61 DR-TB from household-based TB transmission in Peru.	USA	[70]
2019	From a collection of 81 *Mtb*, 18 isolates were clustering into eight clusters when using cut-off <10 SNPs. Lineage 3 was predominated in this setting.	Pakistan	[118]
2019	The researchers found that the isolates carrying *katG*-Ser315Thr shared similar genetic variation with isolates from South Africa.	Vietnam	[119]
2019	The researchers used WGS to reveal the genetic diversity of 178 *Mtb* isolates. The Manila was the highest prevalent follow by European-American and East Asian lineage. Some of MDR-TB cases showed identical variants.	Philippines	[120]
2019	Molecular epidemiology of MDR-TB in Peru, Spain and Italy were investigated. The results showed that, the transmission in Peru is predominate which comparable to the transmission in Europe during 2007-2017.	Peru	[121]
2020	Factors that caused XDR-TB transmission were identified, for example 2–3 months of cough and contact with urban area.	South Africa	[122]
2020	WGS analysis of 87 MDR-TB revealed that 60 isolates formed 10 clusters when using 5 SNPs cut-off.	Brazil	[123]
2020	Among 278 isolates (189 patients), there were 61 isolates that formed H3/4.2.1 clade. Also, WGS was also used to identify the re-activation, re-infection and mixed infection within patient.	Moldova	[124]
2020	The clustering analysis revealed that 39% of sample size were defined and grouped in to eight clusters when using 15 SNPs cut-off.	Liberia	[125]
2020	The researchers used WGS to characterize the polymorphisms of cold and hot spot areas of Guangxi Zhuang. There were three clusters, using <13 SNPs cut-off. One cluster was from cold spots and another two clusters were from hot spots. The hot spot area may contains higher transmissibility of the *Mtb* when compared to the cold spot.	China	[126]

Table 12 Review articles: implementation of WGS in TB research in Thailand

Year	Finding	Ref
2014	The researchers revealed the representative genomic draft of one MDR isolate, DS6701, from isolates causing outbreak in Thailand during 2002 to 2010.	[127]
2015	The outbreak of 148 MDR Beijing isolates was investigated. WGS was used to characterize the genetic background of four isolates, the first and the last three isolates. Their genome were clonal and highly stable, two or three SNPs were uniquely found in each of them. The low fitness cost mutation with additional SNP in *rpoB* was found, Leu731Pro, this might account for their transmissibility.	[8]
2016	WGS analysis for the Nontaburi genotype, isolated from TB meningitis, reveals several genetic signatures including genome size (4,364,461 bp), number of gene (4,154 genes), 48 RNAs, 64 pseudogenes and commonly 2,202 SNPs. The studied isolates were identified as Indo-Oceanic lineage.	[128]
2016	WGS was applied in order to differentiate between re-infection and persistent infection (relapse) among the serially MDR-TB and XDR-TB isolates from tree patients across two years.	[129]
2018	The researchers compared the genetic of *Mtb* isolated from TB meningitis (73 cases) and pulmonary TB (220 cases) using WGS. There were 242 SNPs that commonly found to all TB meningitis isolates, 28 were missense variants and normally found in the *pks* and the PE/PPE gene.	[130]
2019	The researchers compared phenotypic DST using the standard proportional method with WGS-based DST for 226 *Mtb* isolates. There were 51 drug-sensitive isolates, six Mono-drug resistant TB (Mono DR-TB) isolates, two Poly-drug resistant TB (Poly DR-TB) isolates, 88 MDR-TB isolates, 95 pre-XDR-TB isolates and 24 XDR-TB isolates. Two *in silico* tests, PhyResSE and TB-Profiler, were used to identify drug-resistance mutations. INH, RIF and AMK showed the highest concordance between two DSTs. However, Low concordance was found in EMB, ETO and FQs.	[28]
2020	The researchers used WGS to characterize rarer proto-Beijing (L2.1) strains that spanning 13 years in Thailand. Of 43.2% were clustered in MDR-TB or XDR-TB transmission, using <13 SNPs cut-off. All XDR-TB cases were caused by primary resistance rather than inadequate treatment. The 47 signature mutations and partial deletion of *fadD14* were identified in an XDR-TB cluster.	[131]

CHAPTER III
CLUSTERS OF DRUG-RESISTANT *MYCOBACTERIUM TUBERCULOSIS* DETECTED BY WHOLE-GENOME SEQUENCE ANALYSIS OF A NATIONWIDE SAMPLE; THATLAND, 2014-2017

1. Introduction

TB, caused by *Mtb*, is a major global public health issue. Southeast Asia contributes significantly (44%) to global TB cases, with Thailand in the top 14 countries for DR-TB incidence [1]. DR-TB, including RR-TB, and strains with additional resistance to INH (MDR-TB), remains a great challenge for TB control. In 2018, there were ~500k new cases of RR-TB globally, of which 78% were MDR-TB [1]. More worrying is XDR-TB, which further exhibits resistance to one fluoroquinolone and one injectable second-line drug. The average proportion of global MDR-TB cases with XDR-TB is 6.2% [1]. In Thailand, despite reducing incidence of TB, the number of MDR-TB cases nearly doubled between 2014 and 2018 [1], some likely to be XDR-TB. Treatment for patients with DR-TB is prolonged, expensive and outcomes are poor.

WGS of *Mtb* provides insights into drug resistance, where mechanisms almost exclusively involve mutations (mostly SNPs, but also insertions and deletions (indels)) in genes coding for drug targets or drug-converting enzymes. WGS data can also provide insights into transmission and the dating of clusters [74], where strains with near-identical genetic variants are likely to be part of a transmission chain [86]. Analysis of *Mtb* WGS data from isolates across Thailand could provide much-needed insights into MDR/XDR-TB transmission. Previous studies of DR-TB have used genotyping techniques (e.g. spoligotyping, MIRU-VNTR and RFLP) [16, 132], but these methods have limited resolution for inferring transmission as they investigate less than 1% of the *Mtb* genome. A recent WGS analysis revealed possible clonal transmission of four MDR-TB isolates in Kanchanaburi Province [8]. However, the extent of MDR and XDR-TB clusters across Thailand is currently unknown. Here, we

aimed to investigate the clustering patterns and risk factors of possible MDR, pre-XDR and XDR-TB transmission clusters, across Thailand using WGS data.

2. Materials and methods

2.1 Studied population and setting

Between 2014 and 2017, 2,071 *Mtb* culture-confirmed MDR-TB, pre-XDR and XDR-TB cases were listed in the laboratory records of the NTRL; Ministry of Public Health and Siriraj Hospital, Mahidol University, Thailand. These two laboratories cover 230 hospitals handling the majority of DR-TB cases in Thailand (Supplementary Table 1, Supplementary Table 2) [50]. We selected 547 *Mtb* isolates from MDR-TB and pre-XDR cases across 6 regions and 71/77 provinces nationally. We also included all retrievable (n = 32) XDR-TB isolates (Supplementary Table 3). For eleven cases, pairs of isolates collected at different times were used as internal controls for SNP distances. In each control pair, the isolate with the most mutations associated with drug resistance and/or the chronologically earlier isolate was included in the studied population (n = 579). Demographic data were retrieved from laboratory records. The study protocol was approved by the Center for Ethics in Human Research, Khon Kaen University (HE601249).

2.2 Definition and pattern of DR-TB used in this study

DR-TB types were diagnosed using phenotypic DST assay. This method determines ability of *Mtb* growth in medium containing anti-TB drugs at the critical concentration (CC) value recommend by WHO. The DST results were reported as either susceptible or resistant to tested drug. According to DST profiles determined by phenotypic DST, DR-TB patterns that used in present study are classified in Table 13.

Table 13 Definition of DR-TB that used in this study

Pattern of DR-TB	Definition
Susceptible TB	Pan susceptible to both of first and SLDs
Mono DR-TB	Resistance to only one first-line drug
Poly DR-TB	Resistance to more than one first-line drugs, except INH and RIF
RR-TB	Resistance to RIF alone
IR-TB	Resistance to INH alone

Table 13 Definition of DR-TB that used in this study (Cont.)

Pattern of DR-TB	Definition
MDR-TB	Resistance to both of INH and RIF
Pre-XDR-TB	MDR-TB and additional resistance to either any FQ or SLID
XDR-TB	MDR-TB and additional resistance to at least one FQ and any SLID

2.3 Phenotypic drug-susceptibility testing

Phenotypic DST was performed using the standard agar proportional method in LJ medium [19]. Drug concentrations used were 0.2 µg/mL for INH, 40.0 µg/mL for RIF, ETO, CAP and DCS, 2.0 µg/mL for EMB, OFX and LFX, 4.0 µg/mL for STR, 30.0 µg/mL for KAN and 0.5 µg/mL for PAS. *Mtb* H37Rv was used as the susceptible reference strain.

2.4 Whole-genome sequence analysis

Multiple loops of *Mtb* colonies were used for genomic DNA extraction (using the cetyl-trimethyl-ammonium bromide-sodium chloride (CTAB) method) [133]. WGS data of 590 *Mtb* isolates were produced by NovogeneAIT, Hong Kong, using the HiSeq (Illumina) platform generating 150-bp paired-end reads. The quality of sequence reads was checked using FastQC (version 0.11.7) [134]. High-quality reads from each isolate were mapped onto the *Mtb* H37Rv reference genome (NC_000962.3) using BWA-MEM (version 0.7.12) [135]. The average depth of sequencing coverage was high (341.01±61.98). SAMtools (version 0.1.19) [136] and GATK (version 3.4.0) [137] were used to call SNPs and indels. Variants were filtered based on a minimum coverage depth of 10-fold and Q20 minimum base-call quality score, and the intersection set of GATK and SAMtools variants was retained. An online tool, TB-Profiler (version 2.8.6) [138, 139], was used to infer drug resistance and *Mtb* lineage membership based on SNPs from the WGS data. The WGS data are available in the ENA Sequence Read Archive (accession numbers PRJNA598981 and PRJNA613706).

2.5 Phylogenetic analysis

A phylogenetic tree was constructed based on 26,541 high-confidence SNPs among 590 isolates, using the maximum-likelihood method with the selected general time-reversible with gamma-distribution model, implemented within MEGA (version

10.1) [140]. The 130 SNPs known to be associated with DR-TB found in this study were excluded to ensure they would not affect the phylogenetic analysis. A bootstrap consensus tree was inferred from 1,000 replicates. The phylogenetic tree image was produced using iTOL [141].

2.6 Data analysis

Isolates forming monophyletic groups in which many or all pairs differed by ≤25 SNPs were placed in the same clade. Clusters included isolates differing between 0 and 11 SNPs. Members of a single cluster we regard as possibly descended from a single clone via recent transmission. Less-recently transmitted isolates within a clade differed between 12 and 25 SNPs. The clustering percentage was calculated by (no. of clustering isolates/total no. of isolates) × 100. Isolates with acquired DR-TB were differentiated from possible primary DR-TB (MDR-TB, pre-XDR and XDR-TB) isolates based on acquisition of additional resistance-associated mutations, especially those associated with resistance to FQs, KAN and/or CAP, drugs that are used for DR-TB classification. In clusters containing isolates with different types of DR-TB (such as MDR-TB and XDR-TB), the acquisition of additional drug-resistance SNPs and co-ancestral relationships were used to differentiate between two patterns of acquired-resistance: chronological (ancestral strain had fewer mutations and/or lesser type of DR) or non-chronological (ancestral strain had more mutations and/or stronger type of DR). Although XDR-TB and pre-XDR could be considered as subsets of MDR-TB, we have treated all three as separate categories in our analyses.

All data were analyzed using R statistical software (version 3.6.1). P values <0.05 were considered statistically significant. Association between clades/clusters and geography was analyzed using χ^2 tests and visualized by the R package "vcd" (version 1.4-8). Odds ratios (ORs) with 95% confidence intervals (CIs) were calculated using the R package "epiR" (version 1.0-4). Factors associated with clustering isolates were tested using the Student *t*-test (numerical data), χ^2 test or Fisher's exact test (categorical data), where applicable. Graphs were constructed using the R package "ggplot2" (version 3.2.1). Phylo-maps were build using the package "phytools" (version 0.7-20).

3. Results

3.1 Study population and characteristics

Most (466; 80.5%) of the 579 culture-confirmed DR-TB cases in the studied population were MDR-TB, followed by 81 pre-XDR (14.0%) (Supplementary Table 2). We included all available XDR-TB isolates (n = 32), constituting 5.5% of our samples, but only 1.5% of the culture-confirmed 2,071 DR-TB isolates collected nationally from 2014 to 2017. Most patients were male (n = 419; 73.1%) and mean age was 43.5 (±14.7) years (Supplementary Table 3). Central and northeast regions of Thailand had the highest DR-TB proportions (Figure 11). The three provinces with the highest number of DR-TB cases were Bangkok (n = 85; 14.7%), Kanchanaburi (n = 51; 8.8%) and Chonburi (n = 37; 6.4%) (Figure 11, Supplementary Table 4).

3.2 Phylogenetic analysis

Most of the *Mtb* isolates belonged to Lineage 2 (n = 482; 83.2%), followed by Lineage 1 (n = 67; 11.6%), Lineage 4 (n = 29; 5.0%) and Lineage 3 (n = 1; 0.2%) (Figure 12, Supplementary Table 5). Lineage 2.2.1 (n = 413; 71.3%) was the main sub-lineage causing MDR/pre-XDR/XDR-TB.

3.3 Clustering and possible transmission clusters

The phylogenetic tree (Figure 12) showed enormous diversity among the DR-TB isolates from Thailand. Many isolates were quite distinct, differing from all others at (mean±SD) 657±626 SNPs. The majority (n = 319; 55.1%) grouped into 13 clades each consisting of 5-86 isolates (Figure 13**,** Figure 14 and Figure 15). Clades #1, #6, #11 and #13 each consisted of a single small cluster of closely related isolates and the remaining clades included one or more possible clusters (Figure 15). The isolates grouped in each clade were significantly associated with a particular geographical region ($p < 0.001$; Figure 16). Clade #1 (Figure 14, panel A) was only found in Trat Province and clade #13 predominated in Kanchanaburi (Figure 14, panel M).

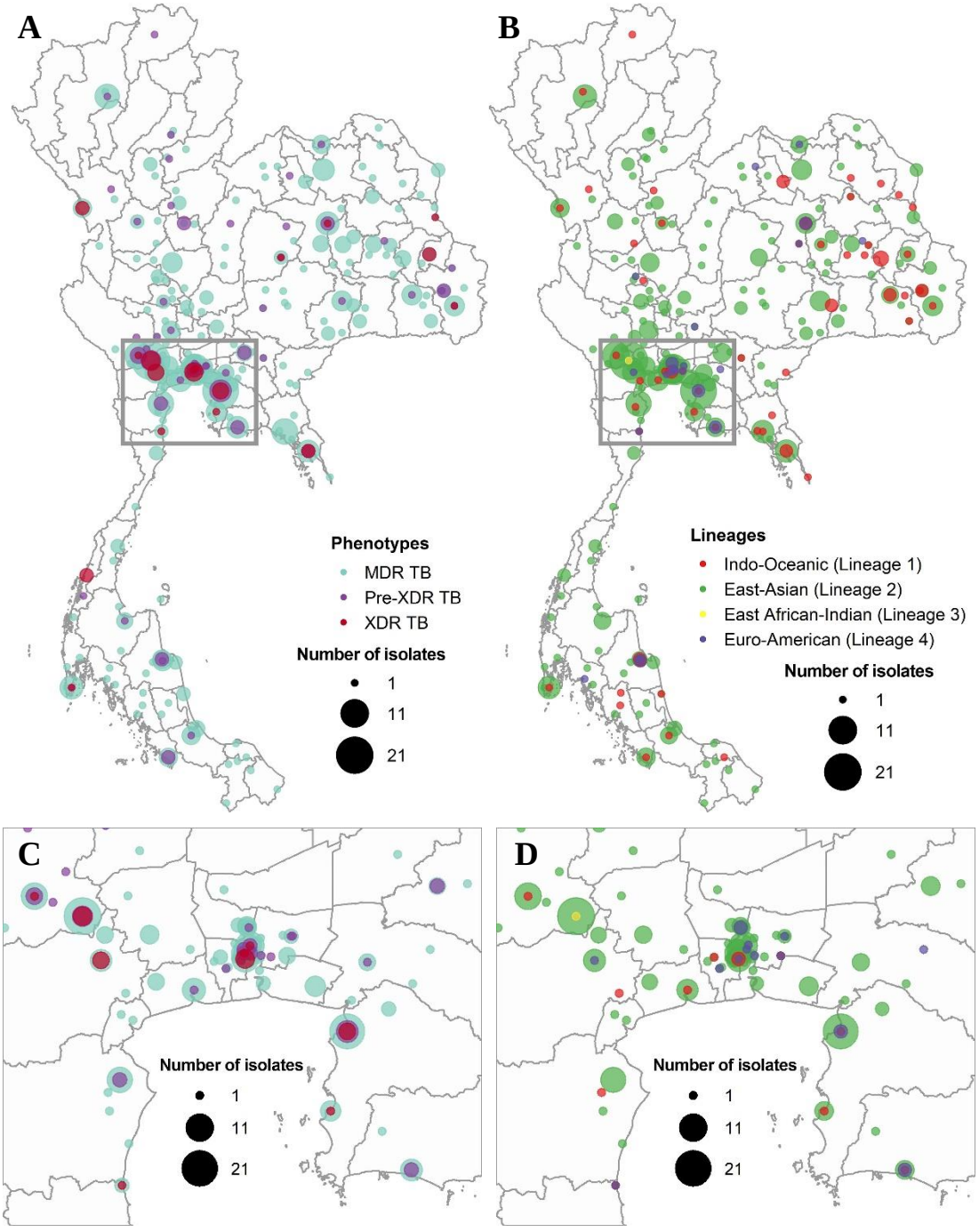

Figure 11 Geographical and lineage distribution of 579 DR *Mtb* isolates in Thailand from 2014 through 2017. (A) Geographical distribution of MDR-TB, pre-XDR-TB and XDR-TB. (B) Lineage distribution of DR *Mtb*. Boxed insets, expanded on the right, of DR types (C) and lineages (D). The size of each circle is proportional to the number of isolates.

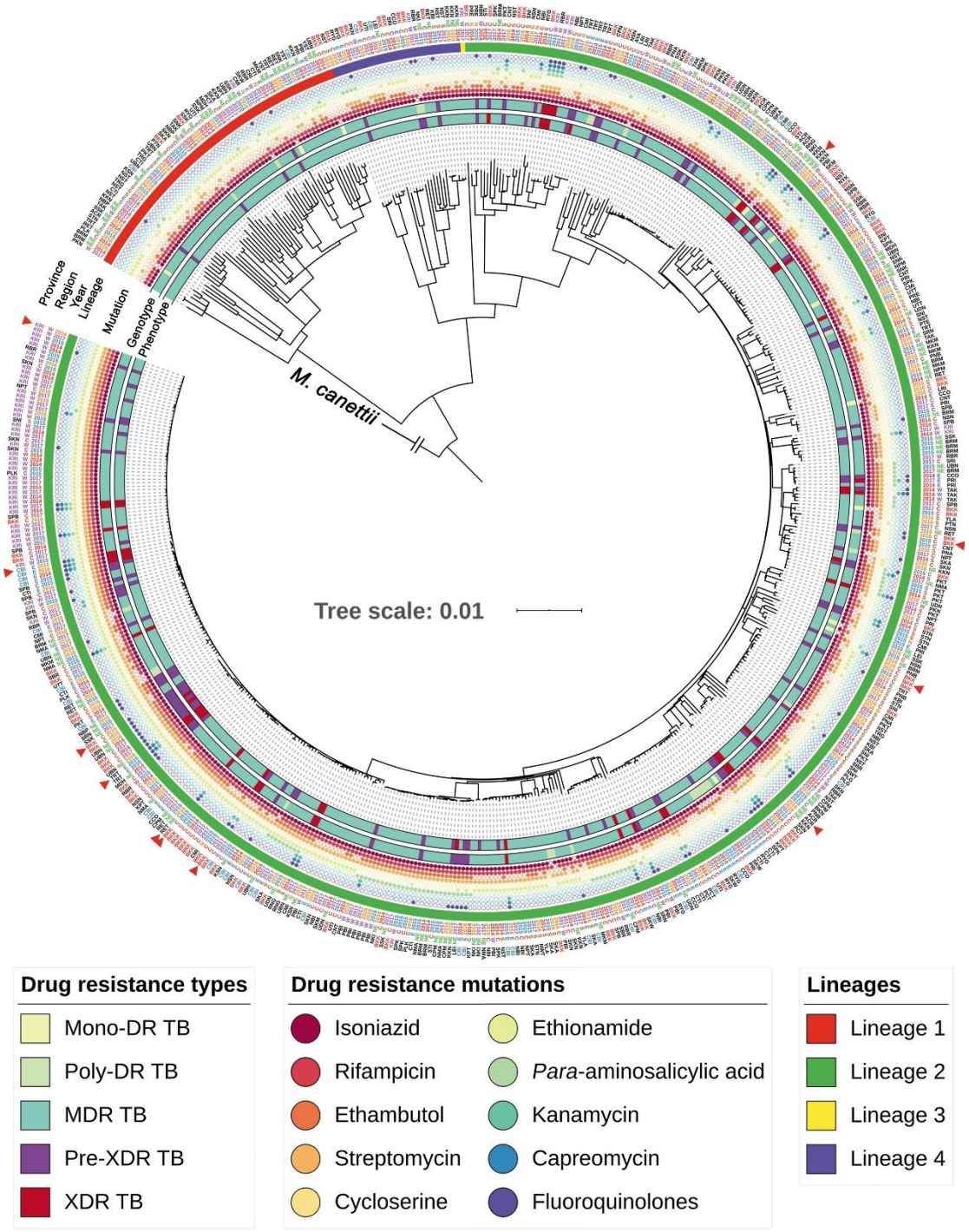

Figure 12 Phylogenetic tree for the 590 DR *Mtb* isolates. From inner to the outer circles: culture-based phenotypic drug-susceptibility test, whole genome sequencing-based drug-resistance profile, drug-resistance mutations, lineage, year of collection, regions and provinces. The red triangles indicate the paired isolates from the same patients (n = 11).

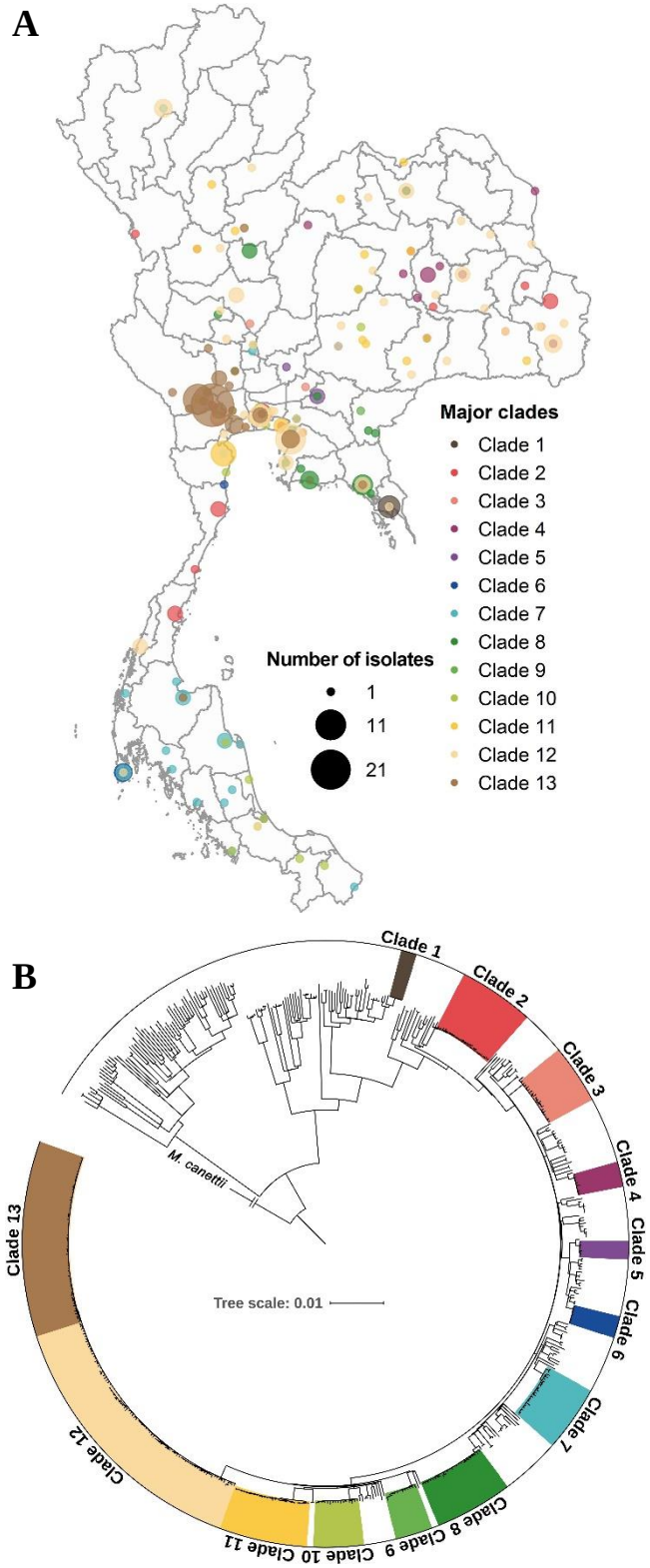

Figure 13 Geographical distribution of 13 major clades (319 isolates) across Thailand. (A) The size of each circle is proportional to the number of isolates. (B) The 13 clades are identified and highlighted in the outer circle.

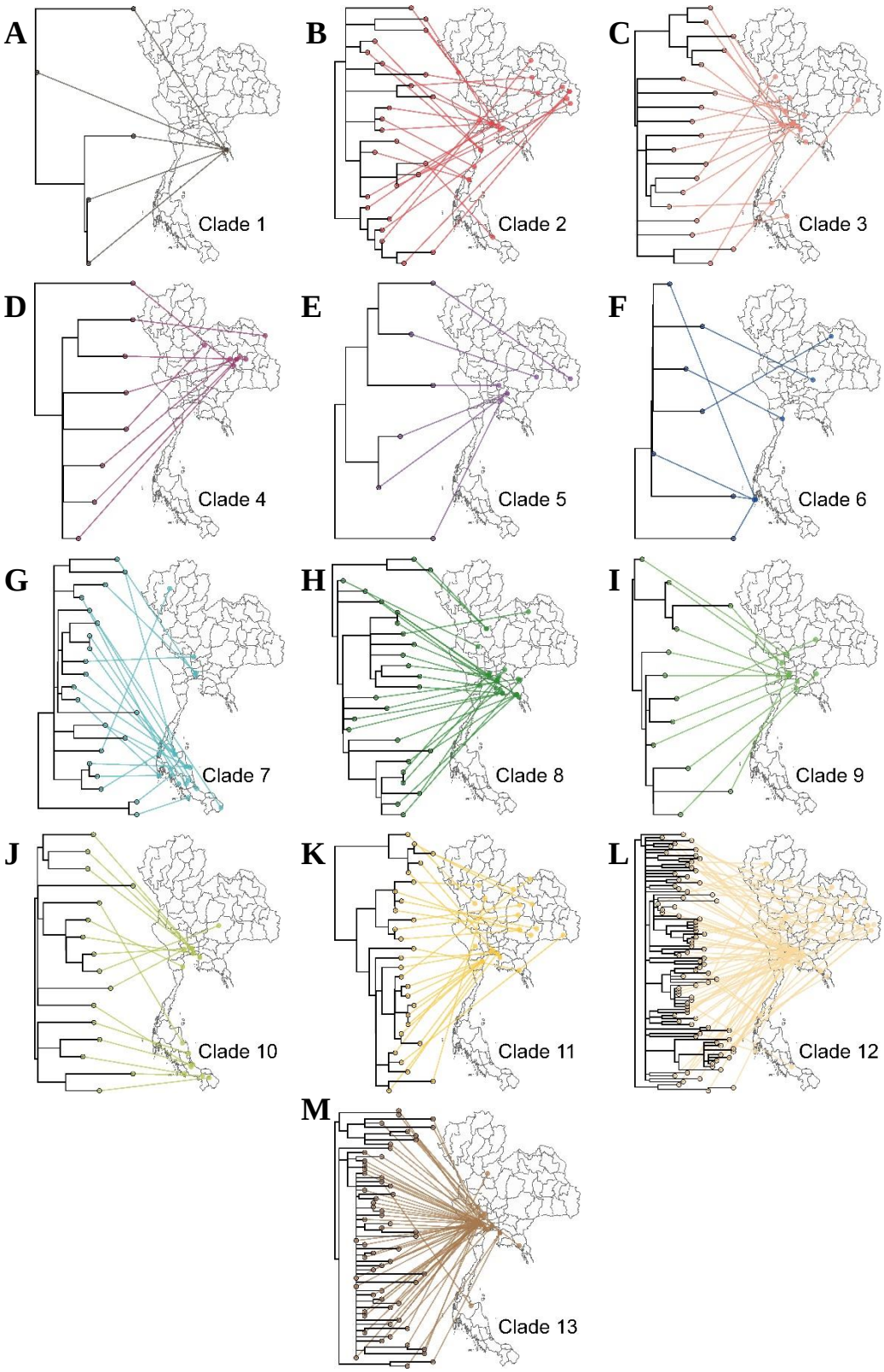

Figure 14 Geographical distribution of thirteen major clades of DR-TB in Thailand. Each of the 13 major clades (A-M) is associated with particular geographical regions.

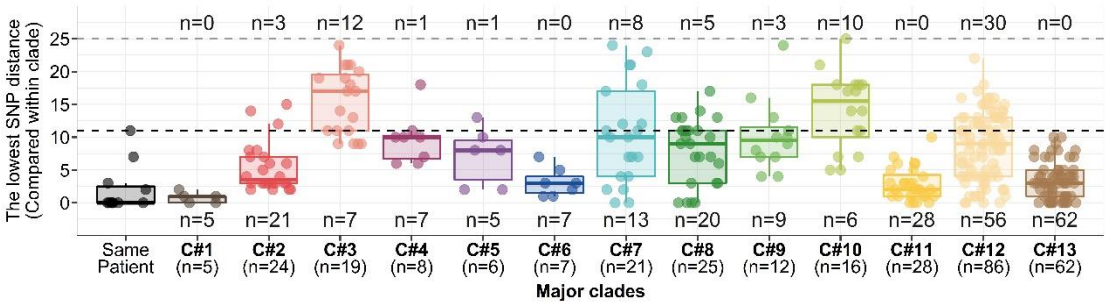

Figure 15 Comparisons of the proportion of isolates in each clade that differ by <11 single nucleotide polymorphisms (SNPs) (suggesting recent transmission) and those that differ by 12–25 SNPs in many or all pairs (suggesting less-recent transmission).

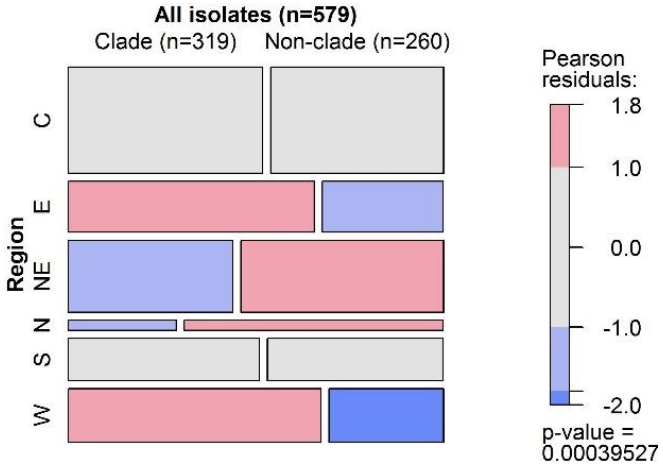

Figure 16 Association between geographical regions and 13 clades.

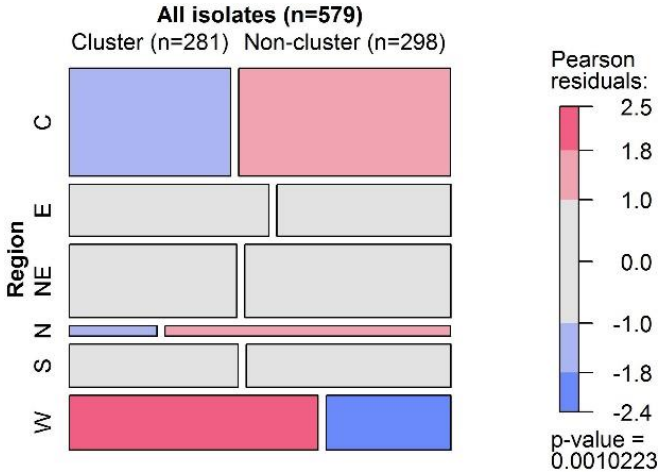

Figure 17 Association between geographical regions and 89 clusters.

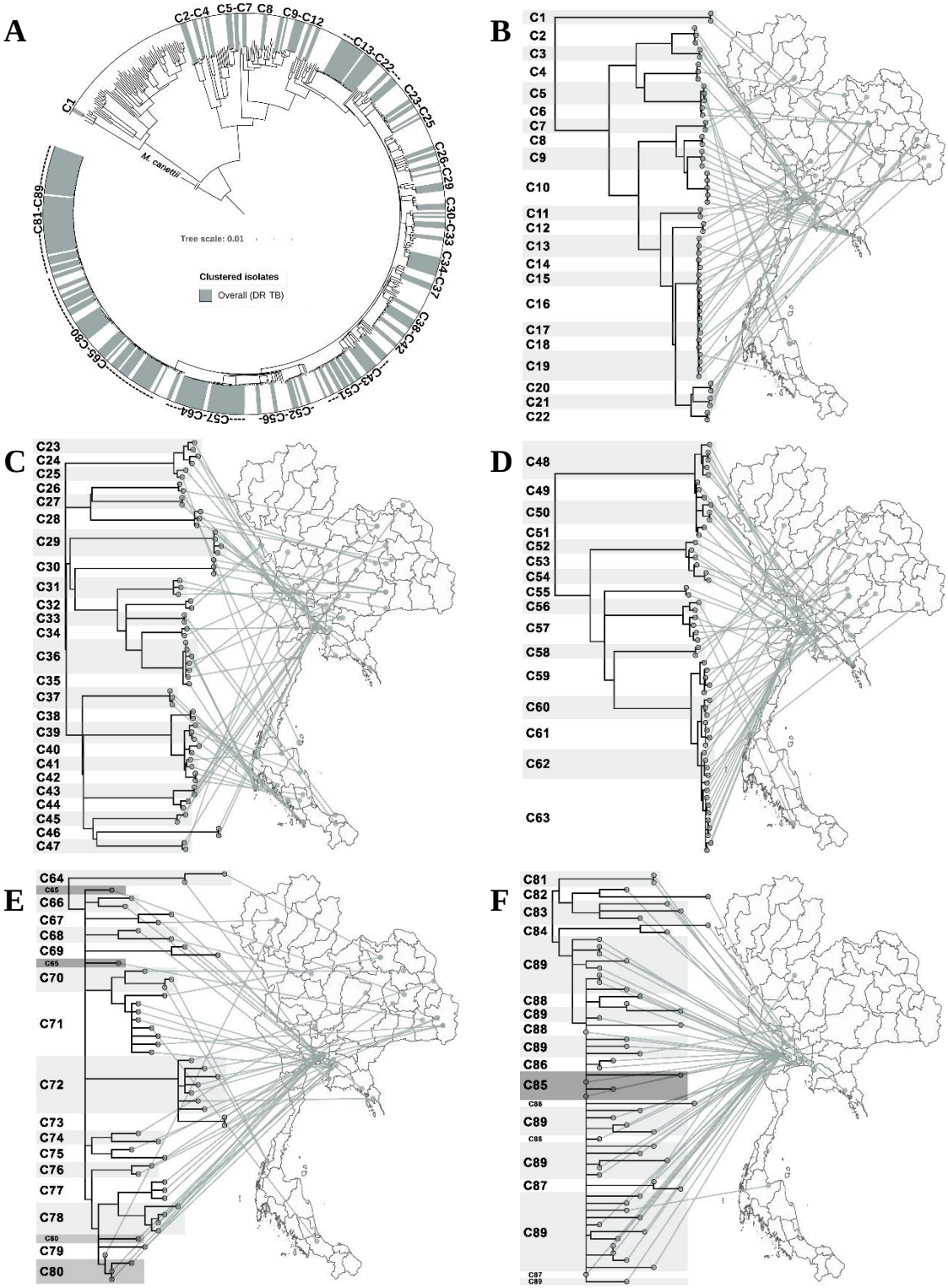

Figure 18 All clusters of DR-TB isolates. (A) 89 clusters are highlighted in the outer circle. (B-F) Phylogeographical links of each cluster are shown. For clarity, clusters are divided among five phylomaps. Some isolates in closely related clusters (C64-65, C79-C80 and C85-C89) crossed localities.

A total of 89 clusters contained 281 isolates (48.5%) (Supplementary Table 6). Sixty clusters, containing between 2 and 34 isolates, fell within the major clades. A further 29 smaller clusters occurred elsewhere in the tree. Most isolates within a cluster shared geographical links (Figure 18, Supplementary Table 6). The percentages of MDR, pre-XDR and XDR-TB isolates (based on phenotypic DST) that fell into clusters were 46.1% (215/466), 49.4% 40/81), and 81.3% (26/32), respectively (Supplementary Table 6). Pairwise SNP distances within and between each of the 89 clusters are given summarized (Supplementary Table 7).

Some clusters included isolates with different types of DR-TB. Nineteen of the 89 clusters (C2, C7, C10, C16, C22, C36, C37, C40, C43, C49, C59, C60, C63, C70, C72, C76, C80, C83 and C89) had a chronological pattern based on progressive increase in numbers of DR mutations from base to tips in the phylogeny (Supplementary Table 8). The pattern of DR mutation changes was non-chronological in clusters C21, C23, C32, C35, C41, C55, C71 and C75. Among the 281 clustering isolates, 81.9% were classified as possible primary DR-TB (n = 230) including MDR-TB (n = 176/205; 85.9%), pre-XDR (n = 29/46; 63.0%) and XDR-TB (n = 14/19; 73.7%). In addition, ten phenotypically MDR isolates and one phenotypically pre-XDR isolate were identified as possible examples of primary IR-TB (n = 11) based on genotypic DR. Other clustering isolates (n = 51/281, 18.1%) exhibited acquired DR-TB (MDR-TB (n = 29/205; 14.1%), pre-XDR (n = 17/46; 37.0%) and XDR-TB (n = 5/19; 26/3%) (Table 14).

Table 14 Characteristics of isolates within 89 clusters

Clustered isolates[a] (n=281)	**DR-TB types[b]**			
	IR-TB (n = 11)	**MDR-TB (n = 205)**	**pre-XDR-TB (n = 46)**	**XDR-TB (n = 19)**
Possible primary DR-TB[c] (n = 230, 81.85%)	11 (100.0%)	176 (85.85%)	29 (63.04%)	14 (73.68%)
Possible acquired DR-TB[c] (n = 51, 18.15%)	0 (0.00%)	29 (14.15%)	17 (36.96%)	5 (26.32%)

[a] Using a pairwise-difference range of 0-11 SNPs, 89 clusters were recognized. [b] DR-TB types were based on genotypic DST. [c] Possible primary DR-TB isolates were differentiated from acquired DR-TB isolates based on the acquisition of mutations associated with drug-resistance and from co-ancestral relationships.

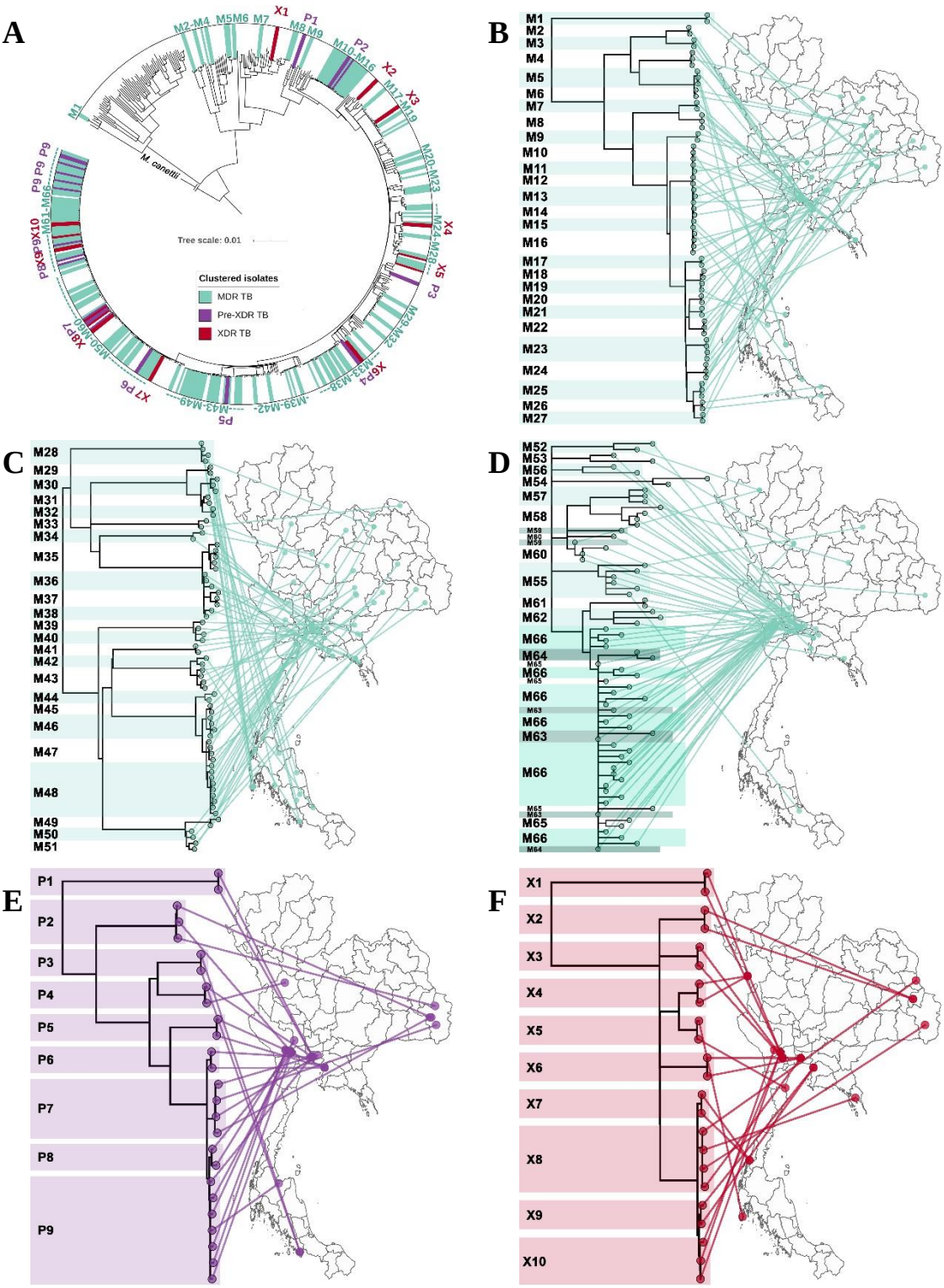

Figure 19 Clusters of DR-TB isolates based on phenotypic DST. (A) 66 MDR-TB (M1–M66), 9 pre-XDR-TB (P1–P9), and 10 XDR-TB (X1–X10) clusters are highlighted in the outer circle. Phylogeographical links of MDR-TB (B–D), pre-XDR-TB (E), and XDR-TB (F) clusters are shown.

Table 15 Demographic and other factors associated with clustering (≤11 SNP difference between) isolates

Characteristic	All isolates (n = 579)	Clustering isolates, no. (%)		OR (95% CI)
		Clusters (n = 281)	Non-clusters (n = 298)	
Gender (n = 573)				
Male	419 (73.12)	198 (70.71)	221 (75.43)	0.79 (0.54-1.14)
Age (n = 508)				
mean±SD (year)	43.51±14.68	42.02±15.23	44.94±14.03	NA
Region				
Central	183 (31.61)	79 (28.11)	104 (34.90)	0.73 (0.51-1.04)
Eastern	88 (15.20)	47 (16.73)	41 (13.76)	1.26 (0.80-1.98)
Northeastern	125 (21.59)	56 (19.93)	69 (23.15)	0.83 (0.56-1.23)
Northern	17 (2.94)	4 (1.42)	13 (4.36)	0.32 (0.10-0.98)
Southern	73 (12.61)	33 (11.74)	40 (13.42)	0.86 (0.52-1.40)
Western	93 (16.06)	62 (22.06)	31 (10.40)	**2.44 (1.53-3.89)[a]**
Lineage				
2.1	31 (5.35)	12 (4.27)	19 (6.38)	0.66 (0.31-1.38)
2.2.1	413 (71.33)	236 (83.99)	177 (59.40)	**3.59 (2.42-5.32)[a]**
2.2.1.1	32 (5.53)	16 (5.69)	16 (5.37)	1.06 (0.52-2.17)
2.2.1.2 and 2.2.2	6 (1.04)	2 (0.71)	4 (1.34)	0.53 (0.05-3.71)
4	29 (5.01)	13 (4.64)	16 (5.35)	0.86 (0.41-1.82)
1	67 (11.57)	2 (0.71)	65 (21.81)	**0.03 (0.01-0.11)[a]**
3	1 (0.17)	0 (0.00)	1 (0.34)	NA
Drug-resistance mutations				
Isoniazid (n = 565)				
katG Ser315Thr	448 (79.29)	252 (89.68)	196 (69.01)	**3.90 (2.46-6.18)[a]**
inhA -15 c/t	52 (9.20)	7 (2.49)	45 (15.85)	**0.14 (0.06-0.31)[a]**
Rifampicin (n = 554)				
rpoB Ser450Leu	279 (50.36)	176 (65.19)	103 (36.27)	**3.29 (2.32-4.66)[a]**
Ethambutol (n = 335)				
embB Met306V	85 (25.37)	44 (20.75)	41 (33.33)	**0.52 (0.32-0.86)[a]**
embB Gly406Asp	66 (19.70)	59 (27.83)	7 (5.69)	**6.39 (2.81-14.51)[a]**
embB Met306Ile	56 (16.72)	27 (12.74)	29 (23.58)	**0.47 (0.26-0.84)[a]**
Streptomycin (n = 349)				
rpsL Lys43Arg	295 (84.53)	188 (89.95)	107 (76.43)	**2.76 (1.52-5.01)[a]**
Ethionamide (n = 268)				
ethA 639-640del	143 (53.36)	105 (73.43)	38 (30.40)	**6.33 (3.72-10.77)[a]**
inhA -15 c/t	65 (24.25)	9 (6.29)	56 (44.80)	**0.08 (0.04-0.18)[a]**
Para-aminosalicylic acid (n = 99)				
folC Ser150Gly	39 (39.39)	32 (50.79)	7 (19.44)	**4.28 (1.63-11.19)[a]**

[a] OR (95% CI) with statistically significant p-values. NA, not applicable

Among clustered isolates, there was some discordance between phenotypic DST findings (MDR-TB (n = 215), pre-XDR (n = 40) and XDR-TB (n = 26)) and genotypic DST results (poly-DR (n = 11), MDR-TB (n = 205), pre-XDR (n = 46) and XDR-TB (n = 19)) (Supplementary Table 8). Based on phenotypic DST, 66 MDR-TB, nine pre-XDR and ten XDR-TB clusters were identified (Supplementary Table 8; Supplementary Table 9; Figure 19, panels A-F). Most pre-XDR and XDR-TB clusters had hospital-based links (Supplementary Table 9). All phenotypic DR-TB clusters and resistance types, stratified by province, are shown (Supplementary Table 10).

3.4 Factors associated with possible DR-TB transmission clusters

TB patients from whom clustering isolates were obtained had an average age of ~42 years. Isolates falling within clusters were significantly associated with geographical regions (p = 0.001; Figure 17). TB patients living in western provinces had a higher risk of being within possible DR-TB transmission clusters than those elsewhere (OR 2.44, 95% CI 1.53-3.89) (Table 15). Lineage 2.2.1 (versus other lineages) was associated with a higher risk of possible DR-TB transmission clusters (OR 3.59, 95% CI 2.42-5.32). Lineage 1 had the lowest risk of being represented in DR-TB transmission clusters (OR 0.03, 95% CI 0.01-0.11). Clustering isolates had drug-resistance mutations such as *katG* Ser315Thr, *rpoB* Ser450Leu and *embB* Gly406Asp (Table 15).

4. Discussion

MDR- and XDR-TB are serious global problems, but few studies have focused on their transmission at a nation-wide resolution. Thailand has a high burden of MDR-TB and increasing numbers of MDR-TB cases [1]. We sourced 579 DR-TB isolates across 71 provinces between 2014 and 2017. Nearly half of these were in possible transmission clusters, mostly involving *Mtb* lineage 2.2.1. Eighty-nine clusters, most distributed among 13 major clades, contributed to multi-clonal MDR-TB outbreaks associated with specific regions in Thailand. Bangkok, Kanchanaburi and Chonburi were the provinces with the highest proportions of MDR-TB, pre-XDR and XDR-TB clusters (i.e. groups of isolates differing by ≤11 SNPs). We used two criteria to select SNP cut-off values. First, the ≤11 SNP difference cut-off for a cluster was derived directly from the maximum number of differences between the 11 paired isolates used

as an internal control. Second, we used a SNP-cutoff concordant with, or more stringent than, previous studies [71, 93, 142, 143]. Our 11-SNP cut-off was proportionally 0.0004 of the 26,541 SNPs in our total set. This proportion was concordant with a previous study [144], and more stringent than other studies [93, 143]. A <12-SNP cutoff has been previously proposed as the upper boundary for possible cluster-transmission events [74].

Phylogenetic analysis identified 13 major clades, each associated with a particular region(s). Pairwise-SNP differences between isolates within clades ranged from <11 to about 25, suggesting a range of divergence times from a common ancestor. Based on the transmission-time estimates (0.5 SNP/genome/year) for *Mtb* [74], some of these major clades might have begun to circulate in Thailand around 20-40 years ago, others more recently. Isolates differing by 12-25 SNPs nevertheless often shared geographical links. For example, 17/21 isolates (81%) in clade #7 (Figure 14, panel G), which had pairwise differences indicating a relatively non-recent common ancestor, were located within neighboring provinces of southern Thailand. Clades #1, #6, #11 and #13 each consisted of isolates differing at very few SNPs, giving us confidence that these were likely examples of recent transmission. Nonetheless, isolates in clade #6 often occurred in different provinces.

The largest and most recent clade was clade #13 (Figure 14, panel M) comprised of 62 cases (46 MDR-TB, 11 pre-XDR and 5 XDR-TB based on phenotypic DST) found in the western region, especially Kanchanaburi. This suggests that clones of pre-XDR and XDR-TB may emerge from recent MDR-TB ancestors. We confirmed a previous report [145] that there was a large MDR-TB outbreak in Kanchanaburi. Additionally, clade #13 is sister to clade #12, which consists of strains that spread in both Central (especially Bangkok) and Northeast Thailand and also contains less-recently transmitted strains. Therefore, the MDR-TB outbreak clade in Kanchanaburi was derived from a less-recently transmitted clade elsewhere in Thailand.

We identified 89 clusters (isolates in each differing by ≤11 SNPs) of DR-TB in Thailand. The clustered isolates showed a strong association with geographical region. The largest cluster (C89), within clade #13 in Kanchanaburi, comprised 34 isolates (27 MDR-TB and 7 pre-XDR-TB based on phenotypic DST). In South Africa, WGS analysis of a large XDR-TB cohort (>400 cases) from a single province showed that

only 30% of participants had clear epidemiological links (person-person or hospital link): 70% of transmission events may have resulted from casual contact between individuals not known to one another [146]. Another study there showed that 19% of XDR-TB patients discharged from the hospital caused secondary XDR-TB cases in the community [95]. Here, we found nine clusters of pre-XDR (the largest with 7 isolates) and ten clusters (the largest with 4 isolates) of XDR-TB in Thailand (Supplementary Table 9; Figure 19).

To reflect the extent of the DR-TB outbreak in Thailand, we calculated the proportion of isolates falling into the 89 DR-TB clusters (Table 14). In some clusters, isolates exhibited different types of DR-TB associated with chronology, revealing the progression of DR mutations in the phylogeny, moving from the ancestor towards the tips of the tree (Supplementary Table 8). Based on mutation-acquisition analysis within this phylogeny, examples of possible primary resistance were seen in 85.9% of MDR-TB, 63% of pre-XDR and 73.7% of XDR-TB cases. Eight clusters included isolates with different types of DR and more resistance-associated mutations in the ancestral strain than in its descendants. This situation might be explained by different durations of the latency stage occurring after transmission events leading to the emergence of less troublesome DR-TB cases (such as MDR-TB) later than the more troublesome cases (such XDR-TB) [147]. Because not all cases from the possible transmission chain could be included, undetected primary resistance might exist. Data from all DR-TB cases in the community, information of treatment history and known exposure are needed to accurately and completely estimate the extent of primary DR-TB. The proportion of primary DR-TB cases could be higher as we reported numbers of MDR-TB cases excluding pre-XDR and XDR-TB (which were each reported as a separate subset). Also, some index cases might not have been included in the selected population.

Previously reported factors contributing to MDR-TB transmission include: illicit drug usage [10]; delayed TB diagnosis and being older than 45 years [93]; being single, low-income, suffering frequent stress and other diseases and lacking medical insurance [148]. Lineage 2 predominated in previous studies of transmission of MDR-TB [10, 93, 149]. We found that infection with Lineage 2.2.1 is the strongest predictor (3.6-fold) of DR-TB clusters whereas infection with Lineage 1 had the lowest risk. Living

in the western region of Thailand increased the risk of being in DR-TB clusters by 2.4-fold. The western region, being close to the border with Myanmar, differs from other regions of the country both in terms of ethnicity and economic development. These differences might explain the increased risk there [150]. Previously, clustering isolates were more likely to have mutations of *rpoB* Ser450Leu [12, 93], *katG* Ser315Thr or the *inhA* promoter [151]. We also found a pattern of drug resistance-associated mutations (*katG* Ser315Thr, *rpoB* Ser450Leu, *embB* Gly406Asp, *rpsL* Lys43Arg, *ethA* 639-640del and *folC* Ser150Gly) in clusters.

The DR-TB situation in Thailand is a major concern and requires urgent implementation of control measures such as active case finding to disrupt the transmission chain. There should be targeted intervention and contact tracing in hotspot regions. The mortality rate and cost of treatment of XDR-TB is very high [152], hence these DR types should be the priority for intervention. The large size of some clusters might reflect their high transmissibility [153]: tracking clade #13 at Kanchanaburi should be a priority. Besides the 13 major clades, there were several small clusters of DR-TB in many provinces. The potential for expansion of these small clusters is unknown. Here, we also identified the hotspot provinces to help prioritize locations for intervention.

Globally, there have been few studies at the nation-wide scale using WGS analysis of MDR-TB, pre-XDR and XDR-TB [9-12, 14]. Older studies have used blunt genotyping tools (e.g. IS*6110* RFLP, spoligotyping and MIRU-VNTR) with limited or convenient sample sizes. DR-TB studies using WGS in Saudi Arabia and Portugal have revealed transmission clusters of MDR-TB, however, they had small samples and provided limited data on epidemiological links [13, 14]. Extrapolating from our findings, primary-resistant TB strains may be the main contributors to the current global problem of high MDR/XDR-TB prevalence.

This study is subject to a number of limitations. First, our study was retrospective rather than prospective. There was a lack of socio-economic data for analysis. There was also a lack of fine-scale data of epidemiological links: possible transmission clusters were presumed only from the genetic distances among isolates and each patient's hospital and province of residence. Also, an accurate estimation of the exact time of the possible transmission cannot be made: clusters originating years ago may

be continuing to spread. We also lacked information about treatment and exposure history, and of the complete population to identify all index cases to differentiate between primary and acquired DR. Second, the prevalence and clustering of MDR-TB, pre-XDR and XDR-TB isolates in some provinces might be underestimated due to the low coverage of DST for the first-line drugs among TB cases [1].

CHAPTER IV
WHOLE-GENOME SEQUENCE ANALYSIS AND COMPARISONS BETWEEN DRUG-RESISTANCE MUTATIONS AND MINIMUM INHIBITORY CONCENTRATIONS OF *MYCOBACTERIUM TUBERCULOSIS* ISOLATES CAUSING MDR/XDR-TB

1. Introduction

Emergence of DR strains of *Mtb* remains the challenge for TB control. In 2018, the WHO estimated that there were 457,000 MDR-TB cases globally and that 8.5% of these were XDR-TB [50]. Early identification of TB and accurate DST are urgently required for appropriate TB treatment and to reduce the risk of further DR-TB development.

The gold standard of DST for *Mtb* is the proportional method [154]. The MIC test is another phenotypic method for quantification of the resistance level. Such phenotypic DSTs are time-consuming and laborious. Hence an alternative approach, genotypic DST, is becoming more readily accepted, provided that the complete database of mutations associated with drug resistance is available. WGS provides the best resolution of the genetic repertoire and is highly applicable for predicting drug-resistance profiles of *Mtb* and simultaneously can determine clustering for transmission analysis [155, 156]. There have been few direct comparisons of these three DST methods [36], especially for second-line drugs.

Quantitative phenotypic resistance (indicated by MIC values) associated with different mutations has been reported [36, 157, 158]. The current guidelines from WHO suggest that mutations detected in *Mtb* isolates can be used to predict resistance levels [36]. However, knowledge of such mutations is still limited in both number of tested strains and number of drugs available in the WHO database, and again especially for the second-line drugs [28].

Heteroresistance of *Mtb*, the mixture of susceptible and resistant strains in a single sample [37], has an effect on quantitative DSTs [38, 39]. A previous study compared different phenotypic DSTs to detect heteroresistance to RIF [38] and genotypic

approaches using WGS have also been described [39]. However, the few relevant studies have not made direct comparisons between genotypic heteroresistance (based on variant frequencies) and MIC levels for *Mtb*.

Thus, we compared DST profiles of a collection of MDR/XDR-TB *Mtb* isolates from Thailand, using phenotypic methods (agar proportion and MIC tests using MYCOTB) and a genotypic method (WGS analysis). The association between specific mutations and levels of drug resistance was analyzed for 11 drugs, including INH, RIF, EMB, STR, SLIDs: KAN and AMK, FQs: OFX and MFX, ETO, PAS and RFB. The possibility of genotypic heteroresistance, based on variant frequencies and quantitative MIC levels, was also investigated.

2. Materials and methods

2.1 *Mtb* isolates and setting

Sixty clinical *Mtb* isolates collected between 2003 and 2017 were obtained from stock cultures deposited at the Drug-Resistant Tuberculosis Research Fund, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand. The clinical specimens were stained for acid-fast bacilli using the Kinyoun method and subjected to phenotypic DST using the agar proportion method. Each selected isolate was from a different patient and each isolate was resistant at least to RIF (Poly-DR TB, n = 1; MDR, n = 28; Pre-XDR, n = 6 and XDR, n = 25). All isolates were sub-cultured on LJ media and incubated at 37°C for four to six weeks. Multiple loops of mycobacterial culture were used for genomic DNA extraction (using the CTAB method [133]) and for MIC-based phenotypic DSTs. This study was approved by the Khon Kaen University Ethics Committee in Human Research (Ethics number HE601249).

2.2 Phenotypic drug susceptibility testing

The standard agar proportion method was performed according to recommendations from the US Centers for Disease Control and Prevention [159]. Briefly, anti-TB drug discs were placed into the centers of individual quadrants of sterile plates, then 5.0 ml of Middlebrook 7H10 (Difco, Detroit, MI, USA) containing 10% oleic acid-albumin dextrose-catalase (BBL, Becton Dickinson, USA) was poured over the plate, and the agar was allowed to solidify overnight at room temperature. The

inoculum was prepared by suspending the *Mtb* colonies in Middlebrook 7H9 (Difco, Detroit, MI, USA) and adjusting the supernatant to turbidity equivalent to a MacFarland standard of one. The suspension was diluted to 10^{-2} and 10^{-4} [159]. The dilutions were inoculated onto the control quadrant, drug-free medium, and drug-containing quadrants. The plate was incubated at 37°C until colonies appeared on the control quadrant after approximately two to four weeks. Percentage of resistance was determined by (no. of colonies on drug-containing quadrant/no. of colonies on control quadrant)×100. An isolate was regarded as resistant when the percentage of resistance was ≥1%.

The MIC-based phenotypic DST was performed using Sensititre MYCOTBI (MYCOTB) plates according to the manufacturer (TREK Diagnostic Systems, West Sussex, United Kingdom). The wells of a MYCOTB plate contain 12 lyophilized anti-TB drugs with ranges of drug concentrations appropriate to each drug [27, 160]. Briefly, *Mtb* colonies were suspended in saline-Tween with glass beads for agitation and the turbidity of the supernatant adjusted to 0.5 MacFarland standard. This suspension (100 µl) was added into Middlebrook 7H9 medium and 100 µl of this mixture was added into each well of the MYCOTB plate. The plates were covered with plastic seals and incubated at 37°C. The plates were read using the Sensititre Vizion Digital MIC Viewing System (TREK Diagnostic Systems) at 10 days, or 21 days if poor growth was observed. The MIC was defined as the lowest concentration of anti-TB that inhibits visible growth.

The CCs used for agar proportion and MYCOTB assays are listed in Table 16. All isolates were tested once: If the test failed, it was repeated. *Mtb* H37Rv ATCC 27294 strain was used as a control for both agar proportion and MYCOTB assays.

2.3 Whole-genome sequencing and *in silico* detection of drug resistance

WGS was done for a subset (n = 27) of the 60 genomic DNA samples at the Genome Institute of Singapore, Singapore, using the TrueSeq DNA sample preparation kit (Illumina, San Diego, CA) and the MiSeq platform (Illumina) generating 250-bp paired-end reads, or using the NEBnext Ultra kit (Illumina, San Diego, CA) for the HiSeq (Illumina) platform generating 150-bp paired-end reads. The remaining 33 samples were sequenced at NovogeneAIT, N.T., Hong Kong, using the HiSeq (Illumina) platform generating 150-bp paired-end reads. The quality of sequence reads

was determined using FastQC version 0.11.7 [134]. The sequencing coverage and percentage of mapped reads against the reference genome of the H37Rv strain were determined using GATK version 3.4.0 [137] and SAMtools version 0.1.19 [161]. The mean genome coverage and the mean mapping rate were 224.5 (±152.4 standard deviation) and 97.9%, respectively. The WGS data are available in the Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) with the accession Nos. PRJNA598981 and PRJNA598949.

To detect drug resistance and determine *Mtb* lineage from the WGS data, raw fastq files were uploaded to an online tool, TB-Profiler version 2.8.6 [138]. To detect heteroresistant isolates, manual analysis was done to calculate frequencies of variants occurring in fewer than 100% of reads. Paired-end raw reads of each isolate were mapped to the *Mtb* H37Rv reference genome (GenBank accession number: NC_000962.3) using BWA-MEM version 0.7.12 [162]. SAMtools was used for SAM-BAM format conversion and sorting of mapped sequences. Local realignment of the mapped reads was performed using GATK. Variants, including single nucleotide polymorphisms (SNPs) and small indels, were called using GATK and SAMtool tools. Variant sites were filtered based on the following criteria: mapping quality >50 (-C in Samtools calling), base quality/base alignment quality >20 (-Q in Samtools calling), >10 reads or ≤2,000 reads (-d in Samtools filter) covering each site. To maximize specificity, the called variants were selected from the intersection of those identified by Samtools and GATK. For detection of heteroresistance, an in-house python script was used to extract the read frequencies supporting the mutations from the mapped reads. When read frequencies of mutant alleles were less than 99% compared to the wild-type background, we regarded this as WGS-based evidence of heteroresistance in that isolate [39]. In addition, the online tool, PhyresSE version 1.0 [163], was used for validation of drug resistance-conferring mutations obtained from TB-Profiler and for detection of heteroresistant TB.

Phylogenetic analysis of the 7,880 high-confidence SNPs identified among the 60 *Mtb* isolates was performed based on the maximum likelihood method with a general time-reversible and gamma distribution model (selected model based on data) using MEGA version 10.1 [164]. The phylogenetic tree was constructed based on 1,000

bootstrap replicates. The visualization of the phylogenetic tree was performed using iTOL [165].

2.4 Data analysis

For all analyzes and visualization, R (version 3.6.1) was used and p-values <0.05 were considered statistically significant. Sensitivity, specificity and categorical agreement with 95% CI were analyzed using the package epiR (version 1.0-4). CompareTests version 1.2 was used for comparisons between DST methods for each drug. Analyses for INH and RIF were not performed because few or no susceptible isolates were available. Also, analyses for RFB, PZA and DCS were not done due to lack of DST results for these from the agar proportion assay. Any association between MIC data and the drug resistance-conferring mutations was tested using the Wilcoxon rank-sum test. Graphs representing genetic information and their corresponding MICs were plotted using package ggplot2 version (3.2.1).

Table 16 Critical concentrations used in this study for phenotypic DST assays

Drug	Agar proportion (µg/ml) [166, 167],	MYCOTB (µg/ml) [27, 160].
Isoniazid	0.2	0.25[b]
Rifampicin	1	1
Ethambutol	5	4[b]
Streptomycin	2	2[b]
Kanamycin	6[a]	5[b]
Amikacin	6[a]	4[b]
Ofloxacin	2	2
Moxifloxacin	2	1[b]
Ethionamide	5	5
Para-aminosalicylic acid	2	1[b]
Rifabutin	-	0.5
D-cycloserine	-	32

The CCs that were different from those previously recommended by WHO. [a] Updated recommendations of CCs from WHO [166, 167], [b] Recommendations of CCs accompanying MYCOTB kit [27, 160].

3. Results

3.1 Characteristics of the studied isolates

The clinical *Mtb* isolates used were isolated from 60 TB patients. Most of the patients were male (79%). The average age was 43.6 years. Based on phylogenetic analysis, 88.3% (n = 53) of the isolates belonged to lineage 2 (East-Asian). There were two small clusters, each of two genetically identical isolates: only in one of these did the isolates share the same drug-resistance patterns (Figure 20).

3.2 Agreement of DST results between phenotypic and genotypic methods

Agreement, sensitivity and specificity among DST methods are shown (Table 1). High levels of agreement between the agar proportion and WGS-based DSTs were found for OFX (95%) and AMK (90%) (Table 17). Agreement between WGS-based DST and MYCOTB was high for all drugs except EMB (65%) and ETO (62%).

3.3 Comparison between WGS-based genotypic DST and MIC results for each drug

3.3.1 Rifampicin (RIF) and rifabutin (RFB)

The *rpoB* Ser450Leu mutation was commonly found (n = 36, 60%) among both RIF- and RFB-resistant isolates (Figure 21 and Supplementary Table 11). However, only RIF-resistant isolates showed distinct MIC values beyond the CC. Many RFB-resistant isolates (n = 18) with *rpoB* mutations (e.g. *rpoB* Asp435Val, Ser441Leu, Leu452Pro) had MIC values below the CC. Isolates with *rpoB* Ser450Leu and Asp435Phe exhibited RIF resistance but were RFB-susceptible according to the MIC test.

Mutations in *rpoB* Asp435Phe/Tyr/Val had MIC values for RFB significantly lower than isolates with *rpoB* Ser450Leu (0.12-1 µg/ml vs. 0.25-16 µg/ml, p = 0.002) (Figure 22). One heteroresistant isolate (79% of reads support *rpoB* Ser450Leu) had MIC of RIF lower than other isolates but had a MIC value below the CC of RFB (Fig 3 and S2 Table). An isolate with 64% reads of Ser441Leu was susceptible to RIF, whereas another isolate with the same mutation (in 96% of reads) was resistant to RIF (Figure 23). However, these two isolates were both susceptible to RFB.

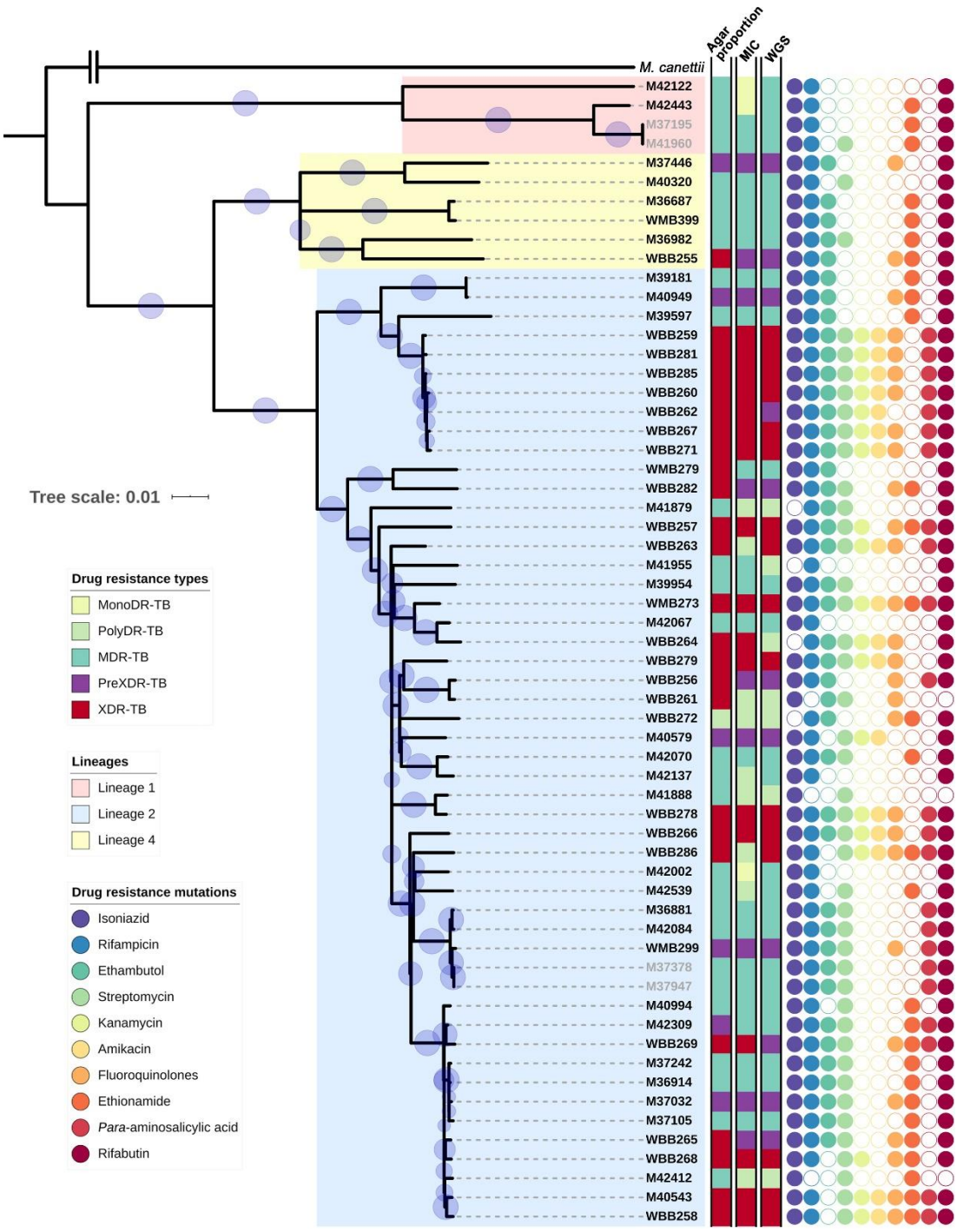

Figure 20 Phylogenetic analysis of 60 *Mtb* isolates. The phylogenetic tree was inferred using the maximum likelihood method with general time reversible and gamma distribution model using 7,880 high-confidence SNPs relative to the H37Rv reference genome. The bootstrap consensus tree was inferred from 1,000 replicates. Blue circles refer to bootstrap values and the size of each circle is proportional to its value (most of the bootstrap values are 100).

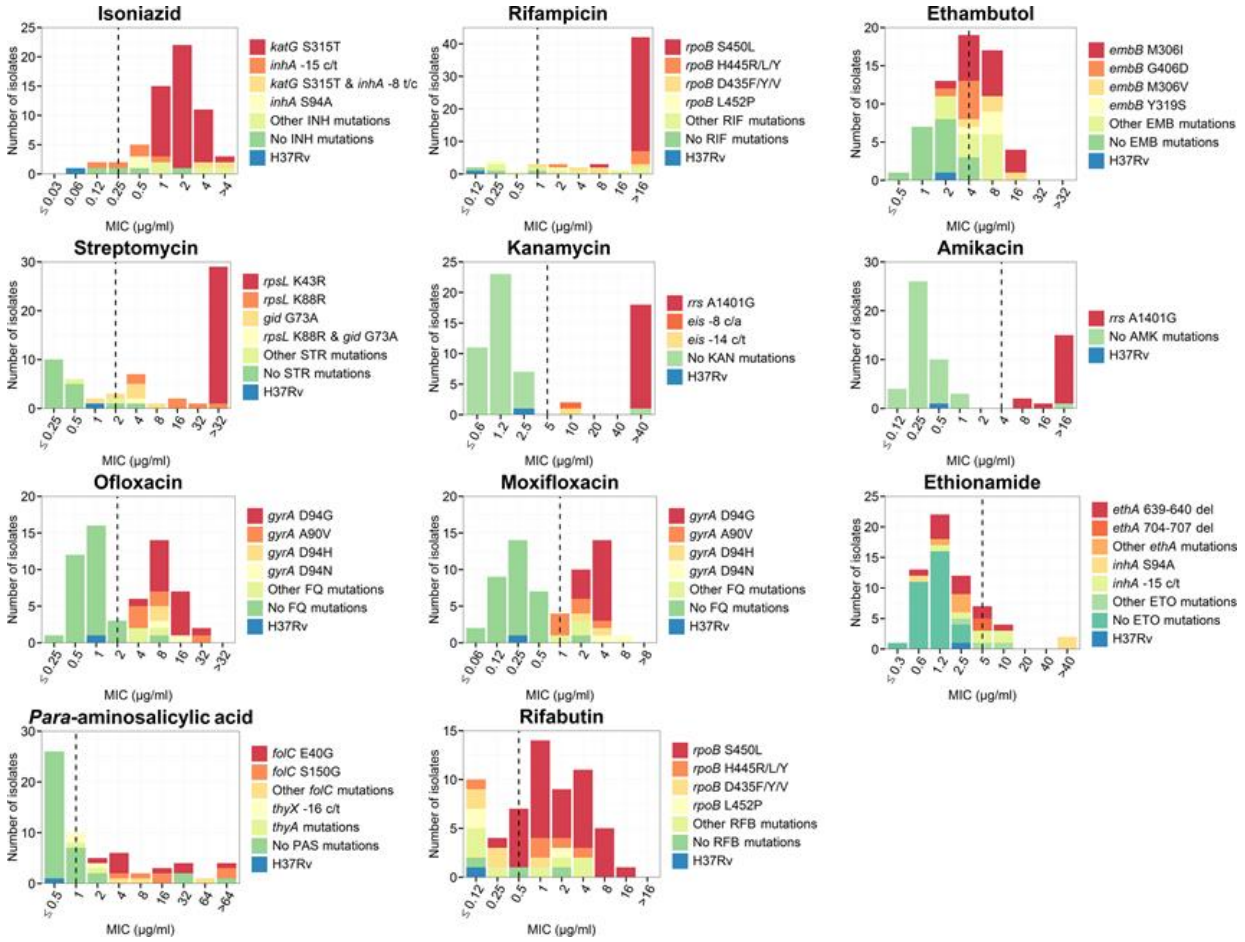

Figure 21 Distributions of drug resistance-conferring mutations with corresponding MIC values. Each stacked column represents a collection of isolates colored by different genetic background. The dashed lines indicate the critical concentrations used for MYCOTB. The H37Rv control strain was susceptible to all anti-tuberculosis drugs and represents the wild-type

Table 17 Agreement among phenotypic and genotypic DST assays

Drug	WGS	Agar prop.		WGS vs. Agar prop.			WGS	MYCOTB		WGS vs. MYCOTB		
		R	S	% Sensitivity (95% CI)	% Specificity (95% CI)	% Categorical agreement (95% CI)		R	S	% Sensitivity (95% CI)	% Specificity (95% CI)	% Categorical agreement (95% CI)
Isoniazid[a]	**R**	56	0	NA	NA	NA	**R**	54	2	NA	NA	NA
	S	3	1				**S**	2	2			
Rifampicin[a]	**R**	57	0	NA	NA	NA	**R**	51	6	NA	NA	NA
	S	3	0				**S**	0	3			
Ethambutol[b]	**R**	35	4	92 (78-97)	79 (55-92)	88 (77-94)	**R**	21	21	100 (NA)	46 (31-62)	65 (54-75)
	S	3	15				**S**	0	18			
Streptomycin	**R**	34	9	94 (80-99)	63 (42-79)	82 (71-89)	**R**	39	4	98 (84-100)	80 (57-92)	92 (82-96)
	S	2	15				**S**	1	16			
Kanamycin[b]	**R**	19	0	70 (50-85)	100 (NA)	86 (76-93)	**R**	19	0	95 (62-100)	100 (NA)	98 (79-100)
	S	8	32				**S**	1	40			
Amikacin	**R**	17	0	74 (52-88)	100 (NA)	90 (80-95)	**R**	17	0	94 (60-99)	100 (NA)	98 (79-100)
	S	6	37				**S**	1	42			
Ofloxacin	**R**	28	0	90 (72-97)	100 (NA)	95 (84-99)	**R**	28	0	97 (71-100)	100 (NA)	98 (79-100)
	S	3	29				**S**	1	31			
Moxifloxacin[b]	**R**	15	11	88 (63-97)	73 (58-84)	78 (65-86)	**R**	24	4	96 (76-99)	89 (73-96)	92 (82-96)
	S	2	30				**S**	1	31			
Ethionamide	**R**	23	6	92 (73-98)	83 (67-92)	87 (76-93)	**R**	6	23	100 (NA)	57 (44-70)	62 (49-73)
	S	2	29				**S**	0	31			
PAS	**R**	22	1	71 (53-84)	97 (79-100)	83 (73-90)	**R**	20	3	80 (60-91)	91 (77-97)	87 (76-93)
	S	9	28				**S**	5	32			

S, susceptible; R, resistant; Agar prop., agar proportion method; NA, not applicable; PAS, *para*-aminosalicylic acid. [a] The number of sensitive isolates based on agar proportion and MYCOTB (MIC-based DST) assays was too low (<10 isolates) to allow for reliable estimation of agreement, sensitivity and specificity. [b] DST results were available for all 60 isolates, except that results for ethambutol, kanamycin and moxifloxacin using agar proportion were only available for 57, 59 and 58 isolates respectively.

3.3.2 Isoniazid (INH) and ethionamide (ETO)

The most frequent mutation for INH resistance was *katG* Ser315Thr (n = 43, 72%) (Figure 21 and Supplementary Table 11). Most isolates with known INH mutations exhibited a MIC values above the CC, except for those harboring *inhA* promoter nutation alone. Two isolates without known INH resistance had MICs higher than the CC. Isolates with the -15 c/t *inhA* promoter mutation had MIC values for INH significantly lower than isolates with *katG* Ser315Thr (0.12-1 µg/ml vs. 1->4 µg/ml, p <0.001) (Figure 22). In addition, one INH-resistant isolate with 81% read frequency of the *katG* Ser315Thr mutation had an MIC value (1 µg/ml) lower than those with 99-100% reads of this mutation (range = 1-2 µg/ml) (Figure 23).

Most (23/29, 79%) isolates with known resistance mutations for ETO (*ethA* and *inhA* promoter) had MIC values lower than the CC (Figure 21 and Supplementary Table 11). Six isolates (21%) with known ETO-resistance mutations had MIC values above the CC and five of them had resistant DST results for both the agar proportion and the MIC tests.

3.3.3 Ethambutol (EMB) and streptomycin (STR)

Half of the isolates with EMB-resistance mutations (19 in *embB* and 2 in *embA*) had MIC values below the CC (Figure 21 and Supplementary Table 11). The agreement between WGS-based DST and MIC was increased from 65% to 85% when the CC was adjusted from 4 µg/ml to 2 µg/ml. The isolates with *embB* Gly406Asp had significantly lower MIC values for EMB compared to those with *embB* Met306Ile (2-4 µg/ml vs. 2-16 µg/ml, p = 0.031) (Figure 22). In addition, one isolate with 73% heteroresistance of *embB* Met306Ile exhibited an EMB-resistant phenotype with 16 µg/ml of MIC (Figure 23).

For STR, isolates with most common mutations (*rpsL* Lys43Arg and Lys88Arg) had MIC values above the CC (Figure 21 and Supplementary Table 11). However, half of the isolates with *gid* mutations had MIC values for STR lower than the CC. The isolates with *gid* Gly73Ala had MIC values for STR significantly lower than isolates with *rpsL* Lys43Arg (1-8 µg/ml vs. >32 µg/ml, p <0.001) (Figure 22). One isolate with *gid* Gly73Ala (100% reads) and 35% heteroresistance of *rpsL* Lys88Arg was resistant to EMB (Figure 23).

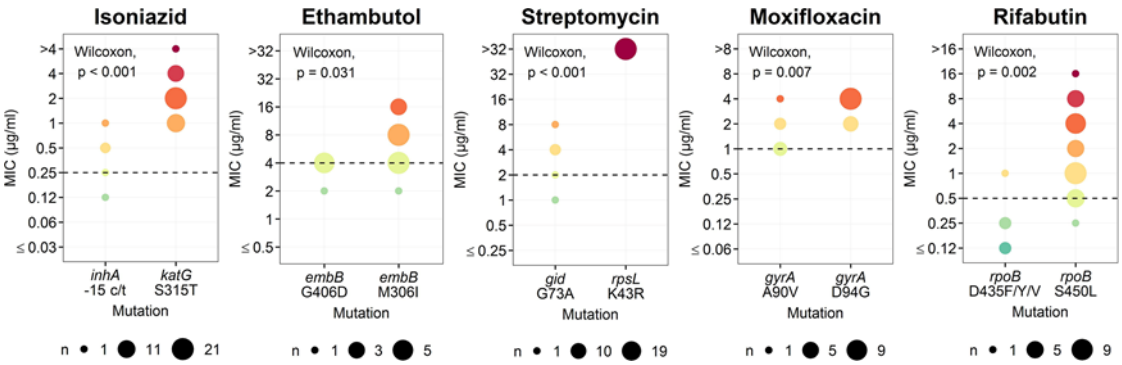

Figure 22 Comparisons between resistance-conferring mutations and MIC values of anti-TB drugs. Only those consensuses are showed for which common mutations are associated with significant differences in MIC levels. The dashed lines indicate the critical concentrations used for MYCOTB. The size of each circle is proportional to the number of isolates. The color of circles indicates the MIC level from low (blue-green) to high (red).

3.3.4 Kanamycin (KAN) and amikacin (AMK)

For KAN and AMK, all isolates (n = 19 and 17 for KAN and AMK, respectively) with known mutations had MIC values above the CC (Figure 21 and Supplementary Table 11). One isolate without known mutations for any of the SLIDs carried 85% reads of *rrs* A1401G (identified by in-house analysis) and this isolate was phenotypically resistant to both KAN and AMK (Figure 21, 23 and Supplementary Table 11). In contrast, another isolate carrying 12% reads of *rrs* A1401G had MIC values (1.2 µg/ml and 1 µg/ml for KAN and AMK, respectively) lower than other phenotypically KAN- and AMK-resistant isolates with high read frequencies for this mutation (KAN: 85-100% reads with MIC >40 µg/ml; AMK: 85-100% reads with MIC 8->16 µg/ml) (Figure 23).

3.3.5 Fluoroquinolones

All isolates with known *gryA* mutations were resistant to OFX but not MFX. Six isolates with *gyrA* Ala90Val had MIC values around the CC of MFX (Figure 21) that were significantly lower than isolates with *gyrA* Asp94Gly (1-4 µg/ml vs. 2-4 µg/ml, p = 0.007) (Figure 22). Discrepancy between WGS-based DST and MIC values for MXF was diminished when the CC was adjusted from 1 µg/ml to 0.5 µg/ml (Figure 21). One isolate which was genotypically wild-type (according to web-based

tools) but carrying heteroresistance detected by in-house analysis (78% and 22% reads of *gryA* Asp94Gly and Asp94Asn, respectively) was resistant to both OFX and MFX (Figure 21 and Figure 23). In addition, genotypic heteroresistance found in *gryA* mutations (Asp94Gly, Ala90Val and Asp94Asn) increased MIC values above the CC for OFX (Figure 23). In contrast, one isolate with 25% heteroresistance and five resistant isolates with 100% reads harboring *gryA* Ala90Val had MIC values at the borderline of the CC for MFX.

3.3.6 *Para*-aminosalicylic acid (PAS)

Most of the isolates with known mutations conferring PAS resistance, especially *folC*, had MIC values higher than the CC (Figure 21 and Supplementary Table 11). However, five isolates without known resistance mutations were resistant to PAS.

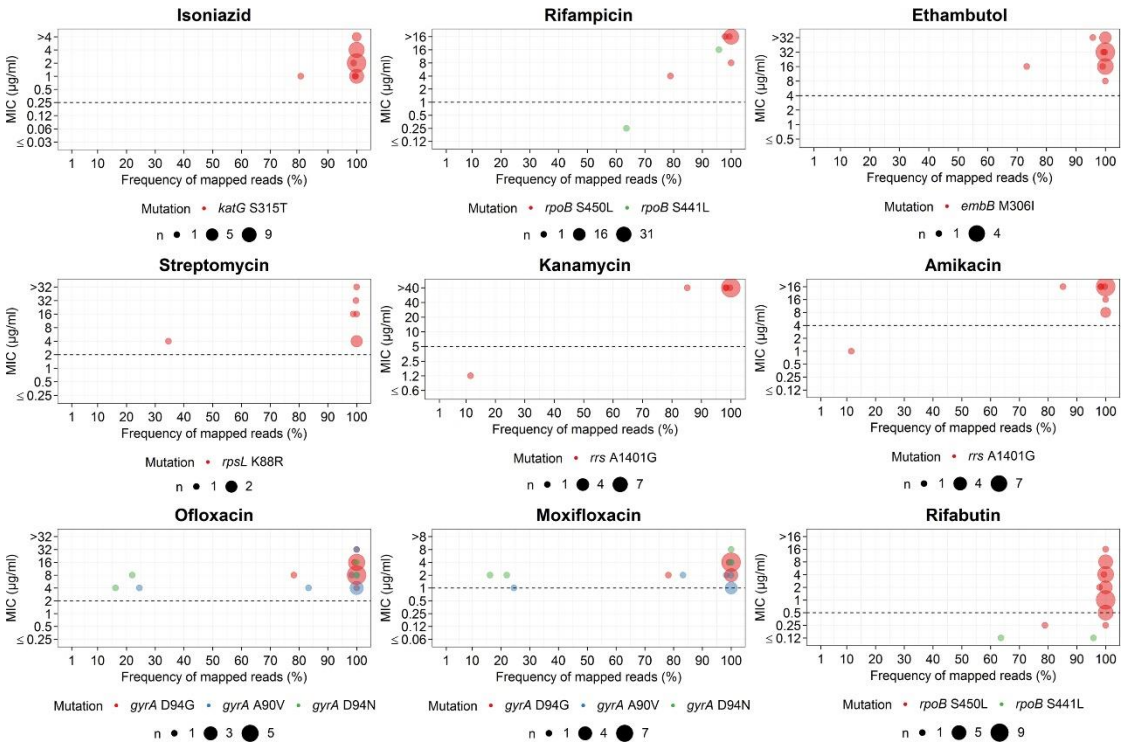

Figure 23 Comparison between heteroresistance (inferred from read frequencies of relevant SNPs) and MIC levels of *Mtb*. The dashed line indicates the critical concentrations used for MYCOTB. Only anti-TB drugs against which heteroresistance was inferred based on read frequencies are shown. The size of each circle is proportional to the number of isolates.

4. Discussion

We compared the DST patterns of MDR/XDR-TB isolates from Thailand using different DST methods including agar proportion tests, MYCOTB (MIC tests) and WGS analysis. Low levels of agreement among these methods were noted for some drugs, especially EMB and ETO. For EMB the agreement between WGS and MYCOTB was low (65%). Possibly the CC (4 μg/ml) used for EMB is too high [168]. When we reduced the CC of EMB to 2 μg/ml, the agreement between MYCOTB vs. WGS was greatly improved (85%). Adjustment of some CCs for MIC-based DSTs might be helpful to improve the agreement between MIC-based DSTs and other methods. For ETO, there was also poor agreement between MYCOTB and WGS methods (62%), but high agreement (87%) between agar proportion and WGS methods. Such discrepancies might be due to an inappropriate CC value for ETO and/or known resistance mutations in *ethA* and the *inhA* promoter might not be associated with ETO resistance in our cohort [169, 170]. Besides an inappropriate CC value and the potential effect of previously unknown mutations or overweighted mutations, the discrepancies between DST methods might also be caused by undetected laboratory error. Taken together, these results identify drugs for which sensitivity tests might be particularly difficult to interpret and the properties of particular DST methods that might contribute to this difficulty.

Although we used CC values close to those recommended by the WHO, genotypically resistant and genotypically susceptible *Mtb* isolates were found with MIC values either side of the CC for many drugs including EMB, ETO and RFB. For example, this applied to isolates with *embB* mutations using the CC value (4 μg/ml) suggested in the test kit instructions. When the WHO-recommended CC value (5 μg/ml) was applied, discordance between genotypic and phenotypic tests was even greater for EMB. Similarly, the agreement of EMB between phenotypic and genotypic DST was low [168]. For ETO, we found isolates that had resistance-conferring mutations in the *ethA* gene and the *inhA* promotor but had MIC values lower than the CC (5 μg/ml). Mutations in the *inhA* promotor confer only low resistance levels against INH [171], and likely also against ETO. For RFB, many isolates with *rpoB* mutations had MIC values both higher and lower than the CC. Although RIF and RFB belong to the same family of anti-TB drugs, the MIC distributions relative to CCs of isolates

harboring known *rpoB* mutations were not the same for both drugs. No wild-type isolates had an MIC above the CC (1 µg/ml) for RIF and few isolates with *rpoB* mutations fell below the CC. However, in the case of RFB, a greater proportion of isolates harboring *rpoB* mutations had MIC values lower than the CC (0.5 µg/ml as recommended by the kit instructions). Possibly, mutations (especially *rpoB* Asp435Val) assumed to confer resistance to RIF might not be highly correlated with RFB resistance, as found previously by others [170, 172, 173]. Furthermore, we found that isolates carrying *rpoB* Asp435Val alone had significantly lower MIC values for RFB than did isolates carrying *rpoB* Ser450Leu. Similarly, a previous study reported that *rpoB* Asp435Val alone had lower IC_{50} values for RIF and RFB than did isolates with *rpoB* Ser450Leu [173]. For STR, eight isolates with *gid* mutations had MIC values between 0.5 and 8 µg/ml, thus falling on and either side of the CC (2 µg/ml). The *gid* mutations have been determined as moderate-confidence mutations for STR resistance [21]. Possibly, mutations in *gid* confer low resistance levels. In the case of AMK and KAN, most isolates lacking specific mutations had MIC values below the CC, whereas MICs for isolates with resistance-conferring mutations fell above the CC. In addition, one isolate with no known mutations for SLIDs (tested by both *in silico* tools) exhibited heteroresistance of *rrs* A1401G (identified by in-house analysis) had MIC values for KAN and AMK above the CC. Conversely, many genotypic wild-type isolates with MIC values higher than CCs were found for several drugs, especially PAS. There are several explanations for this spectrum of results. First, not all mutations confer the same resistance level. The WHO suggested that some mutations confer low, some moderate and some high resistance-levels [36]. Isolates harboring low resistance-level mutations might have MIC values close to the CC. Second, mutation databases are incomplete, especially for the second-line drugs, which might explain why isolates without known resistance-conferring mutations had MIC values higher than the CC. In addition, other drug-resistance mechanisms such as epigenetic mechanisms cannot be identified by genetic analysis [174]. The efflux pump [175] mechanism might fall into this category. Furthermore, we noted that available *in silico* tools were unable to detect certain heteroresistance in *rrs* and *gryA* and gave a false genotypically susceptible result compared to our in-house analysis pipeline for particular drugs. The improvement of the drug-resistance mutation databases, web-based analysis tools and/or use of deep-

sequencing techniques [176] might enhance the sensitivity for identification of heteroresistance. Readjustment of CCs for problematic drugs such as EMB [167] and MXF [168, 170], might also help to overcome these problems.

There are previous reports of mutations in genes associated with low MIC levels for INH (*inhA* promoter: -15 c/t promoter [171]), EMB (*embB*: Gly406Asp and Met306Ile [177]), STR (*gidB* [178]), MFX (*gryA*: Asp94Ala [179]), and RFB (*rpoB*: Asp435Val and Asp435Tyr [170, 172, 173]). However, few of these studies had adequate sample sizes [171, 179]. We used multiple MDR/XDR-TB isolates to test for an association between MIC levels and mutations and found a significant association of the *inhA* promoter -15 c/t, *embB* Gly406Asp, *gid* mutations, *gryA* Ala90Val and *rpoB* Asp435Phe/Tyr/Val with low MIC levels spanning the CCs for INH, EMB, STR, MXF and RFB, respectively. However, the low number of resistance-conferring alleles found in our MDR/XDR-TB isolates limited our ability to investigate other drugs. The WHO database of mutations associated with resistance [36] is still limited in both number of isolates for each mutation and number of drugs. Our findings support the WHO database for known mutations associated with low-level resistance (INH resistance: -15 c/t *inhA* promoter and MFX resistance: *gryA* Ala90Val). In addition, our results suggest additional mutations associated with low vs. high resistance levels for EMB (*embB* Gly406Asp vs. *embB* Met306Ile), STR (*gid* Gly73Ala vs. *rpsL* Lys43Arg) and RFB (*rpoB* Asp435Phe/Tyr/Val vs. *rpoB* Ser450Leu). Further studies using a larger number of drug-resistant isolates will provide more insights into the association between particular mutations and MIC values.

Heteroresistance occurs when subpopulations within an isolate vary in their degree of resistance. Heteroresistance commonly arises during intermittent exposure to subtherapeutic drug levels, leading eventually to the generation of fully resistant populations [37]. Better understanding of the relationship between heteroresistance and MIC level should improve the effective treatment of TB [180], but has been the subject of few previous studies [38, 39]. *In-vitro* phenotypic experiments have demonstrated that low frequencies of *Mtb* cells harboring *rpoB* mutations within an isolate are associated with decreased MIC levels for RIF [38]. Only one study has reported a possible association between genotypic heteroresistance (based on WGS data) and MFX phenotypic heteroresistance [39]. In our study, we attempted to analyze the

association between genotypic heteroresistance based on the proportion of WGS mapped reads of resistance-conferring SNPs and MIC levels for nine drugs. Only RIF, KAN and AMK seemed to show a positive association between read frequencies of relevant mutations and MIC levels. However, the number of genotypically heteroresistant isolates available in our study was also too low for statistical analysis. Overall, our data do indicate a relationship between frequency of resistance-conferring alleles and MIC values in heteroresistant isolates of *Mtb*. This further suggests the considerable applicability of WGS to characterize drug-resistant TB. However, these findings are preliminary, indicating the need for further study with higher sample sizes and systematic analysis.

We found that the WGS method was in good agreement with the MYCOTB system and, for most drugs, in good agreement with the agar proportion test. Although the agar proportion method is still the "gold standard" DST for new drugs for which resistance-conferring mutations are not represented in databases, this method is extremely laborious and time consuming [181]. Similarly, although MIC-based tests can quantify resistance levels, the effort and time required remain obstacles to routine use [181]. The WGS method can shorten the turnaround time, especially when analyzed directly from the samples, and also provides the clustering information needed for epidemiological management [182]. The WGS method provides high-resolution information regarding drug susceptibility and level of resistance. However, a complete database of relevant mutations for each drug and the association of each mutation with resistance level is needed. Our study has contributed part of this information and reinforces the applicability of the WGS method for DST.

Other limitations of our study should be noted. We included a collection of drug-resistant isolates from TB patients in Thailand, including MDR-TB, Pre-XDR-TB and XDR-TB cases. We used these to highlight the effect of drug resistance-conferring mutations on quantitative DSTs for both first-line and second-line anti-TB drugs, except for PZA. PZA is difficult to to use in an agar-based DST because it requires acidity of the culture medium for drug activity [183] and this drug was not included in the MYCOTB MIC plate. Hence, we could not determine the interrelation between phenotypic DST of this drug and likely PZA resistance-conferring mutations which were identified in 26 (43%) isolates. A phylogenetic tree based on whole-genome

variants was inferred to ensure that potentially clonal strains did not affect the association analysis. Although there were two small clusters (each including two isolates) of genetically identical *Mtb* isolates among our samples, only one pair of isolates shared the same drug resistance pattern. Hence, the association results were not confounded by the presence of clonal strains. The diversity of resistance-conferring mutations is generally lower in MDR-TB isolates than in mono- or poly-resistant isolates [20, 184]. Most of our isolates were MDR-TB, Pre-XDR-TB and XDR-TB, which could affect the mutation frequencies and sensitivity comparison between DST methods. The database from TB-Profiler includes some mutations for which there is only a low level of confidence, based on current knowledge, that they are actually associated with resistance. Examples of these are *ethyA* associated with ETO resistance and *eis* promoter -8 c/a associated with KAN resistance). Low-confidence mutations might affect the ability of the WGS method to detect DR and heteroresistance.

CHAPTER V
CONCLUSION

This study has demonstrated the usefulness of WGS for DR-TB epidemiology. It was found that close to half of MDR-TB, pre-XDR-TB and XDR-TB cases in Thailand might be due to transmission clusters. Two-thirds of pre-XDR and three-fourths of MDR-TB and XDR-TB clustering isolates were possible examples of primary resistance. These results indicate that the emergence of MDR-TB, pre-XDR and XDR-TB cases in Thailand might be from a narrow base of ancestral strains. The high prevalence of MDR/XDR-TB in Thailand might be due to multi-clonal outbreaks. People living in the western region of Thailand had a 2.4-fold increased risk of DR-TB clusters. Lineage 2.2.1 conferred a 3.6-fold increased risk of forming DR-TB clusters relative to other lineages.

The comparison of the agreement between phenotypic (agar proportion method and MIC tests using MYCOTB) and genotypic DSTs (WGS) and highlighted problematic drugs, especially EMB and ETO, that can yield different results according to the DST method used. Additional information was provided regarding mutations associated with low vs. high resistance levels against INH (-15 c/t *inhA* promoter vs. *katG* Ser315Thr), EMB (*embB* Gly406Asp vs. *embB* Met306Ile), STR (*gid* Gly73Ala vs. *rpsL* Lys43Arg), MFX (*gyrA* Ala90Val vs. *gyrA* Asp94Gly) and RFB (*rpoB* Asp435Phe/Tyr/Val vs. *rpoB* Ser450Leu), but further evaluation with a larger sample size is required. A possible association between genotypic heteroresistance and MIC level was also suggested. These results emphasize the high applicability of WGS for TB diagnosis including determination of drug resistance, mutated allele association with MIC and heteroresistance associated with MIC.

In conclusion, our study revealed several applications of using WGS for DR-TB epidemiology, tracking transmission of DR-TB clusters and prediction of DR-TB which provide significant information for better management of DR-TB in Thailand.

REFERENCES

1. World Health Organization. **Global tuberculosis report 2019.** Geneva, Switzerland: WHO; 2019.
2. Lee RS, Behr MA. The implications of whole-genome sequencing in the control of tuberculosis. **Ther Adv Infect Dis** 2016; 3(2): 47-62.
3. Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, et al. Whole genome sequencing of *Mycobacterium tuberculosis*: current standards and open issues. **Nat Rev Microbiol** 2019; 17(9): 533-545.
4. van Embden JD, Cave MD, Crawford JT, Dale JW, Eisenach KD, Gicquel B, et al. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. **J Clin Microbiol** 1993; 31(2): 406-409.
5. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. **J Clin Microbiol** 1997; 35(4): 907-914.
6. Supply P, Allix C, Lesjean S, Cardoso-Oelemann M, Rusch-Gerdes S, Willery E, et al. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of *Mycobacterium tuberculosis*. **J Clin Microbiol** 2006; 44(12): 4498-4510.
7. Jiraphongsa C, Wangteeraprasert T, Henpraserttae N, Sanguanwongse N, Panya L, Sukkasitvanichkul J, et al. Community outbreak of multidrug resistance tuberculosis, Kanchanaburi province, Thailand on 2002-June 2010. **J Preventive Med Assoc Thail** 2011; 1(3): 261-271.
8. Regmi SM, Chaiprasert A, Kulawonganunchai S, Tongsima S, Coker OO, Prammananan T, et al. Whole genome sequence analysis of multidrug-resistant *Mycobacterium tuberculosis* Beijing isolates from an outbreak in Thailand. **Mol Genet Genomics** 2015; 290(5): 1933-1941.

9. Somoskovi A, Helbling P, Deggim V, Homke R, Ritter C, Bottger EC. Transmission of multidrug-resistant tuberculosis in a low-incidence setting, Switzerland, 2006 to 2012. **Euro Surveill** 2014; 19(11).
10. Anderson LF, Tamne S, Brown T, Watson JP, Mullarkey C, Zenner D, et al. Transmission of multidrug-resistant tuberculosis in the UK: a cross-sectional molecular and epidemiological study of clustering and contact tracing. **Lancet Infect Dis** 2014; 14(5): 406-415.
11. Jagielski T, Brzostek A, van Belkum A, Dziadek J, Augustynowicz-Kopec E, Zwolska Z. A close-up on the epidemiology and transmission of multidrug-resistant tuberculosis in Poland. **Eur J Clin Microbiol Infect Dis** 2015; 34(1): 41-53.
12. Wang SF, Zhou Y, Pang Y, Zheng HW, Zhao YL. Prevalence and Risk Factors of Primary Drug-Resistant Tuberculosis in China. **Biomed Environ Sci** 2016; 29(2): 91-98.
13. Al-Ghafli H, Kohl TA, Merker M, Varghese B, Halees A, Niemann S, et al. Drug-resistance profiling and transmission dynamics of multidrug-resistant *Mycobacterium tuberculosis* in Saudi Arabia revealed by whole genome sequencing. **Infect Drug Resist** 2018; 11: 2219-2229.
14. Oliveira O, Gaio R, Carvalho C, Correia-Neves M, Duarte R, Rito T. A nationwide study of multidrug-resistant tuberculosis in Portugal 2014-2017 using epidemiological and molecular clustering analyses. **BMC Infect Dis** 2019; 19(1): 567.
15. Palittapongarnpim P, Luangsook P, Tansuphaswadikul S, Chuchottaworn C, Prachaktam R, Sathapatayavongs B. Restriction fragment length polymorphism study of *Mycobacterium tuberculosis* in Thailand using IS6110 as probe. **Int J Tuberc Lung Dis** 1997; 1(4): 370-376.
16. Rienthong D, Ajawatanawong P, Rienthong S, Smithtikarn S, Akarasewi P, Chaiprasert A, et al. Restriction fragment length polymorphism study of nationwide samples of *Mycobacterium tuberculosis* in Thailand, 1997-1998. **Int J Tuberc Lung Dis** 2005; 9(5): 576-581.

17. Faksri K, Drobniewski F, Nikolayevskyy V, Brown T, Prammananan T, Palittapongarnpim P, et al. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family based on IS6110, SNP, LSP and VNTR profiles from Thailand. **Infect Genet Evol** 2011; 11(5): 1142-1149.
18. Sukkasem S, Yanai H, Mahasirimongkol S, Yamada N, Rienthong D, Palittapongarnpim P, et al. Drug resistance and IS6110-RFLP patterns of *Mycobacterium tuberculosis* in patients with recurrent tuberculosis in northern Thailand. **Microbiol Immunol** 2013; 57(1): 21-29.
19. Canetti G, Fox W, Khomenko A, Mahler HT, Menon NK, Mitchison DA, et al. Advances in techniques of testing mycobacterial drug sensitivity, and the use of sensitivity tests in tuberculosis control programmes. **Bulletin of the World Health Organization** 1969; 41(1): 21-43.
20. Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, et al. Whole-genome sequencing for prediction of *Mycobacterium tuberculosis* drug susceptibility and resistance: a retrospective cohort study. **Lancet Infect Dis** 2015; 15(10): 1193-1202.
21. Miotto P, Tessema B, Tagliani E, Chindelevitch L, Starks AM, Emerson C, et al. A standardised method for interpreting the association between mutations and phenotypic drug resistance in *Mycobacterium tuberculosis*. **Eur Respir J** 2017; 50(6).
22. Pandey P, Pant ND, Rijal KR, Shrestha B, Kattel S, Banjara MR, et al. Diagnostic Accuracy of GeneXpert MTB/RIF Assay in Comparison to Conventional Drug Susceptibility Testing Method for the Diagnosis of Multidrug-Resistant Tuberculosis. **PLoS One** 2017; 12(1): e0169798.
23. Bai Y, Wang Y, Shao C, Hao Y, Jin Y. GenoType MTBDRplus Assay for Rapid Detection of Multidrug Resistance in *Mycobacterium tuberculosis*: A Meta-Analysis. **PLoS One** 2016; 11(3): e0150321.
24. Hillemann D, Rusch-Gerdes S, Richter E. Feasibility of the GenoType MTBDRsl assay for fluoroquinolone, amikacin-capreomycin, and ethambutol resistance testing of *Mycobacterium tuberculosis* strains and clinical specimens. **J Clin Microbiol** 2009; 47(6): 1767-1772.

25. Witney AA, Gould KA, Arnold A, Coleman D, Delgado R, Dhillon J, et al. Clinical application of whole-genome sequencing to inform treatment for multidrug-resistant tuberculosis cases. **J Clin Microbiol** 2015; 53(5): 1473-1483.
26. Abuali MM, Katariwala R, LaBombardi VJ. A comparison of the Sensititre(R) MYCOTB panel and the agar proportion method for the susceptibility testing of *Mycobacterium tuberculosis*. **Eur J Clin Microbiol Infect Dis** 2012; 31(5): 835-839.
27. Hall L, Jude KP, Clark SL, Dionne K, Merson R, Boyer A, et al. Evaluation of the Sensititre MycoTB plate for susceptibility testing of the *Mycobacterium tuberculosis* complex against first- and second-line agents. **J Clin Microbiol** 2012; 50(11): 3732-3734.
28. Faksri K, Kaewprasert O, Ong RT, Suriyaphol P, Prammananan T, Teo YY, et al. Comparisons of whole-genome sequencing and phenotypic drug susceptibility testing for *Mycobacterium tuberculosis* causing MDR-TB and XDR-TB in Thailand. **Int J Antimicrob Agents** 2019; 54(2): 109-116.
29. Huitric E, Werngren J, Jureen P, Hoffner S. Resistance levels and rpoB gene mutations among in vitro-selected rifampin-resistant *Mycobacterium tuberculosis* mutants. **Antimicrob Agents Chemother** 2006; 50(8): 2860-2862.
30. Dalla Costa ER, Ribeiro MO, Silva MS, Arnold LS, Rostirolla DC, Cafrune PI, et al. Correlations of mutations in katG, oxyR-ahpC and inhA genes and in vitro susceptibility in *Mycobacterium tuberculosis* clinical strains segregated by spoligotype families from tuberculosis prevalent countries in South America. **BMC Microbiol** 2009; 9: 39.
31. Imperiale BR, Zumarraga MJ, Di Giulio AB, Cataldi AA, Morcillo NS. Molecular and phenotypic characterisation of *Mycobacterium tuberculosis* resistant to anti-tuberculosis drugs. **Int J Tuberc Lung Dis** 2013; 17(8): 1088-1093.
32. Schon T, Jureen P, Chryssanthou E, Giske CG, Kahlmeter G, Hoffner S, et al. Rifampicin-resistant and rifabutin-susceptible *Mycobacterium tuberculosis* strains: a breakpoint artefact? **J Antimicrob Chemother** 2013; 68(9): 2074-2077.

33. Ocheretina O, Escuyer VE, Mabou MM, Royal-Mardi G, Collins S, Vilbrun SC, et al. Correlation between genotypic and phenotypic testing for resistance to rifampin in *Mycobacterium tuberculosis* clinical isolates in Haiti: investigation of cases with discrepant susceptibility results. **PLoS One** 2014; 9(3): e90569.
34. Nosova EY, Zimenkov DV, Khakhalina AA, Isakova AI, Krylova LY, Makarova MV, et al. A Comparison of the Sensititre MycoTB Plate, the Bactec MGIT 960, and a Microarray-Based Molecular Assay for the Detection of Drug Resistance in Clinical *Mycobacterium tuberculosis* Isolates in Moscow, Russia. **PLoS One** 2016; 11(11): e0167093.
35. Rukasha I, Said HM, Omar SV, Koornhof H, Dreyer AW, Musekiwa A, et al. Correlation of rpoB Mutations with Minimal Inhibitory Concentration of Rifampin and Rifabutin in *Mycobacterium tuberculosis* in an HIV/AIDS Endemic Setting, South Africa. **Front Microbiol** 2016; 7: 1947.
36. World Health Organization. The use of next-generation sequencing technologies for the detection of mutations associated with drug resistance in *Mycobacterium tuberculosis* complex: technical guide. Geneva, Switzerland: WHO; 2018.
37. Muller B, Borrell S, Rose G, Gagneux S. The heterogeneous evolution of multidrug-resistant *Mycobacterium tuberculosis*. **Trends Genet** 2013; 29(3): 160-169.
38. Zhang Z, Wang Y, Pang Y, Liu C. Comparison of different drug susceptibility test methods to detect rifampin heteroresistance in *Mycobacterium tuberculosis*. **Antimicrob Agents Chemother** 2014; 58(9): 5632-5635.
39. Operario DJ, Koeppel AF, Turner SD, Bao Y, Pholwat S, Banu S, et al. Prevalence and extent of heteroresistance by next generation sequencing of multidrug-resistant tuberculosis. **PLoS One** 2017; 12(5): e0176522.
40. Lawn SD, Zumla AI. Tuberculosis. **Lancet** 2011; 378(9785): 57-72.
41. Forrellad MA, Klepp LI, Gioffre A, Sabio y Garcia J, Morbidoni HR, de la Paz Santangelo M, et al. Virulence factors of the *Mycobacterium tuberculosis* complex. **Virulence** 2013; 4(1): 3-66.
42. Gagneux S, Small PM. Global phylogeography of *Mycobacterium tuberculosis* and implications for tuberculosis product development. **Lancet Infect Dis** 2007; 7(5): 328-337.

43. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, et al. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. **Nature** 1998; 393(6685): 537-544.
44. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. **Nat Rev Dis Primers** 2016; 2: 16076.
45. Dheda K, Gumbo T, Gandhi NR, Murray M, Theron G, Udwadia Z, et al. Global control of tuberculosis: from extensively drug-resistant to untreatable tuberculosis. **Lancet Respir Med** 2014; 2(4): 321-338.
46. Almeida Da Silva PE, Palomino JC. Molecular basis and mechanisms of drug resistance in *Mycobacterium tuberculosis*: classical and new drugs. **J Antimicrob Chemother** 2011; 66(7): 1417-1430.
47. Zhang Y, Yew WW. Mechanisms of drug resistance in *Mycobacterium tuberculosis*: update 2015. **Int J Tuberc Lung Dis** 2015; 19(11): 1276-1289.
48. Dominguez J, Boettger EC, Cirillo D, Cobelens F, Eisenach KD, Gagneux S, et al. Clinical implications of molecular drug resistance testing for *Mycobacterium tuberculosis*: a TBNET/RESIST-TB consensus statement. **Int J Tuberc Lung Dis** 2016; 20(1): 24-42.
49. World Health Organization. **WHO consolidated guidelines on tuberculosis, Module 4: treatment - drug-resistant tuberculosis treatment.** Geneva, Switzerland: WHO; 2020.
50. World Health Organization. **Global tuberculosis report 2018.** Geneva, Switzerland: WHO; 2018.
51. World Health Organization. **Global tuberculosis report 2014.** Geneva, Switzerland: WHO; 2014.
52. World Health Organization. **Global tuberculosis report 2015.** Geneva, Switzerland: WHO; 2015.
53. World Health Organization. **Global tuberculosis report 2016.** Geneva, Switzerland: WHO; 2016.
54. World Health Organization. **Global tuberculosis report 2017.** Geneva, Switzerland: WHO; 2017.

55. Nasiri MJ, Haeili M, Ghazi M, Goudarzi H, Pormohammad A, Imani Fooladi AA, et al. New Insights in to the Intrinsic and Acquired Drug Resistance Mechanisms in Mycobacteria. **Front Microbiol** 2017; 8: 681.
56. Malone KM, Gordon SV. Antibiotic Methylation: A New Mechanism of Antimicrobial Resistance. **Trends Microbiol** 2016; 24(10): 771-772.
57. Zhang Y, Yew WW. Mechanisms of drug resistance in *Mycobacterium tuberculosis*. **Int J Tuberc Lung Dis** 2009; 13(11): 1320-1330.
58. Muller B, Borrell S, Rose G, Gagneux S. The heterogeneous evolution of multidrug-resistant *Mycobacterium tuberculosis*. **Trends in Genetics** 2013; 29(3): 160-169.
59. Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. **Nature Genetics** 2014; 46(3): 279-286.
60. de Vos M, Muller B, Borrell S, Black PA, van Helden PD, Warren RM, et al. Putative compensatory mutations in the rpoC gene of rifampin-resistant *Mycobacterium tuberculosis* are associated with ongoing transmission. **Antimicrob Agents Chemother** 2013; 57(2): 827-832.
61. Borrell S, Gagneux S. Strain diversity, epistasis and the evolution of drug resistance in *Mycobacterium tuberculosis*. **Clin Microbiol Infect** 2011; 17(6): 815-820.
62. Trauner A, Borrell S, Reither K, Gagneux S. Evolution of drug resistance in tuberculosis: recent progress and implications for diagnosis and therapy. **Drugs** 2014; 74(10): 1063-1072.
63. Hatherell HA, Colijn C, Stagg HR, Jackson C, Winter JR, Abubakar I. Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review. **BMC Medicine** 2016; 14: 21.
64. Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, et al. A large genome center's improvements to the Illumina sequencing system. **Nat Methods** 2008; 5(12): 1005-1010.
65. Mardis ER. Next-generation sequencing platforms. **Annu Rev Anal Chem (Palo Alto Calif)** 2013; 6: 287-303.

66. Kohl TA, Utpatel C, Schleusener V, De Filippo MR, Beckert P, Cirillo DM, et al. MTBseq: a comprehensive pipeline for whole genome sequence analysis of *Mycobacterium tuberculosis* complex isolates. **PeerJ** 2018; 6: e5895.
67. Ezewudo M, Borens A, Chiner-Oms A, Miotto P, Chindelevitch L, Starks AM, et al. Integrating standardized whole genome sequence analysis with a global *Mycobacterium tuberculosis* antibiotic resistance knowledgebase. **Sci Rep** 2018; 8(1): 15382.
68. Phelan JE, Coll F, Bergval I, Anthony RM, Warren R, Sampson SL, et al. Recombination in pe/ppe genes contributes to genetic variation in *Mycobacterium tuberculosis* lineages. **BMC Genomics** 2016; 17: 151.
69. Galagan JE. Genomic insights into tuberculosis. **Nat Rev Genet** 2014; 15(5): 307-320.
70. Dixit A, Freschi L, Vargas R, Calderon R, Sacchettini J, Drobniewski F, et al. Whole genome sequencing identifies bacterial factors affecting transmission of multidrug-resistant tuberculosis in a high-prevalence setting. **Sci Rep** 2019; 9(1): 5602.
71. Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. **Lancet Infect Dis** 2013; 13(2): 137-146.
72. Kohl TA, Harmsen D, Rothganger J, Walker T, Diel R, Niemann S. Harmonized Genome Wide Typing of Tubercle Bacilli Using a Web-Based Gene-By-Gene Nomenclature System. **EBioMedicine** 2018; 34: 131-138.
73. Merker M, Kohl TA, Niemann S, Supply P. The Evolution of Strain Typing in the *Mycobacterium tuberculosis* Complex. **Adv Exp Med Biol** 2017; 1019: 43-78.
74. Meehan CJ, Moris P, Kohl TA, Pecerska J, Akter S, Merker M, et al. The relationship between transmission time and clustering methods in *Mycobacterium tuberculosis* epidemiology. **EBioMedicine** 2018; 37: 410-416.
75. Schito M, Dolinger DL. A Collaborative Approach for "ReSeq-ing" *Mycobacterium tuberculosis* Drug Resistance: Convergence for Drug and Diagnostic Developers. **EBioMedicine** 2015; 2(10): 1262-1265.

76. Starks AM, Aviles E, Cirillo DM, Denkinger CM, Dolinger DL, Emerson C, et al. Collaborative Effort for a Centralized Worldwide Tuberculosis Relational Sequencing Data Platform. **Clin Infect Dis** 2015; 61Suppl 3: S141-146.
77. Allix-Beguec C, Arandjelovic I, Bi L, Beckert P, Bonnet M, Bradley P, et al. Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing. **N Engl J Med** 2018; 379(15): 1403-1415.
78. Faksri K, Tan JH, Chaiprasert A, Teo YY, Ong RT. Bioinformatics tools and databases for whole genome sequence analysis of *Mycobacterium tuberculosis*. **Infect Genet Evol** 2016; 45: 359-368.
79. Niemann S, Koser CU, Gagneux S, Plinke C, Homolka S, Bignell H, et al., cartographers. Genomic diversity among drug sensitive and multidrug resistant isolates of *Mycobacterium tuberculosis* with identical DNA fingerprints2009.
80. Ioerger TR, Feng Y, Chen X, Dobos KM, Victor TC, Streicher EM, et al. The non-clonality of drug resistance in Beijing-genotype isolates of *Mycobacterium tuberculosis* from the Western Cape of South Africa. **BMC Genomics** 2010; 11: 670.
81. Casali N, Nikolayevskyy V, Balabanova Y, Ignatyeva O, Kontsevaya I, Harris SR, et al. Microevolution of extensively drug-resistant tuberculosis in Russia. **Genome Res** 2012; 22(4): 735-745.
82. Torok ME, Reuter S, Bryant J, Koser CU, Stinchcombe SV, Nazareth B, et al. Rapid whole-genome sequencing for investigation of a suspected tuberculosis outbreak. **J Clin Microbiol** 2013; 51(2): 611-614.
83. Lanzas F, Karakousis PC, Sacchettini JC, Ioerger TR. Multidrug-resistant tuberculosis in panama is driven by clonal expansion of a multidrug-resistant *Mycobacterium tuberculosis* strain related to the KZN extensively drug-resistant *M. tuberculosis* strain from South Africa. **J Clin Microbiol** 2013; 51(10): 3277-3285.
84. Clark TG, Mallard K, Coll F, Preston M, Assefa S, Harris D, et al. Elucidating emergence and transmission of multidrug-resistant tuberculosis in treatment experienced patients by whole genome sequencing. **PLoS One** 2013; 8(12): e83012.

85. Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, et al. Evolution and transmission of drug-resistant tuberculosis in a Russian population. **Nat Genet** 2014; 46(3): 279-286.
86. Guerra-Assuncao JA, Crampin AC, Houben RM, Mzembe T, Mallard K, Coll F, et al. Large-scale whole genome sequencing of M. tuberculosis provides insights into transmission in a high prevalence area. **Elife** 2015; 4.
87. Saelens JW, Lau-Bonilla D, Moller A, Medina N, Guzman B, Calderon M, et al. Whole genome sequencing identifies circulating Beijing-lineage *Mycobacterium tuberculosis* strains in Guatemala and an associated urban outbreak. **Tuberculosis (Edinb)** 2015; 95(6): 810-816.
88. Winglee K, Manson McGuire A, Maiga M, Abeel T, Shea T, Desjardins CA, et al. Whole Genome Sequencing of Mycobacterium africanum Strains from Mali Provides Insights into the Mechanisms of Geographic Restriction. **PLoS Negl Trop Dis** 2016; 10(1): e0004332.
89. Ssengooba W, Meehan CJ, Lukoye D, Kasule GW, Musisi K, Joloba ML, et al. Whole genome sequencing to complement tuberculosis drug resistance surveys in Uganda. **Infect Genet Evol** 2016; 40: 8-16.
90. Arnold A, Witney AA, Vergnano S, Roche A, Cosgrove CA, Houston A, et al. XDR-TB transmission in London: Case management and contact tracing investigation assisted by early whole genome sequencing. **J Infect** 2016; 73(3): 210-218.
91. Casali N, Broda A, Harris SR, Parkhill J, Brown T, Drobniewski F. Whole Genome Sequence Analysis of a Large Isoniazid-Resistant Tuberculosis Outbreak in London: A Retrospective Observational Study. **PLoS Med** 2016; 13(10): e1002137.
92. Wollenberg KR, Desjardins CA, Zalutskaya A, Slodovnikova V, Oler AJ, Quinones M, et al. Whole-Genome Sequencing of *Mycobacterium tuberculosis* Provides Insight into the Evolution and Genetic Composition of Drug-Resistant Tuberculosis in Belarus. **J Clin Microbiol** 2017; 55(2): 457-469.

93. Yang C, Luo T, Shen X, Wu J, Gan M, Xu P, et al. Transmission of multidrug-resistant *Mycobacterium tuberculosis* in Shanghai, China: a retrospective observational study using whole-genome sequencing and epidemiological investigation. **Lancet Infect Dis** 2017; 17(3): 275-284.
94. Shah NS, Auld SC, Brust JC, Mathema B, Ismail N, Moodley P, et al. Transmission of Extensively Drug-Resistant Tuberculosis in South Africa. **N Engl J Med** 2017; 376(3): 243-253.
95. Dheda K, Limberis JD, Pietersen E, Phelan J, Esmail A, Lesosky M, et al. Outcomes, infectiousness, and transmission dynamics of patients with extensively drug-resistant tuberculosis and home-discharged patients with programmatically incurable tuberculosis: a prospective cohort study. **Lancet Respir Med** 2017; 5(4): 269-281.
96. Manson AL, Abeel T, Galagan JE, Sundaramurthi JC, Salazar A, Gehrmann T, et al. *Mycobacterium tuberculosis* Whole Genome Sequences From Southern India Suggest Novel Resistance Mechanisms and the Need for Region-Specific Diagnostics. **Clin Infect Dis** 2017; 64(11): 1494-1501.
97. Rodriguez-Castillo JG, Pino C, Nino LF, Rozo JC, Llerena-Polo C, Parra-Lopez CA, et al. Comparative genomic analysis of *Mycobacterium tuberculosis* Beijing-like strains revealed specific genetic variations associated with virulence and drug resistance. **Infect Genet Evol** 2017; 54: 314-323.
98. Senghore M, Otu J, Witney A, Gehre F, Doughty EL, Kay GL, et al. Whole-genome sequencing illuminates the evolution and spread of multidrug-resistant tuberculosis in Southwest Nigeria. **PLoS One** 2017; 12(9): e0184510.
99. Chatterjee A, Nilgiriwala K, Saranath D, Rodrigues C, Mistry N. Whole genome sequencing of clinical strains of *Mycobacterium tuberculosis* from Mumbai, India: A potential tool for determining drug-resistance and strain lineage. **Tuberculosis (Edinb)** 2017; 107: 63-72.
100. Bainomugisa A, Lavu E, Hiashiri S, Majumdar S, Honjepari A, Moke R, et al. Multi-clonal evolution of multi-drug-resistant/extensively drug-resistant *Mycobacterium tuberculosis* in a high-prevalence setting of Papua New Guinea for over three decades. **Microb Genom** 2018; 4(2).

101. Walker TM, Merker M, Knoblauch AM, Helbling P, Schoch OD, van der Werf MJ, et al. A cluster of multidrug-resistant *Mycobacterium tuberculosis* among patients arriving in Europe from the Horn of Africa: a molecular epidemiological study. **Lancet Infect Dis** 2018; 18(4): 431-440.
102. Kidenya BR, Mshana SE, Fitzgerald DW, Ocheretina O. Genotypic drug resistance using whole-genome sequencing of *Mycobacterium tuberculosis* clinical isolates from North-western Tanzania. **Tuberculosis (Edinb)** 2018; 109: 97-101.
103. Parvaresh L, Crighton T, Martinez E, Bustamante A, Chen S, Sintchenko V. Recurrence of tuberculosis in a low-incidence setting: a retrospective cross-sectional study augmented by whole genome sequencing. **BMC Infect Dis** 2018; 18(1): 265.
104. Koster K, Largen A, Foster JT, Drees KP, Qian L, Desmond EP, et al. Whole genome SNP analysis suggests unique virulence factor differences of the Beijing and Manila families of *Mycobacterium tuberculosis* found in Hawaii. **PLoS One** 2018; 13(7): e0201146.
105. Roof I, Jajou R, Kamst M, Mulder A, de Neeling A, van Hunen R, et al. Prevalence and Characterization of Heterogeneous Variable-Number Tandem-Repeat Clusters Comprising Drug-Susceptible and/or Variable Resistant *Mycobacterium tuberculosis* Complex Isolates in the Netherlands from 2004 to 2016. **J Clin Microbiol** 2018; 56(11).
106. Makhado NA, Matabane E, Faccin M, Pincon C, Jouet A, Boutachkourt F, et al. Outbreak of multidrug-resistant tuberculosis in South Africa undetected by WHO-endorsed commercial tests: an observational study. **Lancet Infect Dis** 2018; 18(12): 1350-1359.
107. Merker M, Barbier M, Cox H, Rasigade JP, Feuerriegel S, Kohl TA, et al. Compensatory evolution drives multidrug-resistant tuberculosis in Central Asia. **Elife** 2018; 7.

108. Mai TQ, Martinez E, Menon R, Van Anh NT, Hien NT, Marais BJ, et al. *Mycobacterium tuberculosis* Drug Resistance and Transmission among Human Immunodeficiency Virus-Infected Patients in Ho Chi Minh City, Vietnam. **Am J Trop Med Hyg** 2018; 99(6): 1397-1406.
109. Conceicao EC, Guimaraes A, Lopes ML, Furlaneto IP, Rodrigues YC, da Conceicao ML, et al. Analysis of potential household transmission events of tuberculosis in the city of Belem, Brazil. **Tuberculosis (Edinb)** 2018; 113: 125-129.
110. Panossian B, Salloum T, Araj GF, Khazen G, Tokajian S. First insights on the genetic diversity of MDR *Mycobacterium tuberculosis* in Lebanon. **BMC Infect Dis** 2018; 18(1): 710.
111. Vaziri F, Kohl TA, Ghajavand H, Kargarpour Kamakoli M, Merker M, Hadifar S, et al. Genetic Diversity of Multi- and Extensively Drug-Resistant *Mycobacterium tuberculosis* Isolates in the Capital of Iran, Revealed by Whole-Genome Sequencing. **J Clin Microbiol** 2019; 57(1).
112. Woodman M, Haeusler IL, Grandjean L. Tuberculosis Genetic Epidemiology: A Latin American Perspective. **Genes (Basel)** 2019; 10(1).
113. Bouzouita I, Cabibbe AM, Trovato A, Daroui H, Ghariani A, Midouni B, et al. Whole-Genome Sequencing of Drug-Resistant *Mycobacterium tuberculosis* Strains, Tunisia, 2012-2016. **Emerging Infectious Diseases** 2019; 25(3): 538-546.
114. Arandjelovic I, Merker M, Richter E, Kohl TA, Savic B, Soldatovic I, et al. Longitudinal Outbreak of Multidrug-Resistant Tuberculosis in a Hospital Setting, Serbia. **Emerging Infectious Diseases** 2019; 25(3): 555-558.
115. Bainomugisa A, Pandey S, Donnan E, Simpson G, Foster J, Lavu E, et al. Cross-Border Movement of Highly Drug-Resistant *Mycobacterium tuberculosis* from Papua New Guinea to Australia through Torres Strait Protected Zone, 2010-2015. **Emerging Infectious Diseases** 2019; 25(3): 406-415.
116. Gomez-Gonzalez PJ, Andreu N, Phelan JE, de Sessions PF, Glynn JR, Crampin AC, et al. An integrated whole genome analysis of *Mycobacterium tuberculosis* reveals insights into relationship between its genome, transcriptome and methylome. **Sci Rep** 2019; 9(1): 5204.

117. Madrazo-Moya CF, Cancino-Munoz I, Cuevas-Cordoba B, Gonzalez-Covarrubias V, Barbosa-Amezcua M, Soberon X, et al. Whole genomic sequencing as a tool for diagnosis of drug and multidrug-resistance tuberculosis in an endemic region in Mexico. **PLoS One** 2019; 14(6): e0213046.
118. Jabbar A, Phelan JE, de Sessions PF, Khan TA, Rahman H, Khan SN, et al. Whole genome sequencing of drug resistant *Mycobacterium tuberculosis* isolates from a high burden tuberculosis region of North West Pakistan. **Sci Rep** 2019; 9(1): 14996.
119. Hang NTL, Hijikata M, Maeda S, Thuong PH, Ohashi J, Van Huan H, et al. Whole genome sequencing, analyses of drug resistance-conferring mutations, and correlation with transmission of *Mycobacterium tuberculosis* carrying katG-S315T in Hanoi, Vietnam. **Sci Rep** 2019; 9(1): 15354.
120. Phelan JE, Lim DR, Mitarai S, de Sessions PF, Tujan MAA, Reyes LT, et al. *Mycobacterium tuberculosis* whole genome sequencing provides insights into the Manila strain and drug-resistance mutations in the Philippines. **Sci Rep** 2019; 9(1): 9305.
121. Acosta F, Agapito J, Cabibbe AM, Caceres T, Sola C, Perez-Lago L, et al. Exportation of MDR TB to Europe from Setting with Actively Transmitted Persistent Strains in Peru. **Emerging Infectious Diseases** 2019; 25(3): 596-598.
122. Nelson KN, Jenness SM, Mathema B, Lopman BA, Auld SC, Shah NS, et al. Social Mixing and Clinical Features Linked With Transmission in a Network of Extensively Drug-resistant Tuberculosis Cases in KwaZulu-Natal, South Africa. **Clin Infect Dis** 2020; 70(11): 2396-2402.
123. Salvato RS, Costa ERD, Reis AJ, Schiefelbein SH, Halon ML, Barcellos RB, et al. First insights into circulating XDR and pre-XDR *Mycobacterium tuberculosis* in Southern Brazil. **Infect Genet Evol** 2020; 78: 104127.
124. Wollenberg K, Harris M, Gabrielian A, Ciobanu N, Chesov D, Long A, et al. A retrospective genomic analysis of drug-resistant strains of *M. tuberculosis* in a high-burden setting, with an emphasis on comparative diagnostics and reactivation and reinfection status. **BMC Infect Dis** 2020; 20(1): 17.

125. Lopez MG, Dogba JB, Torres-Puente M, Goig GA, Moreno-Molina M, Villamayor LM, et al. Tuberculosis in Liberia: high multidrug-resistance burden, transmission and diversity modelled by multiple importation events. **Microb Genom** 2020; 6(1).
126. Lin D, Cui Z, Chongsuvivatwong V, Palittapongarnpim P, Chaiprasert A, Ruangchai W, et al. The geno-spatio analysis of *Mycobacterium tuberculosis* complex in hot and cold spots of Guangxi, China. **BMC Infect Dis** 2020; 20(1): 462.
127. Coker OO, Regmi SM, Suriyaphol P, Chininmanu K, Prammananan T, Chaiprasert A. Whole-Genome Sequence of a Multidrug-Resistant *Mycobacterium tuberculosis* Beijing Sequence Type 10 Isolate from an Outbreak in Thailand. **Genome Announc** 2014; 2(4).
128. Coker OO, Chaiprasert A, Ngamphiw C, Tongsima S, Regmi SM, Clark TG, et al. Genetic signatures of *Mycobacterium tuberculosis* Nonthaburi genotype revealed by whole genome analysis of isolates from tuberculous meningitis patients in Thailand. **PeerJ** 2016; 4: e1905.
129. Faksri K, Tan JH, Disratthakit A, Xia E, Prammananan T, Suriyaphol P, et al. Whole-Genome Sequencing Analysis of Serially Isolated Multi-Drug and Extensively Drug Resistant *Mycobacterium tuberculosis* from Thai Patients. **PLoS One** 2016; 11(8): e0160992.
130. Faksri K, Xia E, Ong RT, Tan JH, Nonghanphithak D, Makhao N, et al. Comparative whole-genome sequence analysis of *Mycobacterium tuberculosis* isolated from tuberculous meningitis and pulmonary tuberculosis patients. **Sci Rep** 2018; 8(1): 4910.
131. Srilohasin P, Prammananan T, Faksri K, Phelan JE, Suriyaphol P, Kamolwat P, et al. Genomic evidence supporting the clonal expansion of extensively drug-resistant tuberculosis bacteria belonging to a rare proto-Beijing genotype. **Emerg Microbes Infect** 2020; 9(1): 2632-2641.
132. 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** 2015; 32: 432-439.

133. Larsen MH, Biermann K, Tandberg S, Hsu T, Jacobs Jr WR. Genetic Manipulation of *Mycobacterium tuberculosis*. **Curr Protoc Microbiol** 2007; 6: 10A.12.11-10A.12.21.
134. Andrews S. **FastQC: A Quality Control tool for High Throughput Sequence Data.** [online] 2010 [cited 2020 June 17]. Available from: https://www.bioinformatics.babraham.ac.uk/index.html
135. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. **arXiv preprint arXiv** 2013; 1303.3997.
136. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics** 2009; 25(16): 2078-2079.
137. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. **Genome Res** 2010; 20(9): 1297-1303.
138. Coll F, McNerney R, Preston MD, Guerra-Assuncao JA, Warry A, Hill-Cawthorne G, et al. Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences. **Genome Med** 2015; 7(1): 51.
139. Phelan JE, O'Sullivan DM, Machado D, Ramos J, Oppong YEA, Campino S, et al. Integrating informatics tools and portable sequencing technology for rapid detection of resistance to anti-tuberculous drugs. **Genome Med** 2019; 11(1): 41.
140. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Mol Biol Evol** 2018; 35(6): 1547-1549.
141. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. **Nucleic Acids Res** 2019: W256–W259.
142. Xu Y, Cancino-Munoz I, Torres-Puente M, Villamayor LM, Borras R, Borras-Manez M, et al. High-resolution mapping of tuberculosis transmission: Whole genome sequencing and phylogenetic modelling of a cohort from Valencia Region, Spain. **PLoS Med** 2019; 16(10): e1002961.

143. Shuaib YA, Khalil EAG, Wieler LH, Schaible UE, Bakheit MA, Mohamed-Noor SE, et al. *Mycobacterium tuberculosis* Complex Lineage 3 as Causative Agent of Pulmonary Tuberculosis, Eastern Sudan(1). **Emerging Infectious Diseases** 2020; 26(3): 427-436.
144. Merker M, Nikolaevskaya E, Kohl TA, Molina-Moya B, Pavlovska O, Brannberg P, et al. Multidrug- and Extensively Drug-Resistant *Mycobacterium tuberculosis* Beijing Clades, Ukraine, 2015. **Emerging Infectious Diseases** 2020; 26(3): 481-490.
145. Boonthanapat N, Soontornmon K, Pungrassami P, Sukhasitwanichkul J, Mahasirimongkol S, Jiraphongsa C, et al. Use of network analysis multidrug-resistant tuberculosis contact investigation in Kanchanaburi, Thailand. **Trop Med Int Health** 2019; 24(3): 320-327.
146. Auld SC, Shah NS, Mathema B, Brown TS, Ismail N, Omar SV, et al. Extensively drug-resistant tuberculosis in South Africa: genomic evidence supporting transmission in communities. **Eur Respir J** 2018; 52(4).
147. Knight GM, McQuaid CF, Dodd PJ, Houben R. Global burden of latent multidrug-resistant tuberculosis: trends and estimates based on mathematical modelling. **Lancet Infect Dis** 2019; 19(8): 903-912.
148. Li WB, Zhang YQ, Xing J, Ma ZY, Qu YH, Li XX. Factors associated with primary transmission of multidrug-resistant tuberculosis compared with healthy controls in Henan Province, China. **Infect Dis Poverty** 2015; 4: 14.
149. Marais BJ, Mlambo CK, Rastogi N, Zozio T, Duse AG, Victor TC, et al. Epidemic spread of multidrug-resistant tuberculosis in Johannesburg, South Africa. **J Clin Microbiol** 2013; 51(6): 1818-1825.
150. Palittapongarnpim P, Ajawatanawong P, Viratyosin W, Smittipat N, Disratthakit A, Mahasirimongkol S, et al. Evidence for Host-Bacterial Co-evolution via Genome Sequence Analysis of 480 Thai *Mycobacterium tuberculosis* Lineage 1 Isolates. **Sci Rep** 2018; 8(1): 11597.
151. Gagneux S, Burgos MV, DeRiemer K, Encisco A, Munoz S, Hopewell PC, et al. Impact of bacterial genetics on the transmission of isoniazid-resistant *Mycobacterium tuberculosis*. **PLoS Pathog** 2006; 2(6): e61.

152. Manjelievskaia J, Erck D, Piracha S, Schrager L. Drug-resistant TB: deadly, costly and in need of a vaccine. **Trans R Soc Trop Med Hyg** 2016; 110(3): 186-191.
153. Sobkowiak B, Banda L, Mzembe T, Crampin AC, Glynn JR, Clark TG. Bayesian reconstruction of *Mycobacterium tuberculosis* transmission networks in a high incidence area over two decades in Malawi reveals associated risk factors and genomic variants. **Microb Genom** 2020.
154. Canetti G, Fox W, Khomenko A, Mahler HT, Menon NK, Mitchison DA, et al. Advances in techniques of testing mycobacterial drug sensitivity, and the use of sensitivity tests in tuberculosis control programmes. **Bull World Health Organ** 1969; 41(1): 21-43.
155. Witney AA, Gould KA, Arnold A, Coleman D, Delgado R, Dhillon J, et al. Clinical application of whole-genome sequencing to inform treatment for multidrug-resistant tuberculosis cases. **J Clin Microbiol** 2015; 53(5): 1473-1483.
156. Dookie N, Rambaran S, Padayatchi N, Mahomed S, Naidoo K. Evolution of drug resistance in *Mycobacterium tuberculosis*: a review on the molecular determinants of resistance and implications for personalized care. **J Antimicrob Chemother** 2018; 73(5): 1138-1151.
157. Bottger EC. The ins and outs of *Mycobacterium tuberculosis* drug susceptibility testing. **Clin Microbiol Infect** 2011; 17(8): 1128-1134.
158. Cambau E, Viveiros M, Machado D, Raskine L, Ritter C, Tortoli E, et al. Revisiting susceptibility testing in MDR-TB by a standardized quantitative phenotypic assessment in a European multicentre study. **J Antimicrob Chemother** 2015; 70(3): 686-696.
159. Kent PT, GP K. **Public Health Mycobacteriology: A Guide for the Level III Laboratory**. Atlanta, Georgia: Department of Health and Human Services, Centers for Disease Control and Prevention; 1985.
160. Banu S, Rahman SM, Khan MS, Ferdous SS, Ahmed S, Gratz J, et al. Discordance across several methods for drug susceptibility testing of drug-resistant *Mycobacterium tuberculosis* isolates in a single laboratory. **J Clin Microbiol** 2014; 52(1): 156-163.

161. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. **Bioinformatics** 2009; 25(16): 2078-2079.
162. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. **arXiv** 2013: 1303.3997.
163. Feuerriegel S, Schleusener V, Beckert P, Kohl TA, Miotto P, Cirillo DM, et al. PhyResSE: a Web Tool Delineating *Mycobacterium tuberculosis* Antibiotic Resistance and Lineage from Whole-Genome Sequencing Data. **J Clin Microbiol** 2015; 53(6): 1908-1914.
164. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Mol Biol Evol** 2018; 35(6): 1547-1549.
165. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. **Nucleic Acids Res** 2019; 47(W1): W256-W259.
166. World Health Organization. **Companion handbook to the WHO guidelines for the programmatic management of drug-resistant tuberculosis**. Geneva, Switzerland: WHO; 2014.
167. World Health Organization. **Technical report on critical concentrations for drug susceptibility testing of medicines used in the treatment of drug-resistant tuberculosis**. Geneva, Switzerland: WHO; 2018.
168. Gygli SM, Keller PM, Ballif M, Blochliger N, Homke R, Reinhard M, et al. Whole-Genome Sequencing for Drug Resistance Profile Prediction in *Mycobacterium tuberculosis*. **Antimicrob Agents Chemother** 2019; 63(4).
169. Dover LG, Alahari A, Gratraud P, Gomes JM, Bhowruth V, Reynolds RC, et al. EthA, a common activator of thiocarbamide-containing drugs acting on different mycobacterial targets. **Antimicrob Agents Chemother** 2007; 51(3): 1055-1063.
170. Heyckendorf J, Andres S, Koser CU, Olaru ID, Schon T, Sturegard E, et al. What Is Resistance? Impact of Phenotypic versus Molecular Drug Resistance Testing on Therapy for Multi- and Extensively Drug-Resistant Tuberculosis. **Antimicrob Agents Chemother** 2018; 62(2).

171. Lempens P, Meehan CJ, Vandelannoote K, Fissette K, de Rijk P, Van Deun A, et al. Isoniazid resistance levels of *Mycobacterium tuberculosis* can largely be predicted by high-confidence resistance-conferring mutations. **Sci Rep** 2018; 8(1): 3246.
172. Jamieson FB, Guthrie JL, Neemuchwala A, Lastovetska O, Melano RG, Mehaffy C. Profiling of rpoB mutations and MICs for rifampin and rifabutin in *Mycobacterium tuberculosis*. **J Clin Microbiol** 2014; 52(6): 2157-2162.
173. Farhat MR, Sixsmith J, Calderon R, Hicks ND, Fortune SM, Murray M. Rifampicin and rifabutin resistance in 1003 *Mycobacterium tuberculosis* clinical isolates. **J Antimicrob Chemother** 2019; 74(6): 1477-1483.
174. Freihofer P, Akbergenov R, Teo Y, Juskeviciene R, Andersson DI, Bottger EC. Nonmutational compensation of the fitness cost of antibiotic resistance in mycobacteria by overexpression of tlyA rRNA methylase. **RNA** 2016; 22(12): 1836-1843.
175. Pule CM, Sampson SL, Warren RM, Black PA, van Helden PD, Victor TC, et al. Efflux pump inhibitors: targeting mycobacterial efflux systems to enhance TB therapy. **J Antimicrob Chemother** 2016; 71(1): 17-26.
176. Rigouts L, Miotto P, Schats M, Lempens P, Cabibbe AM, Galbiati S, et al. Fluoroquinolone heteroresistance in *Mycobacterium tuberculosis*: detection by genotypic and phenotypic assays in experimentally mixed populations. **Sci Rep** 2019; 9(1): 11760.
177. Ruesen C, Riza AL, Florescu A, Chaidir L, Editoiu C, Aalders N, et al. Linking minimum inhibitory concentrations to whole genome sequence-predicted drug resistance in *Mycobacterium tuberculosis* strains from Romania. **Sci Rep** 2018; 8(1): 9676.
178. Wong SY, Lee JS, Kwak HK, Via LE, Boshoff HI, Barry CE, 3rd. Mutations in gidB confer low-level streptomycin resistance in *Mycobacterium tuberculosis*. **Antimicrob Agents Chemother** 2011; 55(6): 2515-2522.
179. Chien JY, Chiu WY, Chien ST, Chiang CJ, Yu CJ, Hsueh PR. Mutations in gyrA and gyrB among Fluoroquinolone- and Multidrug-Resistant *Mycobacterium tuberculosis* Isolates. **Antimicrob Agents Chemother** 2016; 60(4): 2090-2096.

180. Ellington MJ, Ekelund O, Aarestrup FM, Canton R, Doumith M, Giske C, et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. **Clin Microbiol Infect** 2017; 23(1): 2-22.
181. Schön T, Miotto P, Koser CU, Viveiros M, Bottger E, Cambau E. *Mycobacterium tuberculosis* drug-resistance testing: challenges, recent developments and perspectives. **Clin Microbiol Infect** 2017; 23(3): 154-160.
182. Pankhurst LJ, Del Ojo Elias C, Votintseva AA, Walker TM, Cole K, Davies J, et al. Rapid, comprehensive, and affordable mycobacterial diagnosis with whole-genome sequencing: a prospective study. **Lancet Respir Med** 2016; 4(1): 49-58.
183. Jonmalung J, Prammananan T, Leechawengwongs M, Chaiprasert A. Surveillance of pyrazinamide susceptibility among multidrug-resistant *Mycobacterium tuberculosis* isolates from Siriraj Hospital, Thailand. **BMC Microbiol** 2010; 10: 223.
184. Jajou R, van der Laan T, de Zwaan R, Kamst M, Mulder A, de Neeling A, et al. WGS more accurately predicts susceptibility of *Mycobacterium tuberculosis* to first-line drugs than phenotypic testing. **J Antimicrob Chemother** 2019; 74(9): 2605-2616.
185. Lemeshow S, Hosmer D W, Klar J, Lwanga S K, World Health Organization. **Adequacy of Sample Size in Health Studies by World Health Organization**. West Sussex: John Wiley & Sons; 1990.

APPENDICES

APPENDIX A

Sample size calculation for Molecular epidemiology of DR-TB in Thailand

1. Sample size calculation

The sample size was calculated using the estimated case number of MDR-TB. All available XDR-TB from 2014-2017 will be collected due to small number of cases. The estimated numbers of XDR-TB cases are 10-20 cases per year. In Thailand, the estimated MDR-TB patients in 2014 [52] and 2015 [53] were 2,200 and 2,500 cases respectively. The average sample size of MDR-TB in Thailand is 2,350 (N) cases per year. The value of selected alpha level (90% confidence level) is 1.645 ($Z^2_{1-\alpha/2}$). The proportion of the population is 0.5 (P). The confident interval (margin of error) is 10% (d = 0.1). As a result, 66 samples should be collected per year. Therefore, our estimation of sample size concordance to the calculation of sample size from the equation below.

$$n = \left(Z^2_{1-\alpha/2}P(1-P)\cdot N\right)/\left(d^2(N-1)+Z^2_{1-\alpha/2}P(1-P)\right)$$

$$n = \left(1.645^2 0.5(1-0.5)\cdot 2350\right)/\left(0.1^2(2350-1)+1.645^2 0.5(1-0.5)\right)$$

$$n = 65.78$$

Above formula was obtained from the formula number 27 in cited book [185].

APPENDIX B

List of chemicals and instruments

1. List of chemicals and instruments that used in this study

1.1 Chemicals

Chemicals	Sources
Agarose [$(C_{12}H_8O_9)_n$]	Invitrogen
Bromophenol blue ($C_{19}H_{10}Br_4O_5S$)	BIO-RAD
Calcium chloride ($CaCl_2$)	Sigma
Cetyl trimethyl ammonium bromide (CTAB) ($C_{19}H_{42}NBr$)	AMRESCO
Chloroform ($CHCl_3$)	RCI labscan
Ethanol (C_2H_5OH)	RCI labscan
Ethidium bromide ($C_{21}H_{20}BrN_3$)	AMRESCO
Ethylene diamine tetra acetic acid disodium Salt (EDTA) ($C_{10}H_{16}N_2Na_2O_8.2H_2O$)	Fisher Chemical
Glacial acetic acid ($C_2H_4O_2$)	BDH Laboratory
Glycerol ($C_3H_8O_3$)	Calbiochem
Hydrochloric acid (HCl)	BDH Laboratory
Isoamyl alcohol ($C_5H_{12}O$)	Merck
Löwenstein–Jensen medium	Biomedia
Middlebrook OADC (Oleic Albumin Dextrose Catalase) Enrichment	BD BBL
Middlebrook 7H9 Broth (powder)	BD Difco
Middlebrook 7H9 with OADC	Thermo Scientific
Proteinase K (serine protease)	AMRESCO
Sodium chloride (NaCl)	BDH Laboratory
Sodium dodecyl sulfate (SDS) ($NaC_{12}H_{25}SO_4$)	BDH Laboratory
Sodium hydroxide (NaOH)	BDH Laboratory
Tris (Tris(hydroxymethyl)aminomethane) ($C_4H_{11}NO_3$)	Sigma
Tween-80 (Polyoxyethylene (20) sorbitan monooleate) ($C_{64}H_{124}O_{26}$)	Calbiochem

1.2 Instruments

Instruments	Sources
Analytical balance	Satorius
Autoclave, SX-700	TOMY
Automatic pipette	Biohit and SCILOGEX
Biosafety cabinet class II type A2	LABCONCO
Centrifuge, Allegra-x15R	Beckman coulter
Centrifuge, D2012 (Micro-centrifuge)	SCILOGEX
Densitometer DEN-1B	Biosan
Freezer -20ºC and -80ºC	Sanyo and Thermo Scientific
Gel electrophoresis	BIO-RAD
Gel Doc XR+ System	BIO-RAD
Heat block	Benchmark
Hot air oven	Memmert
Incubator 37 ºC	Memmert
Magnetic Stirrer C-MAG`MS4	IKA
Microwave	Sharp
Nanodrop 2000c Spectrophotometer	Thermo Scientific
Oven	DAIHAN
PCR Vertical Laminar Flow Cabinets	Esco Airstream
pH meter	Beckman coulter
Sensititre AIM Automated Inoculation Delivery System	Thermo Scientific
Sensititre Doseheads for plate inoculation	Thermo Scientific
Sensititre MYCOTBI AST Plate	Thermo Scientific
Sensititre Vizion Digital MIC Viewing System	Thermo Scientific
Vortex mixer	Scientific Industries

APPENDIX C

Preparation of reagents for *Mtb* culture and *Mtb* DNA extraction

1. Reagents for *Mtb* culture

1.1 Middlebrook 7H9 broth with 10% OADC (200 ml)

Add 0.94 g of Middle brook 7H9 powder, 0.4 ml of glycerol and 180 ml of distilled water. Mix using magnetic stirrer. Sterilization at 121 °C for 10 min. Cooling at room temperature. Store the media at 2-8 °C up to 3 months. Add 20 ml of Middlebrook OADC supplement by aseptic technique before use.

1.2 50% Glycerol (100 ml)

Add 50 ml of glycerol into 50 ml of DI water. Mix using magnetic stirrer. Sterilization at 121 °C for 15 min.

1.3 Normal saline with 0.2% tween 80 (200 ml)

Add 1.8 g of NaCl, 400 µl of tween 80 and 200 ml of distilled water. Mix using magnetic stirrer. Sterilization at 121 °C for 15 min.

2. Reagents for *Mtb* DNA extraction

2.1 0.5 M EDTA pH 8.0 (100 ml)

g/MW = CV/1000

g = (0.5)(100)(372.24)/1000 = 18.6 gram of EDTA disodium dihydrate

Add 18.6 g of EDTA disodium dehydrate and 50 ml of DI water. Mix using magnetic stirrer and add NaOH simultaneously until the solution is well dissolve at the pH of 8. Adjust the volume to 100 ml with DI water. Sterilization at 121 °C for 15 min.

2.2 1 M Tris-HCl pH 8.0 (100 ml)

g/MW = CV/1000

g = (1)(100)(121.14)/1000 = 12.1 gram of Tris

Add 12.1 g of Tris and 50 ml of DI water. Mix using magnetic stirrer and adding HCl simultaneously until the solution has the pH of 8. Adjust the volume to 100 ml with DI water. Sterilization at 121 °C for 15 min.

2.3 1X TE buffer (100 ml)

Composition: 10 mM Tris-HCl (pH 8), 1 mM EDTA (pH 8)

Stock solution: 1 M Tris-HCl pH 8, 0.5 M EDTA pH 8 C1V1 = C2V2

How much of 1 M Tris-HCl need to added? V1 = (10x0.001)(100)/(1) ∴ V1 = 1 ml

How much of 0.5 M EDTA need to added? V1 = (1x0.001)(100)/(0.5) ∴ V1 = 200 µl

Add 1 ml of 1 M Tris-HCl pH 8 and 200 µl of 0.5 M EDTA pH 8 and adjust the volume to 100 ml with DI water. Sterilization at 121 °C for 15 min.

2.4 10% SDS (100 ml)

Add 10 g of SDS, 100 µl of DI water. Mix using magnetic stirrer.

2.5 Lysozyme 10 mg/ml (30 ml)

And 0.3 g of lysozyme, 30 ml of TE buffer. Vortex and aliquot into several microcentrifuge tubes. Store at -20 °C for long term.

2.6 Proteinase K solution (100 ml)

Composition: 50 mM Tris-HCl (pH 8),1 mM $CaCl_2$ and 50% glycerol.

Stock solution: 1 M Tris-HCl pH 8 $C_1V_1 = C_2V_2$

How much of 1 M Tris-HCl need to added? $V_1 = (50x0.001)(100)/(1) \therefore V_1 = 5$ ml

How much of $CaCl_2$ need to added? g/MW = CV/1000

g = (1x0.001)(100)(110.98)/1000 = 0.01 gram of $CaCl_2$

Add 5 ml of 1 M Tris-HCl pH 8, 0.01 g of $CaCl_2$ into 100 ml of 50% glycerol. Mix using magnetic stirrer and sterilization at 121 °C for 15 min.

2.7 Proteinase K 10 mg/ml (10 ml)

And 0.1 g of proteinase K, 10 ml of proteinase K solution. Vortex and aliquot into several microcentrifuge tubes. Store at -20 °C for long term.

2.8 RNase A 10 mg/ml (5 ml)

And 0.05 g of RNase A, 5 ml of TE buffer. Vortex and aliquot into several microcentrifuge tubes. Store at -20 °C for long term.

2.9 5 M NaCl (100 ml)

Composition: 5 M NaCl, 20 mM EDTA, 40 mM Tris-HCl (pH 8)

Stock solution: 1 M Tris-HCl pH 8, 0.5 M EDTA pH 8 $C_1V_1 = C_2V_2$

How much of 1 M Tris-HCl need to added? $V_1 = (0.04)(100)/(1) \therefore V_1 = 4$ ml

How much of 0.5 M EDTA need to added? $V_1 = (0.02)(100)/(0.5) \therefore V_1 = 4$ ml

How much of NaCl need to added? g/MW = CV/1000

g = (5)(100)(58.44)/1000 = 29.22 gram of NaCl

Add 4 ml of 1 M Tris-HCl pH 8, 4 ml of 0.5 M EDTA pH 8, 29.22 g of NaCl into 100 ml of DI water. Mix using magnetic stirrer and sterilization at 121 °C for 15 min.

2.10 CTAB/NaCl (100 ml)

Composition: 10% CTAB and 0.7 M NaCl

Stock solution: 5 M NaCl $C_1V_1 = C_2V_2$

How much of 5 M NaCl need to added? $V_1 = (0.7)(100)/(5) \therefore V_1 = 14$ ml

Add 10 g of CTAB, 14 ml of 5 M NaCl g of NaCl and 86 ml of DI water. Mix using magnetic stirrer.

2.11 Chloroform:isoamyl alcohol solution 24:1 (100 ml)

Add 96 ml of chloroform and 4 ml of isoamyl alcohol. Mix well.

2.12 70% Ethanol

Add 70 ml of absolute ethanol into 100 ml cylinder after that adjust the volume by DI water until the solution is 100 ml. Mix well.

APPENDIX D

Genomic DNA extraction of *Mtb* colonies using CTAB

1. Genomic DNA extraction of *Mtb* colonies using CTAB

Genomic DNA extraction was performed as following [133]. Multiple loops of *Mtb* colonies were transferred into sterile tube (16x100 mm) containing six glass beads (4 mm.) and 3 drops of saline with tween. Then, vortex until the clumping colonies were breakdown, leave for at least 15 min and adding 800 µl of TE buffer, all steps above were performed under biosafety cabinet class II. The tube was placed at 80°C for 30 min (killing of mycobacteria) and cooling at room temperature. Adding 100 µl of 10 mg/ml lysozyme, thoroughly mix and incubated at 37°C overnight. Adding 140 µl of 10% SDS. Adding 20 µl of 10 mg/ml protenase K. After that, vortex and incubate at 65°C for 20 min. Before transferring the suspension into two microcentrifuge tubes, 100 µl of 5 M NaCl and 100 µl of pre-warmed CTAB/NaCl (pre-warmed at 65°C) were added into each new micro tube. After that, 500 µL of suspension was transferred into each of two micro centrifuge tubes containing 5 M NaCl and CTAB/NaCl solution, mix and incubate at 65°C for 10 min. Then, adding 750 µl of chloroform/isoamyl alcohol solution into each micro tubes, vertex for at least 10 sec and centrifugation (10,000 rpm) for 5 min. Before transferring the clear aqueous phase from the old tubes, 10 µl of 10 mg/ml RNase A was added into each of new microcentrifuge tubes. Next, transfer the aqueous supernatant into each micro tube containing RNaseA and incubated at 37°C for 30 min. After that, 1 ml of cold absolute ethanol was added into each of micro tubes, then, the tubes were gently inverted mixed for a 4-5 times. At this step, the participated DNA could be seen with naked eyes. Then, place the tubes at freezer (-20°C) for 30 min. After that, centraifugation (10,000 rpm) for 10 min in order to collect the DNA pellet and discard the supernatant. The DNA pellet was purified with 1 ml of cold 70% ethanol and centrifugution (10,000 rpm) for 5 min. Then, re-purification the DNA pellet and centrifugution (10,000 rpm) for 1 min. Gently discard the remaining ethanol. Allow the DNA pellet for half-dry (25°C), and re-dissolved the pellet with 50 µl of TE buffer. The extractions were stored at -20°C (long-term storage at -80°C). Quantification of DNA was measured using spectrophotometers at the OD ratio of 260/280 (OD = 1.8-1.9 indicates good quality of the extraction which acceptable and be able used for further analysis). The integrity of DNA can be checked by agarose gel electrophoresis. The purity, concentration and total volume of the extraction that strongly recommend for WGS are OD 260/280 = 1.8-2.0, ≥20 ng/µl and ≥30µl respectively.

APPENDIX E

Supplementary data

Supplementary Table 1 Distribution of all culture-confirmed DR-TB cases (according to laboratory records) during 2014-2017

Regions	Provinces[a]	2014 (n = 573)			2015 (n = 608)			2016 (n = 480)			2017 (n = 410)			Total (n = 2,071)		
		MDR	Pre-XDR	XDR	MDR-TB	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR-TB	MDR	Pre-XDR	XDR
Central	BKK	91	12	5	92	10	9	81	10	2	64	8	1	328	40	17
	CNT	3			3			2	1		1			9	1	
	KPT	8	1		2						3			13	1	
	LRI	2			1			2			3			8		
	NYK										2			2		
	NPT				7			7	1		4	1		18	2	
	NSN	4	1		9	1		10	1		2			25	3	
	NBI	20	1		5	2		4			7			36	3	
	AYA				2									2		
	PTE				1									1		
	PNB	5	2		5			2	1		2			14	3	
	PCT	3			5			1	1		1	1		10	2	
	PLK	8			5			2	1		4			19	1	
	SPK	6			10			1	1		5			22	1	
	SKN	4	1		5			5	1		12	2		26	4	
	SKM				1						3			4		
	SRI	2									4			6		
	SBR	1			2			1						4		
	STI	2			7			4						13		
	SPB	5	1		8			3	1		9	2		25	4	
	UTI	3			1			1						5		
Eastern	CCO	9	2		5			5			7			26	2	
	CTI	8	2	1	13	1	1	9			6			36	3	2
	CBI	34	7	1	19	4	5	25	4	3	24	4	1	102	19	10
	PRI	3	1		5	1		5	1		4	1		17	4	
	RYG	21	1		16			7	1		4	1		48	3	
	SKW	3	1		5		1	2			3			13	1	1
	TRT	3			7	1	1	7	1		2	1	1	19	3	2
Northeastern	ACR	3			1	1	1	2		1	2			8	1	2
	BKN				1			1						2		
	BRM	21			18	2		7	1		6	1		52	4	
	CPM		1		6	1		3	3		2	1	2	11	6	2
	KSN	1			4			9	2				1	14	2	1
	KKN	13	4		12	1		19	3	1	6		1	50	8	2
	LEI	1	1		1			4			1			7	1	
	MKM	5			5	3	1	12			4			26	3	1
	MDH	2	1	1	1		1	1						4	1	2
	NPM	3			6	1		2	1					11	2	
	NMA				8		1	6	3		7	1	1	21	4	2

Supplementary Table 1 Distribution of all culture-confirmed DR-TB cases (according to laboratory records) during 2014-2017 (Cont.)

Regions	Provinces[a]	2014 (n = 573)			2015 (n = 608)			2016 (n = 480)			2017 (n = 410)			Total (n = 2,071)		
		MDR	Pre-XDR	XDR	MDR-TB	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR-TB	MDR	Pre-XDR	XDR
Northeastern	NBP				2				2		2			4	2	
	NKI	2			7			4				2		13	2	
	RET	4	3	1	11	2		6	3		3			24	8	1
	SNK	7	1		3	1		5						15	2	
	SSK	11	2		9	2		10	1		6	1		36	6	
	SRN				5	1		1	1		3	1		9	3	
	UBN	14	2	1	15	2		5	6		6	1		40	11	1
	UDN	8			7			13		1				28		1
	YST	4						2			2			8		
Northern	CMI				19	2		4			10	1		33	3	
	CRI	4			1			1				1		6	1	
	LPG				1									1		
	LPN					1									1	
	PYO				2	1								2	1	
	PRE	6			3	1		1	1					10	2	
	UTT	3			3	1		6	2		1			13	3	
Southern	CPN	4	1		5			1			1			11	1	
	KBI				2			2			1			5		
	NST	19	3		15	1		14	3		8	3		56	10	
	NWT				4			3			1			8		
	PTN	3			1			2			1	1		7	1	
	PNA	2			1			2	1					5	1	
	PLG				4			1			3			8		
	PKT	11	2	1	12		2	2			4	2		29	4	3
	RNG			1		1				1					1	2
	STN	4			5	2		2			1			12	2	
	SKA	11			15	1		6	1		6			38	2	
	SNI	7	1		9	1	1	4			4			24	2	1
	TRG	1			5			3			1			10		
	YLA	2			2			2			2			8		
Western	KRI	40	3	2	34	5	2	24	6		44	11	3	142	25	7
	PBI	3			3	3		11			11	1		28	4	
	PKN	3			2			3			8		1	16		1
	RBR	7		2	11	1	1	6	2		10		2	34	3	5
	TAK	14	5	3	1			4	1		3	1		22	7	3
Total		**491**	**63**	**19**	**523**	**58**	**27**	**402**	**69**	**9**	**346**	**50**	**14**	**1762**	**240**	**69**

Note: Geographic locations (provinces) were based on the hospital location associated with the residential location of the patients according to the national health coverage

[a] Full name of all provinces were listed in Supplementary Table 12

Supplementary Table 2 Studied population of DR-TB cases in Thailand, arranged by year and type of DR

Years	Estimated Thai TB prevalence (WHO global TB report)				Culture confirmed MDR/XDR-TB cases (Lab records)				Sample size in this study			
	Total TB cases	MDR/RR-TB	Lab-confirmed MDR/RR-TB	XDR-TB	MDR-TB	Pre-XDR-TB	XDR-TB	Total	MDR-TB	Pre-XDR-TB	XDR-TB*	Total
2014	71,618	2,200	506	NA	491	63	19	**573**	109	18	9	**136**
2015	66,179	2,500	466	5	523	58	27	**608**	112	9	8	**129**
2016	72,014	2,700	955	13	402	69	9	**480**	111	27	6	**144**
2017	82,008	2,700	1,339	7	346	50	14	**410**	134	27	9	**170**
Total			**3,266**					**2,071**	**466**	**81**	**32**	**579**

Note: The sample size represents the WHO global TB report in 2014-2016 (except 2017). *These XDR-TB isolates were all culturable according to the stock culture records (some isolates did not grow). Therefore, all retrievable XDR-TB were included in this study.

Supplementary Table 3 Demographic data of the drug-resistant tuberculosis patients

Demographic data		Phenotypic drug-resistant tuberculosis types			
		MDR-TB (n = 466)	Pre-XDR-TB (n = 81)	XDR-TB (n = 32)	Total
Age*	Mean (SD)	43.71 (±14.84)	44.19 (±14.39)	38.93 (±12.58)	43.51 (±14.68)
	<60	352 (85.44)	56 (82.35)	27 (96.43)	435 (85.63)
	≥60	60 (14.56)	12 (17.65)	1 (3.57)	73 (14.37)
Gender*	Female	119 (25.81)	24 (30)	11 (34.38)	154 (26.88)
	Male	342 (74.19)	56 (70)	21 (65.63)	419 (73.12)

*Data for age and gender were available for 508 (87.47%) and 573 (98.96%) cases respectively.

Supplementary Table 4 Distribution of DR isolates (according to laboratory records) used in this study

Regions	Abbreviations[a]	2014 (n = 136)			2015 (n = 129)			2016 (n = 144)			2017 (n = 170)			Total (n = 579)		
		MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR
Central	BKK	13	2	1	14		2	20	6	1	18	7	1	65	15	5
	CNT	1			2				1		1			4	1	
	KPT	2	1		1						1			4	1	
	LRI	1			1			1						3		
	NYK										1			1		
	NPT				2			2			2			6		
	NSN	2			2			3			1			8		
	NBI	3	1		2			3			4			12	1	
	AYA				1									1		
	PTE				1									1		
	PNB	1			1				1		1			3	1	
	PCT	1			1				1			1		2	2	
	PLK	1			1				1		1			3	1	
	SPK	1			3			1			3			8		
	SKN	1			3				1		6			10	1	
	SKM										2			2		
	SRI	1									1			2		
	SBR	1			1			1						3		
	STI	1			1			1						3		
	SPB	2			3				1		3	1		8	2	
	UTI	2				1		1						3	1	
Eastern	CCO	1	1		1			1			3			6	1	
	CTI	3			4			3			2			12		
	CBI	7	3	1	3	1		5	1	2	12	1	1	27	6	4
	PRI	1	1		1			1	1		2	1		5	3	
	RYG	3	1					2	1		2			7	2	
	SKW	1			1						2			4		
	TRT	3			2		1	2	1			1	1	7	2	2
Northeastern	ACR	1			1		1			1	1			3		2
	BKN				1									1		
	BRM	4			3	1		3			2			12	1	
	CPM				2				2		1		1	3	2	1
	KSN				1			2						3		
	KKN	3	1		1			5	1	1	3			12	2	1
	LEI				1			1			1			3		
	MKM				1			3			3			7		
	MDH	1		1	1									2		1
	NPM	1			1			1						3		
	NMA				2			2			2	1		6	1	

Supplementary Table 4 Distribution of DR isolates (according to laboratory records) used in this study (Cont.)

Regions	Abbreviations[a]	2014 (n = 136)			2015 (n = 129)			2016 (n = 144)			2017 (n = 170)			Total (n = 579)		
		MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR
Northeastern	NBP								1		2			2	1	
	NKI	1			1			2				1		4	1	
	RET	1	1		2			4			1			8	1	
	SNK	2			1			1						4		
	SSK	2			2	1		3			2			9	1	
	SRN				1			1			1			3		
	UBN	4	1	1	3				4		3			10	5	1
	UDN	2			2			2						6		
	YST	1						1			1			3		
Northern	CMI				5			1			3	1		9	1	
	CRI											1			1	
	PRE	1				1		1						2	1	
	UTT				1			1	1					2	1	
Southern	CPN	1			1			1						3		
	KBI				1			1			1			3		
	NST	3	1		3			4			3	2		13	3	
	NWT				1			1						2		
	PTN							1			1			2		
	PNA	1			1			1						3		
	PLG				1						1			2		
	PKT	3	1	1	2			1			2			8	1	1
	RNG			1		1				1					1	2
	STN	1				1	1	2			1			4	1	1
	SKA	2			3			2	1		2			9	1	
	SNI	2			1	1		2			1			6	1	
	TRG				1			1			1			3		
	YLA	1						1			1			3		
Western	KRI	10	2	1	7		2	4	1		14	8	2	35	11	5
	PBI	2			1	1		4			5	1		12	2	
	PKN	1						1			3		1	5		1
	RBR	2			3		1	2			3		2	10		3
	TAK	3	1	2	1			1			1			6	1	2
Total		**109**	**18**	**9**	**112**	**9**	**8**	**111**	**27**	**6**	**134**	**27**	**9**	**466**	**81**	**32**

[a] Full name of all provinces were listed in Supplementary Table 12

Supplementary Table 5 Distribution by *Mtb* of DR isolates (according to laboratory records) used in this study

Region	Abbreviation[a]	2014 (n = 136)			2015 (n = 129)			2016 (n = 144)			2017 (n = 170)				Total (n = 579)			
		L1	L2	L4	L1	L2	L4	L1	L2	L4	L1	L2	L3	L4	L1	L2	L3	L4
Central	BKK	1	14	1	1	13	2	2	23	2		24		2	4	74		7
	CNT		1			2			1			1				5		
	KPT		3			1					1				1	4		
	LRI		1			1			1							3		
	NYK											1				1		
	NPT					2			2			2				6		
	NSN		2			2			3			1				8		
	NBI		3	1		2			3			3		1		11		2
	AYA					1										1		
	PTE					1										1		
	PNB		1			1			1			1				4		
	PCT	1				1			1			1			1	3		
	PLK	1				1			1			1			1	3		
	SPK		1			3			1			3				8		
	SKN		1		1	2			1			6			1	10		
	SKM											2				2		
	SRI			1								1				1		1
	SBR		1			1			1							3		
	STI		1			1			1							3		
	SPB		2			3			1			4				10		
	UTI	1		1		1			1						1	2		1
Eastern	CCO		2			1				1		3				6		1
	CTI	1	2			4			3		2				3	9		
	CBI		11		1	3		1	6	1		13		1	2	33		2
	PRI		2			1			2		1	2			1	7		
	RYG	1	3						2	1		1		1	1	6		2
	SKW		1			1					1	1			1	3		
	TRT		3		1	2		2	1			2			3	8		
Northeastern	ACR	1				2			1			1			1	4		
	BKN					1										1		
	BRM	1	3			4		1	2			2			2	11		
	CPM					2			2			2				6		
	KSN					1		1	1						1	2		
	KKN	2	1	1		1			5	2	1	2			3	9		3
	LEI					1				1		1				2		1
	MKM					1			3		1	2			1	6		
	MDH		2		1										1	2		
	NPM		1			1		1							1	2		
	NMA					2			2			3				7		
	NBP								1		2				2	1		

Supplementary Table 5 Distribution by *Mtb* of DR isolates (according to laboratory records) used in this study (Cont.)

Region	Abbreviation[a]	2014 (n = 136)			2015 (n = 129)			2016 (n = 144)			2017 (n = 170)				Total (n = 579)			
		L1	L2	L4	L1	L2	L4	L1	L2	L4	L1	L2	L3	L4	L1	L2	L3	L4
Northeastern	NKI		1				1		2			1				4		1
	RET		1	1		2		2	2		1				3	5		1
	SNK	1	1		1				1						2	2		
	SSK	1	1		1	2		2	1			2			4	6		
	SRN					1			1			1				3		
	UBN	2	4		1	2			4		1	2			4	12		
	UDN		2			2			2							6		
	YST	1						1			1				3			
Northern	CMI				1	4			1			4			1	9		
	CRI										1				1			
	PRE		1			1			1							3		
	UTT					1			2							3		
Southern	CPN		1			1			1							3		
	KBI						1		1			1				2		1
	NST	1	3		1	2		1	2	1	1	2		2	4	9		3
	NWT					1			1							2		
	PTN								1		1				1	1		
	PNA		1			1			1							3		
	PLG					1						1				2		
	PKT		5			2		1				2			1	9		
	RNG		1			1			1							3		
	STN		1			2		1	1			1			1	5		
	SKA		2		1	2			3		1	1			2	8		
	SNI		2			2			2			1				7		
	TRG				1				1		1				2	1		
	YLA		1						1			1				3		
Western	KRI		13			9		1	4			22	1	1	1	48	1	1
	PBI		2			2			4		1	5			1	13		
	PKN		1						1		1	2		1	1	4		1
	RBR		2		1	3			2			4		1	1	11		1
	TAK	2	4			1			1			1			2	7		
Total		**18**	**112**	**6**	**13**	**112**	**4**	**17**	**118**	**9**	**19**	**140**	**1**	**10**	**67**	**482**	**1**	**29**

[a] Full name of all provinces were listed in Supplementary Table 12

Supplementary Table 6 Characteristics of 89 (C1-C89) DR-TB clusters defined only by SNP pairwise differences ≤11

No.	Name of cluster	Number of isolates	Member of major clade	Drug resistant types* (no.)	Geographical link			Time link, year (no.)
					Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
1	C1	2	No	MDR (2)	Northeastern (2)	Buri Ram (2)	Krasang Hospital (2)	2014 (1), 2016 (1)
2	C2	3	No	MDR (2), pre-XDR (1)	Central (1), Eastern (2)	Bangkok (1), Rayong (2)	Sirinthorn (1), Rayong (2)	2016 (2), 2017 (1)
3	C3	2	No	MDR (2)	Central (1), Eastern (1)	Nonthaburi (1), Chon Buri (1)	Bamrasnaradura Institute (1), Chonburi Hospital (1)	2016 (1), 2017 (1)
4	C4	3	No	MDR (3)	Northeastern (1), Central (2)	Loei (1), Nonthaburi (1), Saraburi (1)	Naduang Hospital (1), Bamrasnaradura Institute (1), Saraburi Hospital (1)	2014 (2), 2016 (1)
5	C5	3	No	MDR (3)	Southern (3)	Krabi (1), Nakhon Si Thammarat (2)	Nueklong Hospital (1), Maharajnakhonsithammarat Hospital (1), Office of Disease Prevention & Control 11 Nakhon Si Thammarat (1)	2015 (1), 2017 (2)
6	C6	2	No	MDR (2)	Northeastern (2)	Khon Kaen (2)	Khonkaen Hospital (2)	2014 (1), 2016 (1)
7	C7	2	No	MDR (1), pre-XDR (1)	Northern (2)	Phrae (2)	Phrae Hospital (1), Sungmen Hospital (1)	2014 (1), 2015 (1)
8	C8	2	No	MDR (2)	Central (2)	Bangkok (1), Nonthaburi (1)	Devision of Tuberculosis (1), Bamrasnaradura Institute (1)	2015 (1), 2016 (1)
9	C9	3	No	MDR (1), XDR (2)	Western (3)	Kanchanaburi (2), Ratchaburi (1)	Makarak Hospital (2), Ratchaburi Hospital (1)	2015 (1), 2016 (1), 2017 (1)
10	C10	5	**Yes, Clade1**	MDR (3), pre-XDR (1), XDR (1)	Eastern (5)	Trat (5)	Trat Hospital (5)	2014 (2), 2015 (1), 2016 (1), 2017 (1)
11	C11	2	No	pre-XDR (2)	Central (2)	Bangkok (2)	Rajavithi Hospital (2)	2016 (2)
12	C12	2	No	MDR (2)	Central (2)	Lop Buri (2)	Khoksamrong Hospital (2)	2014 (1), 2015 (1)
13	C13	3	**Yes, Clade 2**	MDR (3)	Central (2), Western (1)	Bangkok (2), Prachuap Khiri Khan (1)	Rajavithi Hospital (1), Klang Hospital (1), Bangsabhan Hospital (1)	2016 (1), 2017 (2)
14	C14	2	**Yes, Clade 2**	MDR (2)	Northeastern (1), Central (1)	Maha Sarakham (1), Samut Prakan (1)	Phayakkhaphumphisai Hospital (1), Bangbo Hospital (1)	2017 (2)
15	C15	2	**Yes, Clade 2**	MDR (2)	Western (2)	Prachuap Khiri Khan (2)	Samroiyod Hospital (2)	2014 (1), 2016 (1)
16	C16	5	**Yes, Clade 2**	MDR (2), pre-XDR (3)	Northeastern (3), Central (1), Western (1)	Ubon Ratchathani (2), Udon Thani (1), Bangkok (1), Kanchanaburi (1)	Fort sunpasitthiprasong Hospital (1), Trakanphuetpol Hospital (1), Udonthani Hospital (1), Rajavithi Hospital (1), Makarak Hospital (1)	2014 (2), 2016 (2), 2017 (1)
17	C17	2	**Yes, Clade 2**	MDR (2)	Northeastern (1), Central (1)	Khon Kaen (1), Bangkok (1)	Khonkaen Hospital (1), Rajavithi Hospital (1)	2016 (1), 2017 (1)

Supplementary Table 6 Characteristics of 89 (C1-C89) DR-TB clusters defined only by SNP pairwise differences ≤11 (Cont.)

No.	Name of cluster	Number of isolates	Member of major clade	Drug resistant types* (no.)	Geographical link			Time link, year (no.)
					Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
18	C18	2	**Yes, Clade 2**	MDR (2)	Southern (2)	Chumphon 2	Chumphonkhetudomsakdi Hospital 2	2015 (1), 2016 (1)
19	C19	4	**Yes, Clade 2**	MDR (4)	Northeastern (2), Eastern (1), Southern (1)	Amnat Charoen (1), Ubon Ratchathani (1), Chon Buri (1), Songkhla (1)	Amnatcharoen Hospital (1), Somdetphrayuphrarat Detudom Hospital (1), Chonburi Hospital (1), Songkhla Hospital (1)	2017 (3), 2016 (1)
20	C20	2	No	XDR (2)	Northeastern (2)	Amnat Charoen (2)	Amnatcharoen Hospital (2)	2015 (1), 2016 (1)
21	C21	2	No	MDR (1), pre-XDR (1)	Northeastern (2)	Khon Kaen (2)	Khonkaen Hospital (1), Srinagarind Hospital (1)	2015 (1), 2016 (1)
22	C22	2	No	MDR (2)	Central (1), Western (1)	Samut Songkhram (1), Phetchaburi (1)	Somdejphraphutthaloetla Hospital (1), Phrachomklao Hospital (1)	2016 (1), 2017 (1)
23	C23	2	**Yes, Clade 3**	XDR (2)	Western (2)	Ratchaburi (2)	Ratchaburi Hospital (2)	2015 (1), 2017 (1)
24	C24	2	**Yes, Clade 3**	MDR (2)	Eastern (2)	Rayong (1), Chon Buri (1)	Rayong Hospital (1), Chonburi Hospital (1)	2014 (1),2016 (1)
25	C25	2	**Yes, Clade 3**	MDR (2)	Central (1), Southern (2)	Samut Prakan (1), Surat Thani (2)	Bangbo Hospital (1), Suratthani Hospital (1)	2016 (1), 2017 (1)
26	C26	2	No	MDR (2)	Northeastern (1), Southern (1)	Udon Thani (1), Surat Thani (1)	Udonthani Hospital (1), Kohsamui Hospital (1)	2014 (1), 2016 (1)
27	C27	2	No	MDR (2)	Central (1), Eastern (1)	Pathum Thani (1), Trat (1)	Ladlumkaew Hospital (1), Trat Hospital (1)	2014 (1), 2015 (1)
28	C28	3	**Yes, Clade 4**	MDR (3)	Northeastern (3)	Khon Kaen (1), Maha Sarakham (2)	Banphai Hospital (1), Borabue Hospital (2), Mahasarakham Hospital (3)	2016 (2), 2017 (1)
29	C29	4	No	MDR (4)	Central (2), Eastern (2)	Chai Nat (1), Suphan Buri (1), Chachoengsao (1), Prachin Buri (1)	Hankha Hospital (1), Chaophrayayommarat Hospital (1), Buddhasothorn Hospital (1), Prachantakham Hospital (1)	2014 (1), 2015 (1), 2017 (2)
30	C30	3	No	MDR (3)	Northeastern (3)	Buri Ram (3)	Banruat Hospital (1), Buriram Hospital (2)	2015 (3)
31	C31	3	**Yes, Clade 5**	MDR (3)	Northeastern (1), Central (1), Eastern (1)	Buri Ram (1), Saraburi (1), Prachin Buri (1)	Buriram Hospital (1), Saraburi Hospital (1), Chaopraya Abhaiphubet Hospital (1)	2014 (1), 2017 (2)
32	C32	2	No	XDR (2)	Western (1)	Tak (1)	Maesot Hospital (1)	2014 (2)
33	C33	2	No	MDR (2)	Siuthern (2)	Pattani (1), Yala (1)	Pattani Hospital (1), Yala Hospital (1)	2016 (2)
34	C34	2	No	MDR (2)	Northeastern (1), Central (1)	Khon Kaen (1), Bangkok (1)	Somdetphrayuphrarat Kranuan Hospital (1), Public Health Center 27 (1)	2015 (1), 2016 (1)
35	C35	2	**Yes, Clade 6**	XDR (2)	Southern (1), Western (1)	Phuket (1), Prachuap Khiri Khan (1)	Vachiraphuket Hospital (1), Hua-Hin Hospital (1)	2014 91), 2017 (1)

Supplementary Table 6 Characteristics of 89 (C1-C89) DR-TB clusters defined only by SNP pairwise differences ≤11 (Cont.)

No.	Name of cluster	Number of isolates	Member of major clade	Drug resistant types* (no.)	Geographical link			Time link, year (no.)
					Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
36	C36	5	**Yes, Clade 6**	MDR (4), pre-XDR (1)	Northeastern (2), Southern (2)	Udon Thani (1), Nakhon Ratchasima (1), Phuket (2)	Udonthani Hospital (1), Sikhio Hospital (1), Patong Hospital (1), Vachiraphuket Hospital (1)	2014 (1), 2015 (2), 2016 (1), 2017 (1)
37	C37	3	No	MDR (1), pre-XDR (2)	Southern (3)	Satun (3)	Satun Hospital (3)	2014 (1), 2015 (2)
38	C38	2	**Yes, Clade7**	MDR (2)	Southern (2)	Krabi (1), Satun (1)	Khlongthom Hospital (1), Satun Hospital (1)	2016 (2)
39	C39	3	**Yes, Clade7**	MDR (3)	Southern (3)	Nakhon Si Thammarat (1), Phuket (1), Phang Nga (1)	Office of Disease Prevention & Control 11 Nakhon Si Thammarat (1), Vachiraphuket Hospital (1), Khuraburichaipat Hospital (1)	2014 (1), 2016 (1), 2017 (1)
40	C40	2	**Yes, Clade7**	MDR (1), pre-XDR (1)	Central (1), Southern (1)	Nonthaburi (1), Trang (1)	National Institue of Health of Thailand (1), Kantang Hospital (1)	2014 (1), 2016 (1)
41	C41	2	**Yes, Clade7**	MDR (2)	Southern (2)	Surat Thani (1), Phuket (1)	Suratthani Hospital (1), Vachiraphuket Hospital (1)	2014 (1), 2015 (1)
42	C42	2	**Yes, Clade7**	MDR (2)	Southern (2)	Phatthalung (2)	Kongrha Hospital (1), Phatthalung Hospital (1)	2015 (1), 2017 (1)
43	C43	2	No	pre-XDR (1), XDR (1)	Northeastern (2)	Chaiyaphum (2)	Chaiyaphum Hospital (2)	2016 (1), 2017 (1)
44	C44	2	No	MDR (2)	Northeastern (1), Central (1)	Bungkan (1), Phetchabun (1)	Sriwilai Hospital (1), Nongphai Hospital (1)	2015 (2)
45	C45	2	No	MDR (2)	Central (2)	Bangkok (1), Phichit (1)	Sirinthorn Hospital (1), Wangsaiphun Hospital (1)	2014 (1), 2015 (1)
46	C46	2	No	XDR (2)	Central (2)	Bangkok (2)	Devision of Tuberculosis (2)	2015 (1), 2016 (1)
47	C47	2	No	pre-XDR (2)	Central (2)	Bangkok (1), Kamphaeng Phet (1)	Police Hospital (1), Kamphaengphet Hospital (1)	2014 (2)
48	C48	5	**Yes, Clade8**	MDR (5)	Central (1), Eastern (4)	Nonthaburi (1), Rayong (1), Chachoengsao (1), Chanthaburi (1), Sa Kaeo (1)	Bamrasnaradura Institute (1), Nikompattana Hospital (1), Buddhasothorn Hospital (1), Prapokklao Hospital (1), Khlonghat Hospital (1)	2015 (3), 2017 (2)
49	C49	3	**Yes, Clade8**	MDR (3)	Eastern (3)	Rayong (2), Sa Kaeo (1)	Rayong Hospital (2), Wangnamyen Hospital (1)	2014 (2), 2015 (1)
50	C50	3	**Yes, Clade8**	MDR (3)	Eastern (3)	Chon Buri (1), Chanthaburi (2)	Chonburi Hospital (1), Khlung Hospital (1), Prapokklao Hospital (1)	2014 (1), 2015 (1), 2016 (1)
51	C51	2	**Yes, Clade8**	MDR (2)	Central (1), Eastern (1)	Uthai Thani (1), Chon Buri (1)	Nongchang Hospital (1), Banglamung Hospital (1)	2014 (1), 2016 (1)

Supplementary Table 6 Characteristics of 89 (C1-C89) DR-TB clusters defined only by SNP pairwise differences ≤11 (Cont.)

No.	Name of cluster	Number of isolates	Member of major clade	Drug resistant types* (no.)	Geographical link			Time link, year (no.)
					Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
52	C52	2	**Yes, Clade9**	MDR (2)	Eastern (2)	Chon Buri (1), Sa Kaeo (1)	Chonburi Hospital (1), Sakaeo Hospital (1)	2014 (1), 2017 (1)
53	C53	2	**Yes, Clade9**	MDR (2)	Central (2)	Bangkok (1), Sing Buri (1)	Devision of Tuberculosis (1), Singburi Hospital (1)	2014 (1), 2016 (1)
54	C54	2	**Yes, Clade9**	MDR (1), pre-XDR (1)	Central (2)	Bangkok (1), Suphan Buri (1)	Taksin Hospital (1), Chaophrayayommarat Hospital (1)	2014 (1), 2017 (1)
55	C55	2	No	MDR (2)	Central (2)	Samut Sakhon (2)	Samutsakhon Hospital (2)	2014 (1), 2017 (1)
56	C56	2	**Yes, Clade10**	MDR (2)	Southern (2)	Songkhla (1), Satun (1)	Songkhla Hospital (1), Satun Hospital (1)	2015 (1), 2017 (1)
57	C57	4	**Yes, Clade10**	MDR (4)	Northeastern (1), Central (2), Western (1)	Nakhon Ratchasima (1), Nonthaburi (1), Samut Prakan (1), Phetchaburi (1)	Nonthai Hospital (1), Pranangklao Hospital (1), Samutprakan Hospital (1), Cha-am Hospital (1)	2014(2), 2016 (1), 2017 (1)
58	C58	2	No	MDR (1), pre-XDR (1)	Northeastern (2)	Nong Khai (2)	Nongkhai Hospital (2)	2016 (1), 2017 (1)
59	C59	5	**Yes, Clade11**	MDR (2), pre-XDR (2), XDR (1)	Northeastern (2), Central (1), Eastern (2)	Loei (1), Khon Kaen (1), Kamphaeng Phet (1), Chon Buri (2)	Wangsaphung Hospital (1), Khonkaen Hospital (1), Kamphaengphet Hospital (1), Chonburi Hospital (2)	2014 (2), 2016 (2), 2017 (1)
60	C60	3	**Yes, Clade11**	MDR (2), pre-XDR (1)	Northeastern (3)	Nakhon Ratchasima (1), Chaiyaphum (2)	Office of Disease Prevention & Control 9 Nakhon Ratchasima (1), Chaiyaphum Hospital (1), Phukieo Hospital (1)	2015 (1), 2016 (1), 2017 (1)
61	C61	4	**Yes, Clade11**	MDR (4)	Northeastern (3), Central (1)	Buri Ram (2), Nakhon Ratchasima (1), Sukhothai (1)	Buriram Hospital (1), Nangrong Hospital (1), The Golden Gate Hospital (1), Sisatchanalai Hospital (1)	2014 (2), 2015 (2)
62	C62	4	**Yes, Clade11**	MDR (4)	Northeastern (1), Central (2), Western (1)	Si Sa Ket (1), Bangkok (1), Samut Prakan (1), Kanchanaburi (1)	Kantharalak Hospital (1), Public Health Center 4 (1), Bangbo Hospital (1), Paholpolpayuhasena Hospital (1)	2015 (2), 2017 (2)
63	C63	10	**Yes, Clade11**	MDR (9), pre-XDR (1)	Northeastern (1), Central (2), Weastern (7)	Nong Khai (1), Bangkok (1), Samut Prakan (1), Phetchaburi (7)	Nongkhai Hospital (1), Charoenkrung Pracharak Hospital (1), Bangbo Hospital (1), Phrachomklao Hospital (7)	2014 (1), 2015 (2), 2016 (4), 2017 (3)
64	C64	2	**Yes, Clade12**	MDR (2)	Northeastern (1), Central (1)	Udon Thani (1), Bangkok (1)	Udonthani Hospital (1), Public Health Center 28 (1)	2014 (1), 2017 (1)
65	C65	2	**Yes, Clade12**	MDR (2)	Northeastern (2)	Si Sa Ket (1), Roi Et (1)	Sisaket Hospital (1), Roi-et Hospital (1)	2015 (1), 2017 (1)
66	C66	2	**Yes, Clade12**	MDR (2)	Eastern (2)	Chon Buri (1), Chanthaburi (1)	Chonburi Hospital (1), Prapokklao Hospital (1)	2014 (1), 2016 (1)

Supplementary Table 6 Characteristics of 89 (C1-C89) DR-TB clusters defined only by SNP pairwise differences ≤11 (Cont.)

No.	Name of cluster	Number of isolates	Member of major clade	Drug resistant types* (no.)	Geographical link			Time link, year (no.)
					Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
67	C67	2	**Yes, Clade12**	MDR (2)	Eastern (1), Northern (1)	Chon Buri (1), Chiang Mai (1)	Chonburi Hospital (1), Office of Disease Prevention & Control 1 Chiangmai (1)	2017 (2)
68	C68	2	**Yes, Clade12**	MDR (2)	Central (1), Eastern (1)	Phitsanulok (1), Chon Buri (1)	Buddhachinaraj Hospital (1), Banglamung Hospital (1)	2014 (1), 2017 (1)
69	C69	2	**Yes, Clade12**	MDR (2)	Central (2)	Bangkok (2)	Taksin Hospital (1), Public Health Center 36 (1)	2016 (2)
70	C70	3	**Yes, Clade12**	MDR (1), XDR (2)	Northeastern (1), Southern (2)	Kalasin (1), Ranong (2)	Khammuang Hospital (1), Ranong Hospital (2)	2014 (1), 2016 (2)
71	C71	8	**Yes, Clade12**	MDR (6), pre-XDR (2)	Northeastern (2), Central (2), Eastern (3), Southern (1)	Ubon Ratchathani (1), Udon Thani (1), Bangkok (2), Chon Buri (2), Songkhla (1)	Somdetphrayuphrarat Detudom Hospital (1), Udonthani Hospital (1), Taksin Hospital (1), Nopparat Rajathanee Hospital (1), Chonburi Hospital (3), Hatyai Hospital (1)	2014 (2), 2015 (1), 2016 (2), 2017 (3)
72	C72	7	**Yes, Clade12**	pre-XDR (4), XDR (3)	Northeastern (4),Central (2), Eastern (1)	Ubon Ratchathani (3), Mukdahan (1), Bangkok (2), Trat (1)	Somdetphrayuphrarat Detudom Hospital (2), Warinchamrap Hospital (1), Mukdahan Hospital (1), Klang Hospital (1), Navamin Hospital 9 (1), Trat Hospital (1)	2014 (2), 2015 (1), 2016 (3), 2017 (1)
73	C73	2	**Yes, Clade12**	pre-XDR (1), XDR (1)	Central (2)	Bangkok (2)	Devision of Tuberculosis (1), Chulalongkorn Hospital (1)	2014 (1), 2017 (1)
74	C74	2	**Yes, Clade12**	MDR (2)	Central (2)	Bangkok (1), Sing Buri (1)	Taksin Hospital (1), Singburi Hospital (1)	2015 (1), 2017 (1)
75	C75	2	**Yes, Clade12**	MDR (1), pre-XDR (1)	Northeastern (2)	Ubon Ratchathani (1), Maha Sarakham (1)	Fort sunpasitthiprasong Hospital (1), Nadun Hospital (1)	2016 (2)
76	C76	2	**Yes, Clade12**	MDR (1), pre-XDR (1)	Eastern (2)	Chon Buri (2)	Chonburi Hospital (1), Banglamung Hospital (1)	2014 (2)
77	C77	3	**Yes, Clade12**	MDR (3)	Central (3)	Bangkok (3)	Devision of Tuberculosis (1), Public Health Center 30 (1), Public Health Center 40 (1)	2014 (1), 2015 (2)
78	C78	4	**Yes, Clade12**	MDR (4)	Northeastern (1), Eastern (3)	Chaiyaphum (1), Chon Buri (2), Chachoengsao (1)	Kaengkhro Hospital (1), Chonburi Hospital (2), Buddhasothorn Hospital (1)	2017 (4)
79	C79	2	**Yes, Clade12**	MDR (2)	Eastern (1), Northeastern (1)	Chon Buri (1), Chiang Mai (1)	Phanatnikhom Hospital (1), Office of Disease Prevention & Control 1 Chiangmai (1)	2015 (1), 2017 (1)
80	C80	4	**Yes, Clade12**	MDR (4)	Central (4)	Bangkok (4)	Devision of Tuberculosis (1), Public Health Center 23 (1), Public Health Center 29 (1), Public Health Center 48 (1)	2015 (1), 2017 (3)
81	C81	2	**Yes, Clade13**	MDR (2)	Central (1), Western (1)	Samut Sakhon (1), Kanchanaburi (1)	Samutsakhon Hospital (1), Makarak Hospital (1)	2017 (2)
82	C82	2	**Yes, Clade13**	pre-XDR (2)	Central (1), Westerun (1)	Suphan Buri (1), Kanchanaburi (1)	Uthong Hospital (1), Makarak Hospital (1)	2016 (1), 2017 (1)

Supplementary Table 6 Characteristics of 89 (C1-C89) DR-TB clusters defined only by SNP pairwise differences ≤11 (Cont.)

No.	Name of cluster	Number of isolates	Member of major clade	Drug resistant types* (no.)	Geographical link			Time link, year (no.)
					Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
83	C83	3	**Yes, Clade13**	MDR (3)	Central (2), Eastern (1)	Suphan Buri (2), Chanthaburi (1)	Chaophrayayommarat Hospital (1), Uthong Hospital (1), Prapokklao Hospital (1)	2015 (1), 2016 (1), 2017 (1)
84	C84	2	**Yes, Clade13**	XDR (2)	Eastern (2)	Chon Buri (2)	Chonburi Hospital (2)	2014 (1), 2016 (1)
85	C85	4	**Yes, Clade13**	MDR (4)	Western (4)	Kanchanaburi (4)	Makarak Hospital (3), Paholpolpayuhasena Hospital (1)	2014 (1), 2015 (1), 2017 (2)
86	C86	3	**Yes, Clade13**	XDR (3)	Western (3)	Kanchanaburi (3)	Makarak Hospital (2), Paholpolpayuhasena Hospital (1)	2014 (1), 2015 (1), 2017 (1)
87	C87	3	**Yes, Clade13**	MDR (3)	Western (3)	Kanchanaburi (3)	Makarak Hospital (2), Paholpolpayuhasena Hospital (1)	2014 (2), 2017 (1)
88	C88	5	**Yes, Clade13**	MDR (5)	Central (3), Western (2)	Samut Sakhon (2), Nakhon Pathom (1), Kanchanaburi (2)	Banphaeo Hospital (2), Nakhonpathom Hospital (1), Makarak Hospital (1), Paholpolpayuhasena Hospital (1)	2014 (2), 2015 (1), 2016 (1), 2017 (1)
89	C89	34	**Yes, Clade13**	MDR (27), pre-XDR (7)	Central (7), Southern (1), Western (26)	Bangkok (3), Suphan Buri (2), Samut Sakhon (1), Phitsanulok (1), Surat Thani (1), Kanchanaburi (25), Ratchaburi (1)	Rajavithi Hospital (1), Klang Hospital (2), Danchang Hospital (1), Somdetphrasangkharat 17 Hospital (1), Banphaeo Hospital (1), Buddhachinaraj Hospital (1), Suratthani Hospital (1), Makarak Hospital (15), Paholpolpayuhasena Hospital (7), Danmakhamtia Hospital (1), Saiyok Hospital (1), Somdetphrasangkharat 19 Hospital (1), Banpong Hospital (1)	2014 (7), 2015 (8), 2016 (4), 2017 (15)

*DR-TB types (MDR, pre-XDR and XDR) were based on phenotypic DST.

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters

No.	Cluster[1]	Cluster[2]	Clade	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
No.	Cluster[1]	Cluster[2]	Clade	C1	C1	C2	C2	C2	C3	C3	C4	C4	C4	C5	C5	C5	C6	C6
				M1	M1	M2	M2	pre-XDR	M3	M3	M4	M4	M4	M5	M5	M5	M6	M6
				NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
1	C1	M1	NA	0	3	1679	1674	1680	1701	1702	1690	1690	1693	1726	1727	1723	1715	1714
2	C1	M1	NA	3	0	1680	1675	1681	1702	1703	1691	1691	1694	1727	1728	1724	1716	1715
3	C2	M2	NA	1679	1680	0	7	9	300	301	705	705	708	742	743	739	731	730
4	C2	M2	NA	1674	1675	7	0	10	295	296	700	700	703	737	738	734	726	725
5	C2	pre-XDR	NA	1680	1681	9	10	0	301	302	706	706	709	743	744	740	732	731
6	C3	M3	NA	1701	1702	300	295	301	0	3	727	727	730	762	763	759	751	750
7	C3	M3	NA	1702	1703	301	296	302	3	0	728	728	731	763	764	760	752	751
8	C4	M4	NA	1690	1691	705	700	706	727	728	0	10	11	653	654	650	642	641
9	C4	M4	NA	1690	1691	705	700	706	727	728	10	0	11	653	654	650	642	641
10	C4	M4	NA	1693	1694	708	703	709	730	731	11	11	0	656	657	653	645	644
11	C5	M5	NA	1726	1727	742	737	743	762	763	653	653	656	0	1	3	31	30
12	C5	M5	NA	1727	1728	743	738	744	763	764	654	654	657	1	0	4	32	31
13	C5	M5	NA	1723	1724	739	734	740	759	760	650	650	653	3	4	0	28	27
14	C6	M6	NA	1715	1716	731	726	732	751	752	642	642	645	31	32	28	0	11
15	C6	M6	NA	1714	1715	730	725	731	750	751	641	641	644	30	31	27	11	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	16	17	18	19	20	21	22	23	24	25	26	27
No.	Cluster[1]	Cluster[2]	Clade	C7	C7	C8	C8	C9	C9	C9	C10	C10	C10	C10	C10
				MDR	pre-XDR	M7	M7	X1	X1	MDR	M8	M8	M8	XDR	pre-XDR
				NA	NA	NA	NA	NA	NA	NA	Clade1	Clade1	Clade1	Clade1	Clade1
16	C7	MDR	NA	0	3	325	324	326	321	322	353	352	352	353	355
17	C7	pre-XDR	NA	3	0	326	327	329	324	325	356	355	355	354	356
18	C8	M7	NA	325	326	0	1	217	212	213	244	243	243	242	244
19	C8	M7	NA	324	327	1	0	216	211	212	243	242	242	243	245
20	C9	X1	NA	326	329	217	216	0	7	8	205	204	204	205	207
21	C9	X1	NA	321	324	212	211	7	0	3	200	199	199	200	202
22	C9	MDR	NA	322	325	213	212	8	3	0	201	200	200	201	203
23	C10	M8	Clade1	353	356	244	243	205	200	201	0	1	1	2	4
24	C10	M8	Clade1	352	355	243	242	204	199	200	1	0	0	1	3
25	C10	M8	Clade1	352	355	243	242	204	199	200	1	0	0	1	3
26	C10	XDR	Clade1	353	354	242	243	205	200	201	2	1	1	0	2
27	C10	pre-XDR	Clade1	355	356	244	245	207	202	203	4	3	3	2	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	No.	28	29	30	31	32	33	34	35	36	37	38
			Cluster[1]	C11	C11	C12	C12	C13	C13	C13	C14	C14	C15	C15
			Cluster[2]	P1	P1	M9	M9	M10	M10	M10	M11	M11	M12	M12
			Clade	NA	NA	NA	NA	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2
28	C11	P1	NA	0	0	482	486	458	460	461	466	458	458	459
29	C11	P1	NA	0	0	482	486	458	460	461	466	458	458	459
30	C12	M9	NA	482	482	0	4	330	332	333	338	330	330	331
31	C12	M9	NA	486	486	4	0	334	336	337	342	334	334	335
32	C13	M10	Clade2	458	458	330	334	0	2	3	16	8	8	9
33	C13	M10	Clade2	460	460	332	336	2	0	3	16	10	8	9
34	C13	M10	Clade2	461	461	333	337	3	3	0	17	11	9	10
35	C14	M11	Clade2	466	466	338	342	16	16	17	0	8	14	15
36	C14	M11	Clade2	458	458	330	334	8	10	11	8	0	8	9
37	C15	M12	Clade2	458	458	330	334	8	8	9	14	8	0	3
38	C15	M12	Clade2	459	459	331	335	9	9	10	15	9	3	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	No.	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54
			Cluster[1]	C16	C16	C16	C16	C16	NA	C17	C17	C18	C18	C19	C19	C19	C19	C20	C20
			Cluster[2]	P2	P2	P2	M13	M13	M13	M14	M14	M15	M15	M16	M16	M16	M16	X2	X2
			Clade	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	Clade2	NA	NA
39	C16	P2	Clade2	0	8	5	5	5	12	21	18	16	13	12	16	19	20	347	347
40	C16	P2	Clade2	8	0	5	7	3	6	21	20	16	15	14	18	21	22	349	349
41	C16	P2	Clade2	5	5	0	4	2	9	20	17	15	12	11	15	18	19	348	348
42	C16	M13	Clade2	5	7	4	0	4	11	18	15	13	10	9	13	16	17	346	346
43	C16	M13	Clade2	5	3	2	4	0	7	20	17	15	12	11	15	18	19	348	348
44	NA	M13	Clade2	12	6	9	11	7	0	25	24	20	19	18	22	25	26	353	353
45	C17	M14	Clade2	21	21	20	18	20	25	0	7	17	16	15	19	22	23	350	350
46	C17	M14	Clade2	18	20	17	15	17	24	7	0	16	13	12	16	19	20	349	349
47	C18	M15	Clade2	16	16	15	13	15	20	17	16	0	7	6	10	13	14	345	345
48	C18	M15	Clade2	13	15	12	10	12	19	16	13	7	0	3	7	10	11	344	344
49	C19	M16	Clade2	12	14	11	9	11	18	15	12	6	3	0	6	9	10	343	343
50	C19	M16	Clade2	16	18	15	13	15	22	19	16	10	7	6	0	3	4	347	347
51	C19	M16	Clade2	19	21	18	16	18	25	22	19	13	10	9	3	0	7	350	350
52	C19	M16	Clade2	20	22	19	17	19	26	23	20	14	11	10	4	7	0	351	351
53	C20	X2	NA	347	349	348	346	348	353	350	349	345	344	343	347	350	351	0	2
54	C20	X2	NA	347	349	348	346	348	353	350	349	345	344	343	347	350	351	2	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	55	56	57	58	59	60	61	62	63	64	65	66	67	68
			No.	55	56	57	58	59	60	61	62	63	64	65	66	67	68
			Cluster[1]	C21	C21	C22	C22	C23	C23	C24	C24	C25	C25	C26	C26	C27	C27
			Cluster[2]	pre-XDR	MDR	M17	M17	X3	X3	M18	M18	M19	M19	M20	M20	M21	M21
55	C21	pre-XDR	NA	0	8	206	206	218	219	217	222	209	212	211	208	211	211
56	C21	MDR	NA	8	0	208	208	220	221	219	224	211	214	213	210	213	213
57	C22	M17	NA	206	208	0	2	200	201	199	204	191	194	193	190	193	193
58	C22	M17	NA	206	208	2	0	200	201	199	204	191	194	193	190	193	193
59	C23	X3	Clade3	218	220	200	200	0	9	15	20	23	26	203	200	201	201
60	C23	X3	Clade3	219	221	201	201	9	0	16	21	24	27	204	201	202	202
61	C24	M18	Clade3	217	219	199	199	15	16	0	9	22	25	202	199	200	200
62	C24	M18	Clade3	222	224	204	204	20	21	9	0	27	30	207	204	205	205
63	C25	M19	Clade3	209	211	191	191	23	24	22	27	0	11	194	191	192	192
64	C25	M19	Clade3	212	214	194	194	26	27	25	30	11	0	197	194	195	195
65	C26	M20	NA	211	213	193	193	203	204	202	207	194	197	0	5	150	150
66	C26	M20	NA	208	210	190	190	200	201	199	204	191	194	5	0	147	147
67	C27	M21	NA	211	213	193	193	201	202	200	205	192	195	150	147	0	0
68	C27	M21	NA	211	213	193	193	201	202	200	205	192	195	150	147	0	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	69	70	71	72	73	74	75	76	77	78	79	80	81
			No.	69	70	71	72	73	74	75	76	77	78	79	80	81
			Cluster[1]	C28	C28	C28	C29	C29	C29	C29	C30	C30	C30	C31	C31	C31
			Cluster[2]	M22	M22	M22	M23	M23	M23	M23	M24	M24	M24	M25	M25	M25
69	C28	M22	Clade4	0	6	7	236	231	233	231	230	230	230	202	204	203
70	C28	M22	Clade4	6	0	9	238	233	235	233	232	232	232	204	206	205
71	C28	M22	Clade4	7	9	0	239	234	236	234	233	233	233	205	207	206
72	C29	M23	NA	236	238	239	0	7	9	7	242	242	242	214	216	215
73	C29	M23	NA	231	233	234	7	0	4	2	237	237	237	209	211	210
74	C29	M23	NA	233	235	236	9	4	0	4	239	239	239	211	213	212
75	C29	M23	NA	231	233	234	7	2	4	0	237	237	237	209	211	210
76	C30	M24	NA	230	232	233	242	237	239	237	0	0	0	200	202	201
77	C30	M24	NA	230	232	233	242	237	239	237	0	0	0	200	202	201
78	C30	M24	NA	230	232	233	242	237	239	237	0	0	0	200	202	201
79	C31	M25	Clade5	202	204	205	214	209	211	209	200	200	200	0	10	9
80	C31	M25	Clade5	204	206	207	216	211	213	211	202	202	202	10	0	11
81	C31	M25	Clade5	203	205	206	215	210	212	210	201	201	201	9	11	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

			No.	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97
			Cluster[1]	C32	C32	C33	C33	C34	C34	C35	C35	C36	C36	C36	C36	C36	C37	C37	C37
			Cluster[2]	X4	X4	M26	M26	M27	M27	X5	X5	M28	M28	M28	M28	pre-XDR	P3	P3	MDR
No.	**Cluster[1]**	**Cluster[2]**	**Clade**	NA	NA	NA	NA	NA	NA	Clade6	Clade6	Clade6	Clade6	Clade6	Clade6	Clade6	NA	NA	NA
82	C32	X4	NA	0	9	103	103	104	106	107	108	108	105	110	106	108	192	192	190
83	C32	X4	NA	9	0	102	102	103	105	106	107	107	104	109	105	107	191	191	189
84	C33	M26	NA	103	102	0	0	97	99	100	101	101	98	103	99	101	185	185	183
85	C33	M26	NA	103	102	0	0	97	99	100	101	101	98	103	99	101	185	185	183
86	C34	M27	NA	104	103	97	97	0	6	73	74	74	71	76	72	74	186	186	184
87	C34	M27	NA	106	105	99	99	6	0	75	76	76	73	78	74	76	188	188	186
88	C35	X5	Clade6	107	106	100	100	73	75	0	9	5	2	7	3	5	189	189	187
89	C35	X5	Clade6	108	107	101	101	74	76	9	0	10	7	12	8	10	190	190	188
90	C36	M28	Clade6	108	107	101	101	74	76	5	10	0	3	8	4	6	190	190	188
91	C36	M28	Clade6	105	104	98	98	71	73	2	7	3	0	5	1	3	187	187	185
92	C36	M28	Clade6	110	109	103	103	76	78	7	12	8	5	0	6	8	192	192	190
93	C36	M28	Clade6	106	105	99	99	72	74	3	8	4	1	6	0	4	188	188	186
94	C36	pre-XDR	Clade6	108	107	101	101	74	76	5	10	6	3	8	4	0	190	190	188
95	C37	P3	NA	192	191	185	185	186	188	189	190	190	187	192	188	190	0	0	2
96	C37	P3	NA	192	191	185	185	186	188	189	190	190	187	192	188	190	0	0	2
97	C37	MDR	NA	190	189	183	183	184	186	187	188	188	185	190	186	188	2	2	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

			No.	98	99	100	101	102	103	104	105	106	107	108
			Cluster[1]	C38	C38	C39	C39	C39	C40	C40	C41	C41	C42	C42
			Cluster[2]	M29	M29	M30	M30	M30	MDR	pre-XDR	M31	M31	M32	M32
No.	**Cluster[1]**	**Cluster[2]**	**Clade**	Clade7	Clade7	Clade7	Clade7	Clade7	Clade7	Clade7	Clade7	Clade7	Clade7	Clade7
98	C38	M29	Clade7	0	4	40	41	38	36	42	35	34	38	38
99	C38	M29	Clade7	4	0	40	41	38	36	42	35	34	38	38
100	C39	M30	Clade7	40	40	0	9	2	18	24	17	16	20	20
101	C39	M30	Clade7	41	41	9	0	7	19	25	18	17	21	21
102	C39	M30	Clade7	38	38	2	7	0	16	22	15	14	18	18
103	C40	MDR	Clade7	36	36	18	19	16	0	10	13	12	16	16
104	C40	pre-XDR	Clade7	42	42	24	25	22	10	0	19	18	22	22
105	C41	M31	Clade7	35	35	17	18	15	13	19	0	7	13	13
106	C41	M31	Clade7	34	34	16	17	14	12	18	7	0	12	12
107	C42	M32	Clade7	38	38	20	21	18	16	22	13	12	0	0
108	C42	M32	Clade7	38	38	20	21	18	16	22	13	12	0	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123
			No.	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123
			Cluster[1]	C43	C43	C44	C44	C45	C45	C46	C46	C47	C47	C48	C48	C48	C48	C48
			Cluster[2]	XDR	pre-XDR	M33	M33	M34	M34	X6	X6	P4	P4	M35	M35	M35	M35	M35
109	C43	XDR	NA	0	0	25	29	167	174	201	201	173	172	186	187	186	186	185
110	C43	pre-XDR	NA	0	0	25	29	167	174	201	201	173	172	186	187	186	186	185
111	C44	M33	NA	25	25	0	8	160	167	194	194	166	165	177	178	177	177	176
112	C44	M33	NA	29	29	8	0	164	171	198	198	170	169	181	182	181	181	180
113	C45	M34	NA	167	167	160	164	0	7	172	172	145	144	157	158	157	157	156
114	C45	M34	NA	174	174	167	171	7	0	179	179	152	151	164	165	164	164	163
115	C46	X6	NA	201	201	194	198	172	179	0	0	173	172	185	186	185	185	184
116	C46	X6	NA	201	201	194	198	172	179	0	0	173	172	185	186	185	185	184
117	C47	P4	NA	173	173	166	170	145	152	173	173	0	5	158	159	158	158	157
118	C47	P4	NA	172	172	165	169	144	151	172	172	5	0	157	158	157	157	156
119	C48	M35	Clade8	186	186	177	181	157	164	185	185	158	157	0	11	8	8	7
120	C48	M35	Clade8	187	187	178	182	158	165	186	186	159	158	11	0	11	11	10
121	C48	M35	Clade8	186	186	177	181	157	164	185	185	158	157	8	11	0	0	7
122	C48	M35	Clade8	186	186	177	181	157	164	185	185	158	157	8	11	0	0	7
123	C48	M35	Clade8	185	185	176	180	156	163	184	184	157	156	7	10	7	7	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	124	125	126	127	128	129	130	131	132	133	134	135
			No.	124	125	126	127	128	129	130	131	132	133	134	135
			Cluster[1]	C49	C49	C49	C50	C50	C50	C51	C51	C52	C52	C53	C53
			Cluster[2]	M36	M36	M36	M37	M37	M37	M38	M38	M39	M39	M40	M40
124	C49	M36	Clade8	0	3	7	11	12	11	3	7	212	209	212	213
125	C49	M36	Clade8	3	0	8	12	13	12	4	8	213	210	213	214
126	C49	M36	Clade8	7	8	0	16	17	16	8	12	217	214	217	218
127	C50	M37	Clade8	11	12	16	0	11	0	12	16	221	218	221	222
128	C50	M37	Clade8	12	13	17	11	0	11	13	17	222	219	222	223
129	C50	M37	Clade8	11	12	16	0	11	0	12	16	221	218	221	222
130	C51	M38	Clade8	3	4	8	12	13	12	0	6	213	210	213	214
131	C51	M38	Clade8	7	8	12	16	17	16	6	0	217	214	217	218
132	C52	M39	Clade9	212	213	217	221	222	221	213	217	0	11	14	15
133	C52	M39	Clade9	209	210	214	218	219	218	210	214	11	0	7	8
134	C53	M40	Clade9	212	213	217	221	222	221	213	217	14	7	0	11
135	C53	M40	Clade9	213	214	218	222	223	222	214	218	15	8	11	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	136 C54 MDR Clade9	137 C54 pre-XDR Clade9	138 C55 M41 NA	139 C55 M41 NA	140 C56 M42 Clade10	141 C56 M42 Clade10	142 C57 M43 Clade10	143 C57 M43 Clade10	144 C57 M43 Clade10	145 C57 M43 Clade10	146 C58 pre-XDR NA	147 C58 MDR NA
136	C54	MDR	Clade9	0	4	163	161	166	169	168	165	166	167	168	170
137	C54	pre-XDR	Clade9	4	0	161	159	164	167	166	163	164	165	166	168
138	C55	M41	NA	163	161	0	2	131	134	133	130	131	132	131	133
139	C55	M41	NA	161	159	2	0	129	132	131	128	129	130	129	131
140	C56	M42	Clade10	166	164	131	129	0	11	20	17	18	19	136	138
141	C56	M42	Clade10	169	167	134	132	11	0	23	20	21	22	139	141
142	C57	M43	Clade10	168	166	133	131	20	23	0	9	10	5	138	140
143	C57	M43	Clade10	165	163	130	128	17	20	9	0	7	8	135	137
144	C57	M43	Clade10	166	164	131	129	18	21	10	7	0	9	136	138
145	C57	M43	Clade10	167	165	132	130	19	22	5	8	9	0	137	139
146	C58	pre-XDR	NA	168	166	131	129	136	139	138	135	136	137	0	2
147	C58	MDR	NA	170	168	133	131	138	141	140	137	138	139	2	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	148 C59 M44 Clade11	149 C59 M44 Clade11	150 C59 P5 Clade11	151 C59 P5 Clade11	152 C59 XDR Clade11	153 C60 M45 Clade11	154 C60 M45 Clade11	155 C60 pre-XDR Clade11	156 C61 M46 Clade11	157 C61 M46 Clade11	158 C61 M46 Clade11	159 C61 M46 Clade11
148	C59	M44	Clade11	0	10	6	8	7	22	21	24	21	24	21	22
149	C59	M44	Clade11	10	0	4	6	3	26	25	28	25	28	25	26
150	C59	P5	Clade11	6	4	0	2	1	22	21	24	21	24	21	22
151	C59	P5	Clade11	8	6	2	0	3	24	23	26	23	26	23	24
152	C59	XDR	Clade11	7	3	1	3	0	23	22	25	22	25	22	23
153	C60	M45	Clade11	22	26	22	24	23	0	1	6	7	10	7	8
154	C60	M45	Clade11	21	25	21	23	22	1	0	5	6	9	6	7
155	C60	pre-XDR	Clade11	24	28	24	26	25	6	5	0	9	12	9	10
156	C61	M46	Clade11	21	25	21	23	22	7	6	9	0	3	0	1
157	C61	M46	Clade11	24	28	24	26	25	10	9	12	3	0	3	4
158	C61	M46	Clade11	21	25	21	23	22	7	6	9	0	3	0	1
159	C61	M46	Clade11	22	26	22	24	23	8	7	10	1	4	1	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	160	161	162	163	164	165	166	167	168	169	170	171	172	173
			No.	160	161	162	163	164	165	166	167	168	169	170	171	172	173
			Cluster[1]	C62	C62	C62	C62	C63	C63	C63	C63	C63	C63	C63	C63	C63	C63
			Cluster[2]	M47	M47	M47	M47	M48	M48	M48	M48	M48	M48	M48	M48	M48	pre-XDR
160	C62	M47	Clade11	0	5	4	7	12	10	10	10	9	11	10	9	9	9
161	C62	M47	Clade11	5	0	3	6	11	9	9	9	8	10	9	8	8	8
162	C62	M47	Clade11	4	3	0	5	10	8	8	8	7	9	8	7	7	7
163	C62	M47	Clade11	7	6	5	0	9	7	7	7	6	8	7	6	6	6
164	C63	M48	Clade11	12	11	10	9	0	6	6	6	11	7	6	9	5	3
165	C63	M48	Clade11	10	9	8	7	6	0	4	4	9	1	0	7	3	3
166	C63	M48	Clade11	10	9	8	7	6	4	0	2	9	5	4	7	3	3
167	C63	M48	Clade11	10	9	8	7	6	4	2	0	9	5	4	7	3	3
168	C63	M48	Clade11	9	8	7	6	11	9	9	9	0	10	9	8	8	8
169	C63	M48	Clade11	11	10	9	8	7	1	5	5	10	0	1	8	4	4
170	C63	M48	Clade11	10	9	8	7	6	0	4	4	9	1	0	7	3	3
171	C63	M48	Clade11	9	8	7	6	9	7	7	7	8	8	7	0	6	6
172	C63	M48	Clade11	9	8	7	6	5	3	3	3	8	4	3	6	0	2
173	C63	pre-XDR	Clade11	9	8	7	6	3	3	3	3	8	4	3	6	2	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188
			No.	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188
			Cluster[1]	C64	C64	C65	C65	C66	C66	C67	C67	C68	C68	C69	C69	C70	C70	C70
			Cluster[2]	M49	M49	M50	M50	M51	M51	M52	M52	M53	M53	M54	M54	X7	X7	MDR
174	C64	M49	Clade12	0	6	25	24	26	27	31	33	33	28	40	35	32	33	29
175	C64	M49	Clade12	6	0	31	30	32	33	37	39	39	34	46	41	38	39	35
176	C65	M50	Clade12	25	31	0	9	11	12	16	18	18	13	25	20	17	18	14
177	C65	M50	Clade12	24	30	9	0	10	11	15	17	17	12	24	19	16	17	13
178	C66	M51	Clade12	26	32	11	10	0	9	17	19	19	14	26	21	18	19	15
179	C66	M51	Clade12	27	33	12	11	9	0	18	20	20	15	27	22	19	20	16
180	C67	M52	Clade12	31	37	16	15	17	18	0	8	24	19	31	26	23	24	20
181	C67	M52	Clade12	33	39	18	17	19	20	8	0	26	21	33	28	25	26	22
182	C68	M53	Clade12	33	39	18	17	19	20	24	26	0	11	33	28	25	26	22
183	C68	M53	Clade12	28	34	13	12	14	15	19	21	11	0	28	23	20	21	17
184	C69	M54	Clade12	40	46	25	24	26	27	31	33	33	28	0	9	32	33	29
185	C69	M54	Clade12	35	41	20	19	21	22	26	28	28	23	9	0	27	28	24
186	C70	X7	Clade12	32	38	17	16	18	19	23	25	25	20	32	27	0	1	9
187	C70	X7	Clade12	33	39	18	17	19	20	24	26	26	21	33	28	1	0	10
188	C70	MDR	Clade12	29	35	14	13	15	16	20	22	22	17	29	24	9	10	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203
			No.	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203
			Cluster[1]	C71	C71	C71	C71	C71	C71	C71	C71	C72	C72	C72	C72	C72	C72	C72
			Cluster[2]	M55	M55	M55	M55	M55	M55	P6	P6	P7	P7	P7	P7	X8	X8	X8
189	C71	M55	Clade12	0	4	4	7	7	10	6	4	26	27	28	30	26	28	27
190	C71	M55	Clade12	4	0	2	5	5	8	4	2	24	25	26	28	24	26	25
191	C71	M55	Clade12	4	2	0	5	5	8	4	2	24	25	26	28	24	26	25
192	C71	M55	Clade12	7	5	5	0	8	11	7	5	27	28	29	31	27	29	28
193	C71	M55	Clade12	7	5	5	8	0	11	7	5	27	28	29	31	27	29	28
194	C71	M55	Clade12	10	8	8	11	11	0	10	8	28	29	30	32	28	30	29
195	C71	P6	Clade12	6	4	4	7	7	10	0	4	26	27	28	30	26	28	27
196	C71	P6	Clade12	4	2	2	5	5	8	4	0	24	25	26	28	24	26	25
197	C72	P7	Clade12	26	24	24	27	27	28	26	24	0	3	6	6	4	4	5
198	C72	P7	Clade12	27	25	25	28	28	29	27	25	3	0	7	7	5	5	6
199	C72	P7	Clade12	28	26	26	29	29	30	28	26	6	7	0	10	6	8	7
200	C72	P7	Clade12	30	28	28	31	31	32	30	28	6	7	10	0	8	8	9
201	C72	X8	Clade12	26	24	24	27	27	28	26	24	4	5	6	8	0	6	5
202	C72	X8	Clade12	28	26	26	29	29	30	28	26	4	5	8	8	6	0	7
203	C72	X8	Clade12	27	25	25	28	28	29	27	25	5	6	7	9	5	7	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Clade	204	205	206	207	208	209	210	211	212	213	214
			No.	204	205	206	207	208	209	210	211	212	213	214
			Cluster[1]	C73	C73	C74	C74	C75	C75	C76	C76	C77	C77	C77
			Cluster[2]	X8	pre-XDR	M56	M56	MDR	pre-XDR	MDR	pre-XDR	M57	M57	M57
204	C73	X8	Clade12	0	0	32	27	33	28	29	31	33	33	33
205	C73	pre-XDR	Clade12	0	0	32	27	33	28	29	31	33	33	33
206	C74	M56	Clade12	32	32	0	11	21	16	19	21	23	23	23
207	C74	M56	Clade12	27	27	11	0	18	13	16	18	20	20	20
208	C75	MDR	Clade12	33	33	21	18	0	11	20	22	24	24	24
209	C75	pre-XDR	Clade12	28	28	16	13	11	0	15	17	19	19	19
210	C76	MDR	Clade12	29	29	19	16	20	15	0	4	18	18	18
211	C76	pre-XDR	Clade12	31	31	21	18	22	17	4	0	20	20	20
212	C77	M57	Clade12	33	33	23	20	24	19	18	20	0	4	4
213	C77	M57	Clade12	33	33	23	20	24	19	18	20	4	0	4
214	C77	M57	Clade12	33	33	23	20	24	19	18	20	4	4	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

			No.	215	216	217	218	219	220	221	222	223	224	225	226	227	228
			Cluster[1]	C78	C78	C78	C78	C79	C79	C80	C80	C80	C80	C81	C81	C82	C82
			Cluster[2]	M58	M58	M58	M58	M59	M59	M60	M60	M60	M60	M61	M61	P8	P8
No.	Cluster[1]	Cluster[2]	Clade	Clade12	Clade12	Clade12	Clade12	Clade12	Clade12	Clade12	Clade12	Clade12	Clade12	Clade13	Clade13	Clade13	Clade13
215	C78	M58	Clade12	0	2	3	7	16	10	11	11	14	15	23	23	27	21
216	C78	M58	Clade12	2	0	1	7	16	10	11	11	14	15	23	23	27	21
217	C78	M58	Clade12	3	1	0	8	17	11	12	12	15	16	24	24	28	22
218	C78	M58	Clade12	7	7	8	0	19	13	14	14	17	18	26	26	30	24
219	C79	M59	Clade12	16	16	17	19	0	8	9	9	12	13	21	21	25	19
220	C79	M59	Clade12	10	10	11	13	8	0	1	1	4	7	15	15	19	13
221	C80	M60	Clade12	11	11	12	14	9	1	0	0	3	8	16	16	20	14
222	C80	M60	Clade12	11	11	12	14	9	1	0	0	3	8	16	16	20	14
223	C80	M60	Clade12	14	14	15	17	12	4	3	3	0	11	19	19	23	17
224	C80	M60	Clade12	15	15	16	18	13	7	8	8	11	0	20	20	24	18
225	C81	M61	Clade13	23	23	24	26	21	15	16	16	19	20	0	0	18	12
226	C81	M61	Clade13	23	23	24	26	21	15	16	16	19	20	0	0	18	12
227	C82	P8	Clade13	27	27	28	30	25	19	20	20	23	24	18	18	0	10
228	C82	P8	Clade13	21	21	22	24	19	13	14	14	17	18	12	12	10	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

			No.	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243
			Cluster[1]	C83	C83	C83	C84	C84	C85	C85	C85	C85	C86	C86	C86	C87	C87	C87
			Cluster[2]	M62	M62	M62	X9	X9	M63	M63	M63	M63	X10	X10	X10	M64	M64	M64
No.	Cluster[1]	Cluster[2]	Clade	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13
229	C83	M62	Clade13	0	9	8	15	18	9	9	16	11	10	17	11	9	16	14
230	C83	M62	Clade13	9	0	11	18	21	12	12	19	14	13	20	14	12	19	17
231	C83	M62	Clade13	8	11	0	17	20	11	11	18	13	12	19	13	11	18	16
232	C84	X9	Clade13	15	18	17	0	7	10	10	17	12	11	18	12	10	17	15
233	C84	X9	Clade13	18	21	20	7	0	13	13	20	15	14	21	15	13	20	18
234	C85	M63	Clade13	9	12	11	10	13	0	0	7	2	1	8	2	0	7	5
235	C85	M63	Clade13	9	12	11	10	13	0	0	7	2	1	8	2	0	7	5
236	C85	M63	Clade13	16	19	18	17	20	7	7	0	9	8	15	9	7	14	12
237	C85	M63	Clade13	11	14	13	12	15	2	2	9	0	3	10	4	2	9	7
238	C86	X10	Clade13	10	13	12	11	14	1	1	8	3	0	9	1	1	8	6
239	C86	X10	Clade13	17	20	19	18	21	8	8	15	10	9	0	10	8	15	13
240	C86	X10	Clade13	11	14	13	12	15	2	2	9	4	1	10	0	2	9	7
241	C87	M64	Clade13	9	12	11	10	13	0	0	7	2	1	8	2	0	7	5
242	C87	M64	Clade13	16	19	18	17	20	7	7	14	9	8	15	9	7	0	2
243	C87	M64	Clade13	14	17	16	15	18	5	5	12	7	6	13	7	5	2	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

			No.	244	245	246	247	248	249	250	251	252	253	254	255
			Cluster[1]	C88	C88	C88	C88	C88	C89	C89	C89	C89	C89	C89	C89
			Cluster[2]	M65	M65	M65	M65	M65	P9	P9	P9	P9	P9	P9	P9
No.	Cluster[1]	Cluster[2]	Clade	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13
244	C88	M65	Clade13	0	1	3	4	7	5	3	3	5	3	3	4
245	C88	M65	Clade13	1	0	4	5	8	6	4	4	6	4	4	5
246	C88	M65	Clade13	3	4	0	5	10	4	6	6	8	6	6	7
247	C88	M65	Clade13	4	5	5	0	11	9	7	7	9	7	7	8
248	C88	M65	Clade13	7	8	10	11	0	12	10	10	12	10	10	11
249	C89	P9	Clade13	5	6	4	9	12	0	8	8	10	8	8	9
250	C89	P9	Clade13	3	4	6	7	10	8	0	6	8	6	6	7
251	C89	P9	Clade13	3	4	6	7	10	8	6	0	8	6	6	7
252	C89	P9	Clade13	5	6	8	9	12	10	8	8	0	8	8	9
253	C89	P9	Clade13	3	4	6	7	10	8	6	6	8	0	6	7
254	C89	P9	Clade13	3	4	6	7	10	8	6	6	8	6	0	7
255	C89	P9	Clade13	4	5	7	8	11	9	7	7	9	7	7	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

			No.	256	257	258	259	260	261	262	263	264	265	266	267	268	269
			Cluster[1]	C89	C89	C89	C89	C89	C89	C89	C89	C89	C89	C89	C89	C89	C89
			Cluster[2]	M66	M66	M66	M66	M66	M66	M66	M66	M66	M66	M66	M66	M66	M66
No.	Cluster[1]	Cluster[2]	Clade	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13	Clade13
256	C89	M66	Clade13	0	6	6	7	7	4	4	6	7	8	4	6	6	8
257	C89	M66	Clade13	6	0	6	7	7	4	4	6	7	8	4	6	6	8
258	C89	M66	Clade13	6	6	0	7	7	4	4	6	7	8	4	6	6	8
259	C89	M66	Clade13	7	7	7	0	8	5	5	7	8	9	5	7	7	9
260	C89	M66	Clade13	7	7	7	8	0	5	5	7	8	9	5	7	7	9
261	C89	M66	Clade13	4	4	4	5	5	0	2	4	5	6	2	4	4	6
262	C89	M66	Clade13	4	4	4	5	5	2	0	4	5	6	2	4	4	6
263	C89	M66	Clade13	6	6	6	7	7	4	4	0	7	8	4	6	6	8
264	C89	M66	Clade13	7	7	7	8	8	5	5	7	0	9	5	7	7	9
265	C89	M66	Clade13	8	8	8	9	9	6	6	8	9	0	6	8	8	10
266	C89	M66	Clade13	4	4	4	5	5	2	2	4	5	6	0	4	4	6
267	C89	M66	Clade13	6	6	6	7	7	4	4	6	7	8	4	0	6	8
268	C89	M66	Clade13	6	6	6	7	7	4	4	6	7	8	4	6	0	8
269	C89	M66	Clade13	8	8	8	9	9	6	6	8	9	10	6	8	8	0

Supplementary Table 7 SNP pairwise distances within clusters and among isolates between clusters (Cont.)

			No.	**270**	**271**	**272**	**273**	**274**	**275**	**276**	**277**	**278**	**279**	**280**	**281**	**282**
			Cluster[1]	**C89**	**C89**	**C89**	**C89**	**C89**	**C89**	**C89**	**C89**	**C89**	**C89**	**C89**	**C89**	**C89**
			Cluster[2]	**M66**	**M66**	**M66**	**M66**	**M66**	**M66**	**M66**	**M66**	**M66**	**M66**	**M66**	**M66**	**M66**
No.	**Cluster[1]**	**Cluster[2]**	**Clade**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**	**Clade13**
270	C89	M66	Clade13	0	5	3	3	4	5	8	6	6	3	6	0	7
271	C89	M66	Clade13	5	0	4	4	5	6	9	7	7	4	5	5	4
272	C89	M66	Clade13	3	4	0	2	3	4	7	5	5	2	5	3	6
273	C89	M66	Clade13	3	4	2	0	3	4	7	5	5	2	5	3	6
274	C89	M66	Clade13	4	5	3	3	0	5	8	6	6	3	6	4	7
275	C89	M66	Clade13	5	6	4	4	5	0	9	7	7	4	7	5	8
276	C89	M66	Clade13	8	9	7	7	8	9	0	10	10	7	10	8	11
277	C89	M66	Clade13	6	7	5	5	6	7	10	0	8	5	8	6	9
278	C89	M66	Clade13	6	7	5	5	6	7	10	8	0	5	8	6	9
279	C89	M66	Clade13	3	4	2	2	3	4	7	5	5	0	5	3	6
280	C89	M66	Clade13	6	5	5	5	6	7	10	8	8	5	0	6	7
281	C89	M66	Clade13	0	5	3	3	4	5	8	6	6	3	6	0	7
282	C89	M66	Clade13	7	4	6	6	7	8	11	9	9	6	7	7	0

Note: Using a pairwise-difference range of 0-11 SNPs, 89 clusters (totaling 281 isolates: minimum cluster size = 2 isolates) could be recognized. When clusters were further defined as consisting only of isolates with the same type of drug resistance, only 85 clusters (255 isolates) were recognized. Among the clusters to disappear were C7, C21 and C75. Some remaining clusters were split (e.g. C16, C59, C89), had fewer members (such as C2 and C63) or members were re-assigned (no.204). An additional isolate (no.44) fell just outside cluster C16 based on SNP differences, but fell within drug-type cluster M13. Therefore, there were 282 isolates fitting the clustering criterion of ≤11 SNP differences, with or without matching the type of drug resistance. Three hundred and nineteen isolates fell into 13 clades (defined by ≤25 pairwise SNP differences among most pairs of isolates). Sixty clusters (both when based on SNP differences only or using the additional criterion of type of drug sensitivity) were included within these clades, along with many non-clustering isolates.

Supplementary Table 8 Distribution of primary and acquired DR-TB among 89 clusters

No.	Cluster[1]	Cluster[2]	Different DR types in cluster	Assumed previous primary transmission events	Assumed acquired resistance status	Primary/acquired DR-TB classification	Chronology of mixed DR types	Phenotypic DST	Genotypic DST
1	C1	M1	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
2	C1	M1	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
3	C2	M2	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
4	C2	M2	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
5	C2	pre-XDR	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
6	C3	M3	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
7	C3	M3	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
8	C4	M4	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
9	C4	M4	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
10	C4	M4	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
11	C5	M5	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
12	C5	M5	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
13	C5	M5	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
14	C6	M6	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
15	C6	M6	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
16	C7	MDR	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
17	C7	pre-XDR	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
18	C8	M7	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
19	C8	M7	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
20	C9	X1	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
21	C9	X1	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
22	C9	MDR	No	Primary XDR	-	Primary XDR	NA	MDR	XDR
23	C10	M8	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
24	C10	M8	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
25	C10	M8	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
26	C10	XDR	Yes	Primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	XDR	pre-XDR
27	C10	pre-XDR	Yes	Primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
28	C11	P1	No	Primary pre-XDR	-	Primary pre-XDR	NA	pre-XDR	pre-XDR
29	C11	P1	No	Primary pre-XDR	-	Primary pre-XDR	NA	pre-XDR	pre-XDR
30	C12	M9	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
31	C12	M9	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
32	C13	M10	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
33	C13	M10	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
34	C13	M10	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
35	C14	M11	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
36	C14	M11	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
37	C15	M12	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
38	C15	M12	No	Primary MDR	-	Primary MDR	NA	MDR	MDR

Supplementary Table 8 Distribution of primary and acquired DR-TB among 89 clusters (Cont.)

No.	Cluster [1]	Cluster[2]	Different DR types in cluster	Assumed previous primary transmission events	Assumed acquired resistance status	Primary/acquired DR-TB classification	Chronology of mixed DR types	Phenotypic DST	Genotypic DST
39	C16	P2	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
40	C16	P2	Yes	Primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
41	C16	P2	Yes	Primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
42	C16	M13	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
43	C16	M13	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
44	C17	M14	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
45	C17	M14	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
46	C18	M15	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
47	C18	M15	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
48	C19	M16	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
49	C19	M16	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
50	C19	M16	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
51	C19	M16	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
52	C20	X2	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
53	C20	X2	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
54	C21	pre-XDR	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	No	pre-XDR	pre-XDR
55	C21	MDR	Yes	Primary MDR	-	Primary MDR	No	MDR	MDR
56	C22	M17	Yes	Primary IR	Acquired MDR	Acquired MDR	Yes	MDR	MDR
57	C22	M17	Yes	Primary IR	-	Primary IR	Yes	MDR	PolyDR
58	C23	X3	Yes	Primary MDR>primary pre-XDR	Acquired XDR	Acquired XDR	No	XDR	XDR
59	C23	X3	Yes	Primary MDR>primary pre-XDR	-	Primary pre-XDR	No	XDR	pre-XDR
60	C24	M18	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
61	C24	M18	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
62	C25	M19	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
63	C25	M19	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
64	C26	M20	No	Primary pre-XDR	-	Primary pre-XDR	NA	MDR	pre-XDR
65	C26	M20	No	Primary pre-XDR	-	Primary pre-XDR	NA	MDR	pre-XDR
66	C27	M21	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
67	C27	M21	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
68	C28	M22	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
69	C28	M22	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
70	C28	M22	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
71	C29	M23	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
72	C29	M23	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
73	C29	M23	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
74	C29	M23	No	Primary MDR	-	Primary MDR	NA	MDR	MDR

Supplementary Table 8 Distribution of primary and acquired DR-TB among 89 clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Different DR types in cluster	Assumed previous primary transmission events	Assumed acquired resistance status	Primary/acquired DR-TB classification	Chronology of mixed DR types	Phenotypic DST	Genotypic DST
75	C30	M24	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
76	C30	M24	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
77	C30	M24	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
78	C31	M25	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
79	C31	M25	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
80	C31	M25	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
81	C32	X4	Yes	Primary MDR>primary pre-XDR	-	Primary pre-XDR	No	XDR	pre-XDR
82	C32	X4	Yes	Primary MDR>primary pre-XDR	Acquired XDR	Acquired XDR	No	XDR	XDR
83	C33	M26	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
84	C33	M26	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
85	C34	M27	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
86	C34	M27	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
87	C35	X5	Yes	Primary MDR	Acquired pre-XDR, acquired XDR	Acquired XDR	No	XDR	XDR
88	C35	X5	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	No	XDR	pre-XDR
89	C36	M28	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
90	C36	M28	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
91	C36	M28	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
92	C36	M28	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
93	C36	pre-XDR	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
94	C37	P3	Yes	Primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
95	C37	P3	Yes	Primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
96	C37	MDR	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
97	C38	M29	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
98	C38	M29	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
99	C39	M30	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
100	C39	M30	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
101	C39	M30	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
102	C40	MDR	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
103	C40	pre-XDR	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
104	C41	M31	Yes	Primary IR	-	Primary IR	No	MDR	PolyDR
105	C41	M31	Yes	Primary IR	Acquired MDR	Acquired MDR	No	MDR	MDR
106	C42	M32	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
107	C42	M32	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
108	C43	XDR	Yes	Primary pre-XDR	Acquired XDR	Acquired XDR	Yes	XDR	XDR
109	C43	pre-XDR	Yes	Primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
110	C44	M33	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
111	C44	M33	No	Primary MDR	-	Primary MDR	NA	MDR	MDR

Supplementary Table 8 Distribution of primary and acquired DR-TB among 89 clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Different DR types in cluster	Assumed previous primary transmission events	Assumed acquired resistance status	Primary/acquired DR-TB classification	Chronology of mixed DR types	Phenotypic DST	Genotypic DST
112	C45	M34	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
113	C45	M34	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
114	C46	X6	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
115	C46	X6	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
116	C47	P4	No	Primary pre-XDR	-	Primary pre-XDR	NA	pre-XDR	pre-XDR
117	C47	P4	No	Primary pre-XDR	-	Primary pre-XDR	NA	pre-XDR	pre-XDR
118	C48	M35	No	Primary IR	-	Primary IR	NA	MDR	PolyDR
119	C48	M35	No	Primary IR	-	Primary IR	NA	MDR	PolyDR
120	C48	M35	No	Primary IR	-	Primary IR	NA	MDR	PolyDR
121	C48	M35	No	Primary IR	-	Primary IR	NA	MDR	PolyDR
122	C48	M35	No	Primary IR	-	Primary IR	NA	MDR	PolyDR
123	C49	M36	Yes	Primary IR,>primary MDR	-	Primary MDR	Yes	MDR	MDR
124	C49	M36	Yes	Primary IR	Acquired MDR	Acquired MDR	Yes	MDR	MDR
125	C49	M36	Yes	Primary IR>primary MDR	-	Primary MDR	Yes	MDR	MDR
126	C50	M37	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
127	C50	M37	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
128	C50	M37	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
129	C51	M38	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
130	C51	M38	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
131	C52	M39	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
132	C52	M39	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
133	C53	M40	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
134	C53	M40	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
135	C54	MDR	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
136	C54	pre-XDR	No	Primary MDR	-	Primary MDR	NA	pre-XDR	MDR
137	C55	M41	Yes	Primary IR	-	Primary IR	No	MDR	PolyDR
138	C55	M41	Yes	Primary IR	Acquired MDR	Acquired MDR	No	MDR	MDR
139	C56	M42	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
140	C56	M42	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
141	C57	M43	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
142	C57	M43	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
143	C57	M43	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
144	C57	M43	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
145	C58	pre-XDR	No	Primary MDR	-	Primary MDR	NA	pre-XDR	MDR
146	C58	MDR	No	Primary MDR	-	Primary MDR	NA	MDR	MDR

Supplementary Table 8 Distribution of primary and acquired DR-TB among 89 clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Different DR types in cluster	Assumed previous primary transmission events	Assumed acquired resistance status	Primary/acquired DR-TB classification	Chronology of mixed DR types	Phenotypic DST	Genotypic DST
147	C59	M44	Yes	Primary IR	Acquired MDR	Acquired MDR	Yes	MDR	MDR
148	C59	M44	Yes	Primary IR>primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	MDR	pre-XDR
149	C59	P5	Yes	Primary IR>primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
150	C59	P5	Yes	Primary IR>primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
151	C59	XDR	Yes	Primary IR>primary MDR>primary pre-XDR	-	Primary pre-XDR	Yes	XDR	pre-XDR
152	C60	M45	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
153	C60	M45	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
154	C60	pre-XDR	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
155	C61	M46	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
156	C61	M46	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
157	C61	M46	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
158	C61	M46	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
159	C62	M47	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
160	C62	M47	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
161	C62	M47	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
162	C62	M47	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
163	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
164	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
165	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
166	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
167	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
168	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
169	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
170	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
171	C63	M48	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
172	C63	pre-XDR	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
173	C64	M49	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
174	C64	M49	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
175	C65	M50	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
176	C65	M50	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
177	C66	M51	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
178	C66	M51	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
179	C67	M52	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
180	C67	M52	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR

Supplementary Table 8 Distribution of primary and acquired DR-TB among 89 clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Different DR types in cluster	Assumed previous primary transmission events	Assumed acquired resistance status	Primary/acquired DR-TB classification	Chronology of mixed DR types	Phenotypic DST	Genotypic DST
181	C68	M53	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
182	C68	M53	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
183	C69	M54	No	Primary IR	-	Primary IR	NA	MDR	PolyDR
184	C69	M54	No	Primary IR	-	Primary IR	NA	MDR	PolyDR
185	C70	X7	Yes	Primary IR>primary MDR>primary XDR	-	Primary XDR	Yes	XDR	XDR
186	C70	X7	Yes	Primary IR>primary MDR>primary XDR	-	Primary XDR	Yes	XDR	XDR
187	C70	MDR	Yes	Primary IR	Acquired MDR	Acquired MDR	Yes	MDR	MDR
188	C71	M55	Yes	Primary IR>primary MDR	-	Primary MDR	No	MDR	MDR
189	C71	M55	Yes	Primary IR>primary MDR	-	Primary MDR	No	MDR	MDR
190	C71	M55	Yes	Primary IR>primary MDR	-	Primary MDR	No	MDR	MDR
191	C71	M55	Yes	Primary IR>primary MDR	-	Primary MDR	No	MDR	MDR
192	C71	M55	Yes	Primary IR>primary MDR	-	Primary MDR	No	MDR	MDR
193	C71	M55	Yes	Primary IR	Acquired MDR	Acquired MDR	No	MDR	MDR
194	C71	P6	Yes	Primary IR	-	Primary IR	No	pre-XDR	PolyDR
195	C71	P6	Yes	Primary IR>primary MDR	-	Primary MDR	No	pre-XDR	MDR
196	C72	P7	Yes	Primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
197	C72	P7	Yes	Primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
198	C72	P7	Yes	Primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
199	C72	P7	Yes	Primary pre-XDR	-	Primary pre-XDR	Yes	pre-XDR	pre-XDR
200	C72	X8	Yes	Primary pre-XDR	-	Primary pre-XDR	Yes	XDR	pre-XDR
201	C72	X8	Yes	Primary pre-XDR	-	Primary pre-XDR	Yes	XDR	pre-XDR
202	C72	X8	Yes	Primary pre-XDR	Acquired XDR	Acquired XDR	Yes	XDR	XDR
203	C73	X8	No	Primary pre-XDR	-	Primary pre-XDR	NA	XDR	pre-XDR
204	C73	pre-XDR	No	Primary pre-XDR	-	Primary pre-XDR	NA	pre-XDR	pre-XDR
205	C74	M56	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
206	C74	M56	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
207	C75	MDR	Yes	Primary IR	Acquired MDR	Acquired MDR	No	MDR	MDR
208	C75	pre-XDR	Yes	Primary IR	Acquired pre-XDR	Acquired pre-XDR	No	pre-XDR	pre-XDR
209	C76	MDR	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
210	C76	pre-XDR	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
211	C77	M57	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
212	C77	M57	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
213	C77	M57	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
214	C78	M58	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
215	C78	M58	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
216	C78	M58	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
217	C78	M58	No	Primary MDR	-	Primary MDR	NA	MDR	MDR

Supplementary Table 8 Distribution of primary and acquired DR-TB among 89 clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Different DR types in cluster	Assumed previous primary transmission events	Assumed acquired resistance status	Primary/acquired DR-TB classification	Chronology of mixed DR types	Phenotypic DST	Genotypic DST
218	C79	M59	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
219	C79	M59	No	Primary IR	Acquired MDR	Acquired MDR	NA	MDR	MDR
220	C80	M60	Yes	Primary IR>primary MDR	-	Primary MDR	Yes	MDR	MDR
221	C80	M60	Yes	Primary IR>primary MDR	-	Primary MDR	Yes	MDR	MDR
222	C80	M60	Yes	Primary IR>primary MDR	-	Primary MDR	Yes	MDR	MDR
223	C80	M60	Yes	Primary IR	Acquired MDR	Acquired MDR	Yes	MDR	MDR
224	C81	M61	No	Primary pre-XDR	-	Primary pre-XDR	NA	MDR	pre-XDR
225	C81	M61	No	Primary pre-XDR	-	Primary pre-XDR	NA	MDR	pre-XDR
226	C82	P8	No	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	NA	pre-XDR	pre-XDR
227	C82	P8	No	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	NA	pre-XDR	pre-XDR
228	C83	M62	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
229	C83	M62	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
230	C83	M62	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	MDR	pre-XDR
231	C84	X9	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
232	C84	X9	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
233	C85	M63	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
234	C85	M63	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
235	C85	M63	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
236	C85	M63	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
237	C86	X10	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
238	C86	X10	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
239	C86	X10	No	Primary XDR	-	Primary XDR	NA	XDR	XDR
240	C87	M64	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
241	C87	M64	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
242	C87	M64	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
243	C88	M65	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
244	C88	M65	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
245	C88	M65	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
246	C88	M65	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
247	C88	M65	No	Primary MDR	-	Primary MDR	NA	MDR	MDR
248	C89	P9	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
249	C89	P9	Yes	Primary MDR	-	Primary MDR	Yes	pre-XDR	MDR
250	C89	P9	Yes	Primary MDR	-	Primary MDR	Yes	pre-XDR	MDR
251	C89	P9	Yes	Primary MDR	-	Primary MDR	Yes	pre-XDR	MDR
252	C89	P9	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR
253	C89	P9	Yes	Primary MDR	-	Primary MDR	Yes	pre-XDR	MDR
254	C89	P9	Yes	Primary MDR	Acquired pre-XDR	Acquired pre-XDR	Yes	pre-XDR	pre-XDR

Supplementary Table 8 Distribution of primary and acquired DR-TB among 89 clusters (Cont.)

No.	Cluster[1]	Cluster[2]	Different DR types in cluster	Assumed previous primary transmission events	Assumed acquired resistance status	Primary/acquired DR-TB classification	Chronology of mixed DR types	Phenotypic DST	Genotypic DST
255	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
256	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
257	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
258	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
259	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
260	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
261	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
262	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
263	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
264	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
265	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
266	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
267	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
268	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
269	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
270	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
271	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
272	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
273	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
274	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
275	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
276	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
277	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
278	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
279	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
280	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR
281	C89	M66	Yes	Primary MDR	-	Primary MDR	Yes	MDR	MDR

Supplementary Table 9 Characteristics of 85 MDR-TB (M1-M66), pre-XDR-TB (P1-P09) and XDR-TB (X1-X10) clusters

No.	Cluster*	Isolates (no.)	Member of major clade	Geographical link			Time link, year (no.)
				Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
1	M1	2	No	Northeastern (2)	Buri Ram (2)	Krasang Hospital (2)	2014 (1), 2016 (1)
2	M2	2	No	Central (1), Eastern (1)	Bangkok (1), Rayong (1)	Sirinthorn Hospital (1), Rayong Hospital (1)	2016 (1), 2017 (1)
3	M3	2	No	Central (1), Eastern (1)	Nonthaburi (1), Chon Buri (1)	Bamrasnaradura Institute (1), Chonburi Hospital (1)	2016 (1), 2017 (1)
4	M4	3	No	Northeastern (1), Central (2)	Loei (1), Nonthaburi (1), Saraburi (1)	Naduang Hospital (1), Bamrasnaradura Institute (1), Saraburi Hospital (1)	2014 (2), 2016 (1)
5	M5	3	No	Southern (3)	Krabi (1), Nakhon Si Thammarat (2)	Nueklong Hospital (1), Maharajnakhonsithammarat Hospital (1), Office of Disease Prevention & Control 11 Nakhon Si Thammarat (1)	2015 (1), 2017 (2)
6	M6	2	No	Northeastern (2)	Khon Kaen (2)	Khonkaen Hospital (2)	2014 (1), 2016 (1)
7	M7	2	No	Central (2)	Bangkok (1), Nonthaburi (1)	Devision of Tuberculosis (1), Bamrasnaradura Institute (1)	2015 (1), 2016 (1)
8	M8	3	**Yes, Clade1**	Eastern (3)	Trat (3)	Trat Hospital (3)	2014 (2), 2015 (1)
9	M9	2	No	Central (2)	Lop Buri (2)	Khoksamrong Hospital (2)	2014 (1), 2015 (1)
10	M10	3	**Yes, Clade2**	Central (2), Western (1)	Bangkok (2), Prachuap Khiri Khan (1)	Klang Hospital (1), Rajavithi Hospital (1), Bangsabhan Hospital (1)	2016 (1), 2017 (2)
11	M11	2	**Yes, Clade2**	Northeastern (1), Central (1)	Maha Sarakham (1), Samut Prakan (1)	Phayakkhaphumphisai Hospital (1), Bangbo Hospital (1)	2017 (2)
12	M12	2	**Yes, Clade2**	Western (2)	Prachuap Khiri Khan (2)	Samroiyod Hospital (2)	2014 (1), 2016 (1)
13	M13	3	**Yes, Clade2**	Northeastern (2), Central (1)	Ubon Ratchathani (1), Udon Thani (1), Bangkok (1)	Trakanphuetpol Hospital (1), Udonthani Hospital (1), Rajavithi Hospital (1)	2014 (2), 2016 (1)
14	M14	2	**Yes, Clade2**	Northeastern (1), Central (1)	Khon Kaen (1), Bangkok (1)	Khonkaen Hospital (1), Rajavithi Hospital (1)	2016 (1), 2017 (1)
15	M15	2	**Yes, Clade2**	Southern (2)	Chumphon (2)	Chumphonkhetudomsakdi Hospital (2)	2015 (1), 2016 (1)
16	M16	4	**Yes, Clade2**	Northeastern (2), Eastern (1), Southern (1)	Ubon Ratchathani (1), Amnat Charoen (1), Chon Buri (1), Songkhla (1)	Somdetphrayuphrarat Detudom Hospital (1), Amnatcharoen Hospital (1), Chonburi Hospital (1), Songkhla Hospital (1)	2016 (1), 2017 (3)
17	M17	2	No	Central (1), Western (1)	Samut Songkhram (1), Phetchaburi (1)	Somdejphraphutthaloetla Hospital (1), Phrachomklao Hospital (1)	2016 (1), 2017 (1)
18	M18	2	**Yes, Clade3**	Eastern (2)	Rayong (1), Chon Buri (1)	Rayong Hospital (1), Chonburi Hospital (1)	2014 (1), 2016 (1)
19	M19	2	**Yes, Clade3**	Central (1), Southern (1)	Samut Prakan, Surat Thani	Bangbo Hospital (1), Suratthani Hospital (1)	2016 (1), 2017 (1)
20	M20	2	No	Northeastern (2), Southern (2)	Udon Thani (1), Surat Thani (1)	Udonthani Hospital (1), Kohsamui Hospital (1)	2014 (1), 2016 (1)
21	M21	2	No	Central (1), Eastern (1)	Pathum Thani (1), Trat (1)	Ladlumkaew Hospital (1), Trat Hospital (1)	2014 (1), 2015 (1)

Supplementary Table 9 Characteristics of 85 MDR-TB (M1-M66), pre-XDR-TB (P1-P09) and XDR-TB (X1-X10) clusters (Cont.)

No.	Cluster*	Isolates (no.)	Member of major clade	Geographical link			Time link, year (no.)
				Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
22	M22	3	**Yes, Clade4**	Northeastern (3)	Khon Kaen (1), Maha Sarakham (2)	Banphai Hospital (1), Borabue Hospital (1), Mahasarakham Hospital 91)	2016 (2), 2017 (1)
23	M23	4	No	Central (2), Eastern (2)	Chai Nat (1), Suphan Buri (1), Chachoengsao (1), Prachin Buri (1)	Hankha Hospital (1), Chaophrayayommarat Hospital (1), Buddhasothorn Hospital (1), Prachantakham Hospital (1)	2014 (1), 2015 (1), 2017(2)
24	M24	3	No	Northeastern (3)	Buri Ram (2)	Banruat Hospital (1), Buriram Hospital (2)	2015 (3)
25	M25	3	**Yes, Clade5**	Northeastern (1), Central (1), Eastern (1)	Buri Ram (1), Saraburi (1), Prachin Buri (1)	Buriram Hospital (1), Saraburi Hospital (1), Chaopraya Abhaiphubet Hospital (1)	2014 (1), 2017 (2)
26	M26	2	No	Siuthern (2)	Pattani (1), Yala (1)	Pattani Hospital (1), Yala Hospital (1)	2016 (2)
27	M27	2	No	Northeastern (1), Central (1)	Khon Kaen (1), Bangkok (1)	Somdetphrayuphrarat Kranuan Hospital (1), Public Health Center 27 (1)	2015 (1), 2016 (1)
28	M28	4	**Yes, Clade6**	Northeastern (1), Southern (3)	Udon Thani (1), Phuket (3)	Udonthani Hospital (1), Patong Hospital (1), Vachiraphuket Hospital (2)	2014 (1), 2015 (2), 2016 (1)
29	M29	2	**Yes, Clade7**	Southern (2)	Krabi (1), Satun (1)	Khlongthom Hospital (1), Satun Hospital (1)	2016 (2)
30	M30	3	**Yes, Clade7**	Southern (3)	Nakhon Si Thammarat (1), Phuket (1), Phang Nga (1)	Office of Disease Prevention & Control 11 Nakhon Si Thammarat (1), Vachiraphuket Hospital (1), Khuraburichaipat Hospital (1)	2014 (1), 2016 (1), 2017 (1)
31	M31	2	**Yes, Clade7**	Southern (2)	Surat Thani (1), Phuket (1)	Suratthani Hospital (1), Vachiraphuket Hospital (1)	2014 (1), 2015 (1)
32	M32	2	**Yes, Clade7**	Southern (2)	Phatthalung (2)	Kongrha Hospital (1), Phatthalung Hospital (1)	2015 (1), 2017 (1)
33	M33	2	No	Northeastern (1), Central (1)	Bungkan (1), Phetchabun (1)	Sriwilai Hospital (1), Nongphai Hospital (1)	2015 (2)
34	M34	2	No	Central (2)	Bangkok (1), Phichit (1)	Sirinthorn Hospital (1), Wangsaiphun Hospital (1)	2014 (1), 2015 (1)
35	M35	5	**Yes, Clade8**	Central (1), Eastern (4)	Nonthaburi (1), Rayong (1), Chachoengsao (1), Chanthaburi (1), Sa Kaeo (1)	Bamrasnaradura Institute (1), Nikompattana Hospital (1), Buddhasothorn Hospital (1), Prapokklao Hospital (1), Khlonghat Hospital (1)	2015 (3), 2017 (2)
36	M36	3	**Yes, Clade8**	Eastern (3)	Rayong (2), Sa Kaeo (1)	Rayong Hospital (2), Wangnamyen Hospital (1)	2014 (2), 2015 (1)
37	M37	3	**Yes, Clade8**	Eastern (3)	Chon Buri (1), Chanthaburi (2)	Chonburi Hospital (1), Khlung Hospital (1), Prapokklao Hospital (1)	2014 (1), 2015 (1), 2016 (1)
38	M38	2	**Yes, Clade8**	Central (1), Eastern (1)	Uthai Thani (1), Chon Buri (1)	Nongchang Hospital (1), Banglamung Hospital (1)	2014 (1), 2016 (1)
39	M39	2	**Yes, Clade9**	Eastern (2)	Chon Buri (1), Sa Kaeo (1)	Chonburi Hospital (1), Sakaeo Hospital (1)	2014 (1), 2017 (1)
40	M40	2	**Yes, Clade9**	Central (2)	Bangkok (1), Sing Buri (1)	Devision of Tuberculosis (1), Singburi Hospital (1)	2014 (1), 2016 (1)
41	M41	2	No	Central (2)	Samut Sakhon (2)	Samutsakhon Hospital (2)	2014 (1), 2017 (1)

Supplementary Table 9 Characteristics of 85 MDR-TB (M1-M66), pre-XDR-TB (P1-P09) and XDR-TB (X1-X10) clusters (Cont.)

No.	Cluster[2]*	Isolates (no.)	Member of major clade	Geographical link			Time link, year (no.)
				Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
42	M42	2	**Yes, Clade10**	Southern (2)	Songkhla (1), Satun (1)	Songkhla Hospital (1), Satun Hospital (1)	2015 (1), 2017 (1)
43	M43	4	**Yes, Clade10**	Northeastern (1), Central (2), Western (1)	Nakhon Ratchasima (1), Nonthaburi (1), Samut Prakan (1), Phetchaburi (1)	Nonthai Hospital (1), Pranangklao Hospital (1), Samutprakan Hospital (1), Cha-am Hospital (1)	2014(2), 2016 (1), 2017 (1)
44	M44	2	**Yes, Clade11**	Northeastern (1), Central (1)	Loei (1), Kamphaeng Phet (1)	Wangsaphung Hospital (1), Kamphaengphet Hospital (1)	2014 (1), 2017 (1)
45	M45	2	**Yes, Clade11**	Northeastern (2)	Nakhon Ratchasima (1), Chaiyaphum (1)	Office of Disease Prevention & Control 9 Nakhon Ratchasima (1), Chaiyaphum Hospital (1)	2015 (1), 2017 (1)
46	M46	4	**Yes, Clade11**	Northeastern (3), Central (1)	Buri Ram (2), Nakhon Ratchasima (1), Sukhothai (1)	Buriram Hospital (1), Nangrong Hospital (1), The Golden Gate Hospital (1), Sisatchanalai Hospital (1)	2014 (2), 2015 (2)
47	M47	4	**Yes, Clade11**	Northeastern (1), Central (2), Western (1)	Si Sa Ket (1), Bangkok (1), Samut Prakan (1), Kanchanaburi (1)	Kantharalak Hospital (1), Public Health Center 4 (1), Bangbo Hospital (1), Paholpolpayuhasena Hospital (1)	2015 (2), 2017 (2)
48	M48	9	**Yes, Clade11**	Northeastern (1), Central (2), Weastern (6)	Nong Khai (1), Bangkok (1), Samut Prakan (1), Phetchaburi (6)	Nongkhai Hospital (1), Charoenkrung Pracharak Hospital (1), Bangbo Hospital (1), Phrachomklao Hospital (6)	2014 (1), 2015 (2), 2016 (4), 2017 (2)
49	M49	2	**Yes, Clade12**	Northeastern (1), Central (1)	Udon Thani (1), Bangkok (1)	Udonthani Hospital (1), Public Health Center 28 (1)	2014 (1), 2017 (1)
50	M50	2	**Yes, Clade12**	Northeastern (2)	Si Sa Ket (1), Roi Et (1)	Sisaket Hospital (1), Roi-et Hospital (1)	2015 (1), 2017 (1)
51	M51	2	**Yes, Clade12**	Eastern (2)	Chon Buri (1), Chanthaburi (1)	Chonburi Hospital (1), Prapokklao Hospital (1)	2014 (1), 2016 (1)
52	M52	2	**Yes, Clade12**	Eastern (1), Northern (1)	Chon Buri (1), Chiang Mai (1)	Chonburi Hospital (1), Office of Disease Prevention & Control 1 Chiangmai (1)	2017 (2)
53	M53	2	**Yes, Clade12**	Central (1), Eastern (1)	Phitsanulok (1), Chon Buri (1)	Buddhachinaraj Hospital (1), Banglamung Hospital (1)	2014 (1), 2017 (1)
54	M54	2	**Yes, Clade12**	Central (2)	Bangkok (2)	Taksin Hospital (1), Public Health Center 36 (1)	2016 (2)
55	M55	6	**Yes, Clade12**	Northeastern (2), Central (1), Eastern (2), Southern (1)	Ubon Ratchathani (1), Udon Thani (1), Bangkok (1), Chon Buri (2), Songkhla (1)	Somdetphrayuphrarat Detudom Hospital (1), Udonthani Hospital (1), Nopparat Rajathanee Hospital (1), Chonburi Hospital (2), Hatyai Hospital (1)	2014 (1), 2015 (1), 2016 (1), 2017 (3)
56	M56	2	**Yes, Clade12**	Central (2)	Bangkok (1), Sing Buri (1)	Taksin Hospital (1), Singburi Hospital (1)	2015 (1), 2017 (1)
57	M57	3	**Yes, Clade12**	Central (3)	Bangkok (3)	Devision of Tuberculosis (1), Public Health Center 30 (1), Public Health Center 40 (1)	2014 (1), 2015 (2)
58	M58	4	**Yes, Clade12**	Northeastern (1), Eastern (3)	Chaiyaphum (1), Chon Buri (2), Chachoengsao (1)	Kaengkhro Hospital (1), Chonburi Hospital (2), Buddhasothorn Hospital (1)	2017 (4)

Supplementary Table 9 Characteristics of 85 MDR-TB (M1-M66), pre-XDR-TB (P1-P09) and XDR-TB (X1-X10) clusters (Cont.)

No.	Cluster[2]*	Isolates (no.)	Member of major clade	Geographical link			Time link, year (no.)
				Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
59	M59	2	**Yes, Clade12**	Eastern (1), Northeastern (1)	Chon Buri (1), Chiang Mai (1)	Phanatnikhom Hospital (1), Office of Disease Prevention & Control 1 Chiangmai (1)	2015 (1), 2017 (1)
60	M60	4	**Yes, Clade12**	Central (4)	Bangkok (4)	Devision of Tuberculosis (1), Public Health Center 23 (1), Public Health Center 29 (1), Public Health Center 48 (1)	2015 (1), 2017 (3)
61	M61	2	**Yes, Clade13**	Central (1), Western (1)	Samut Sakhon (1), Kanchanaburi (1)	Samutsakhon Hospital (1), Makarak Hospital (1)	2017 (2)
62	M62	3	**Yes, Clade13**	Central (2), Eastern (1)	Suphan Buri (2), Chanthaburi (1)	Chaophrayayommarat Hospital (1), Uthong Hospital (1), Prapokklao Hospital (1)	2015 (1), 2016 (1), 2017 (1)
63	M63	4	**Yes, Clade13**	Western (4)	Kanchanaburi (4)	Makarak Hospital (3), Paholpolpayuhasena Hospital (1)	2014 (1), 2015 (1), 2017 (2)
64	M64	3	**Yes, Clade13**	Western (3)	Kanchanaburi (3)	Makarak Hospital (2), Paholpolpayuhasena Hospital (1)	2014 (2), 2017 (1)
65	M65	5	**Yes, Clade13**	Central (3), Western (2)	Samut Sakhon (2), Nakhon Pathom (1), Kanchanaburi (2)	Banphaeo Hospital (2), Nakhonpathom Hospital (1), Makarak Hospital (1), Paholpolpayuhasena Hospital (1)	2014 (2), 2015 (1), 2016 (1), 2017 (1)
66	M66	27	**Yes, Clade13**	Central (6), Western (21)	Bangkok (2), Suphan Buri (2), Samut Sakhon (1), Phitsanulok (1), Kanchanaburi (20), Ratchaburi (1)	Rajavithi Hospital (1), Klang Hospital (1), Danchang Hospital (1), Somdetphrasangkharat 17 Hospital (1), Banphaeo Hospital (1), Buddhachinaraj Hospital (1), Makarak Hospital (13), Paholpolpayuhasena Hospital (5), Danmakhamtia Hospital (1), Saiyok Hospital (1), Banpong Hospital (1)	2014 (7), 2015 (8), 2016 (4), 2017 (14)
67	P1	2	No	Central (2)	Bangkok (2)	Rajavithi Hospital (2)	2016 (2)
68	P2	3	**Yes, Clade2**	Northeastern (2), Western (1)	Ubon Ratchathani (2), Kanchanaburi (1)	Trakanphuetpol Hospital (1), Fort sunpasitthiprasong Hospital (1), Makarak Hospital(1)	2014 (1), 2016 (1), 2017 (1)
69	P3	2	No	Southern (2)	Satun (2)	Satun Hospital (2)	2015 (2)
70	P4	2	No	Central (2)	Bangkok (1), Kamphaeng Phet (1)	Police Hospital (1), Kamphaengphet Hospital (1)	2014 (2)
71	P5	2	**Yes, Clade11**	Eastern (2)	Chon Buri (2)	Chon Buri Hospital (2)	2014 (1), 2016 (1)
72	P6	2	**Yes, Clade12**	Central (1), Eastern (1)	Bangkok(1), Chon Buri (1)	Taksin Hospital (1), Chonburi Hospital (1)	2014 (1), 2016 (1)
73	P7	4	**Yes, Clade12**	Northeastern (2), Central (2)	Ubon Ratchathani (2), Bangkok (2)	Somdetphrayuphrarat Detudom Hospital (1), Warinchamrap Hospital (1), Klang Hospital (1), Navamin Hospital 9 (1)	2016 (3), 2017 (1)
74	P8	2	**Yes, Clade13**	Central (1), Western (1)	Suphan Buri (1), Kanchanaburi (1)	Uthong Hospital (1), Makarak Hospital (1)	2016 (1), 2017 (1)
75	P9	7	**Yes, Clade13**	Central (1), Southern (1), Western (5)	Bangkok (1), Surat Thani (1), Kanchanaburi (5)	Klang Hospital (1), Suratthani Hospital (1), Paholpolpayuhasena Hospital (2), Makarak Hospital (2), Somdetphrasangkharat 19 Hospital (1)	2014 (1), 2015 (1), 2016 (1), 2017 (4)

Supplementary Table 9 Characteristics of 85 MDR-TB (M1-M66), pre-XDR-TB (P1-P09) and XDR-TB (X1-X10) clusters (Cont.)

No.	Cluster[2]*	Isolates (no.)	Member of major clade	Geographical link			Time link, year (no.)
				Region-based link (no.)	Province-based link (no.)	Hospital-based link (no.)	
76	X1	2	No	Western (2)	Kanchanaburi (1), Ratchaburi (1)	Makarak Hospital (1), Ratchaburi Hospital (1)	2015 (1), 2017 (1)
77	X2	2	No	Northeastern (2)	Amnat Charoen (2)	Amnatcharoen Hospital (2)	2015 (1), 2016 (1)
78	X3	2	**Yes, Clade3**	Western (2)	Ratchaburi (2)	Ratchaburi Hospital (2)	2015 (1), 2017 (1)
79	X4	2	No	Western (2)	Tak	Maesot Hospital (2)	2014 (2)
80	X5	2	**Yes, Clade6**	Southern (1), Western (1)	Phuket (1), Prachuap Khiri Khan 91)	Vachiraphuket Hospital (1), Hua-Hin Hospital (1)	2014 (1), 2017 (1)
81	X6	2	No	Central (2)	Bangkok (2)	Devision of Tuberculosis (2)	2015 (1), 2016 (1)
82	X7	2	**Yes, Clade12**	Southern (2)	Ranong (2)	Ranong Hospital (2)	2014 (1), 2016 (1)
83	X8	4	**Yes, Clade12**	Northeastern (2), Central (1), Eastern (1)	Ubon Ratchathani (1), Mukdahan (1), Bangkok (1), Trat (1)	Somdetphrayuphrarat Detudom Hospital (1), Mukdahan Hospital (1), Devision of Tuberculosis (1), Trat Hospital (1)	2014 (3),2015 (1)
84	X9	2	**Yes, Clade13**	Eastern (2)	Chonburi (2)	Chonburi Hospital (1)	2014 (1), 2016 (1)
85	X10	3	**Yes, Clade13**	Western (3)	Kanchanaburi (3)	Paholpolpayuhasena Hospital (1), Makarak Hospital (2)	2014 (1), 2015 (1), 2017 (1)

*DR-TB types (MDR), pre-XDR and XDR) were based on phenotypic DST. Clusters defined by SNP pairwise differences ≤11 and by type of phenotypic drug resistance.

Supplementary Table 10 Distribution (by year, phenotypic DR type, region and province) of 281 clustering isolates

Region	Abbreviation[a]	2014			2015			2016			2017			Total, no. (%)			
		MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	Total
Central	BKK	5	1	1	5		1	7	4	1	9	3		26 (12.09)	8 (20.00)	3 (11.54)	37 (13.17)
	CNT										1			1 (0.47)			1 (0.36)
	KPT	1	1											1 (0.47)	1 (2.50)		2 (0.71)
	LRI	1			1									2 (0.93)			2 (0.71)
	NPT							1						1 (0.47)			1 (0.36)
	NBI	1	1		1			2			1			5 (2.33)	1 (2.50)		6 (2.14)
	PTE				1									1 (0.47)			1 (0.36)
	PNB				1									1 (0.47)			1 (0.36)
	PCT				1									1 (0.47)			1 (0.36)
	PLK				1						1			2 (0.93)			2 (0.71)
	SPK	1			2						2			5 (2.33)			5 (1.78)
	SKN	1			2						3			6 (2.79)			6 (2.14)
	SKM										1			1 (0.47)			1 (0.36)
	SRI	1									1			2 (0.93)			2 (0.71)
	SBR				1			1						2 (0.93)			2 (0.71)
	STI				1									1 (0.47)			1 (0.36)
	SPB	1			2				1		2	1		5 (2.33)	2 (5.00)		7 (2.49)
	UTI							1						1 (0.47)			1 (0.36)
Eastern	CCO	1			1						1			3 (1.40)			3 (1.07)
	CTI	1			2			2						5 (2.33)			5 (1.78)
	CBI	5	3	1				2	1	1	8			15 (6.98)	4 (10.00)	2 (7.69)	21 (7.47)
	PRI	1									1			2 (0.93)			2 (0.71)
	RYG	2						1	1		2			5 (2.33)	1 (2.50)		6 (2.14)
	SKW	1			1						1			3 (1.40)			3 (1.07)
	TRT	3			1		1		1				1	4 (1.86)	1 (2.50)	2 (7.69)	7 (2.49)
Northeastern	ACR						1			1	1			1 (0.47)		2 (7.69)	3 (1.07)
	BKN				1									1 (0.47)			1 (0.36)
	BRM	3			3			1			1			8 (3.72)			8 (2.85)
	CPM				1				2		1		1	2 (0.93)	2 (5.00)	1 (3.85)	5 (1.78)
	KSN							1						1 (0.47)			1 (0.36)
	KKN	1			1			3	1	1	1			6 (2.79)	1 (2.50)	1 (3.85)	8 (2.85)
	LEI							1			1			2 (0.93)			2 (0.71)
	MKM							3			1			4 (1.86)			4 (1.42)
	MDH			1												1 (3.85)	1 (0.36)
	NMA				1						2	1		3 (1.40)	1 (2.50)		4 (1.42)
	NKI							2				1		2 (0.93)	1 (2.50)		3 (1.07)
	RET				1									1 (0.47)			1 (0.36)
	SSK										2			2 (0.93)			2 (0.71)
	UBN	1	1	1					4		1			2 (0.93)	5 (12.50)	1 (3.85)	8 (2.85)
	UDN	2			1			2						5 (2.33)			5 (1.78)

Supplementary Table 10 Distribution (by year, phenotypic DR type, region and province) of 281 clustering isolates (Cont.)

Region	Abbreviation[a]	2014			2015			2016			2017			Total, no. (%)			
		MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	MDR	Pre-XDR	XDR	Total
Northern	CMI				1						1			2 (0.93)			2 (0.71)
	PRE	1				1								1 (0.47)	1 (2.50)		2 (0.71)
Southern	CPN				1			1						2 (0.93)			2 (0.71)
	KBI				1			1						2 (0.93)			2 (0.71)
	NST							1			2			3 (1.40)			3 (1.07)
	PTN							1						1 (0.47)			1 (0.36)
	PNA	1												1 (0.47)			1 (0.36)
	PLG				1						1			2 (0.93)			2 (0.71)
	PKT	2		1	2						1			5 (2.33)		1 (3.85)	6 (2.14)
	RNG			1						1						2 (7.69)	2 (0.71)
	STN	1				2		1			1			3 (1.40)	2 (5.00)		5 (1.78)
	SKA				1			1			1			3 (1.40)			3 (1.07)
	SNI	1			1	1		1						3 (1.40)	1 (2.50)		4 (1.42)
	TRG							1						1 (0.47)			1 (0.36)
	YLA							1						1 (0.47)			1 (0.36)
Western	KRI	10	1	1	6		2	3	1		13	5	1	32 (14.88)	7 (17.50)	4 (15.38)	43 (15.30)
	PBI	1			1			4			2	1		8 (3.72)	1 (2.50)		9 (3.20)
	PKN	1						1			1		1	3 (1.40)		1 (3.85)	4 (1.42)
	RBR				1		1						2	1 (0.47)		3 (11.54)	4 (1.42)
	TAK			2												2 (7.69)	2 (0.71)
Total		**51**	**8**	**9**	**49**	**4**	**6**	**47**	**16**	**5**	**68**	**11**	**7**	**215**	**40**	**26**	**281**

Note: DR-TB types (MDR, pre-XDR and XDR) were based on phenotypic DST.

[a] Full name of all provinces were listed in Supplementary Table 12

Supplementary Table 11 Frequency and MIC-distribution of isolates with drug resistance-conferring mutations.

Drug	Mutations [a]	MIC (µg/ml) distribution for relevant isolate									No. of isolates	%
Isoniazid		**≤0.03**	**0.06**	**0.12**	**0.25**	**0.5**	**1**	**2**	**4**	**>4**		
(CC = 0.25 µg/ml)	No INH mutations			1	1	1		1			4	7
	katG Ser315Thr						12	21	9	1	43	72
	katG Ser315Thr, *katG* Ala424Gly								1		1	2
	katG Ser315Asn						1				1	2
	inhA -15 c/t			1	1	2	1				5	8
	inhA -15 c/t, *inhA* Ser94Ala									1	1	2
	inhA -15 c/t, *katG* Met257Ile						1				1	2
	inhA -8 t/c, *katG* Ser315Thr								1	1	2	3
	inhA Ser94Ala					2					2	3
Rifampicin		**≤0.12**	**0.25**	**0.5**	**1**	**2**	**4**	**8**	**16**	**>16**		
(CC = 1 µg/ml)	No RIF mutations	1	1		1						3	5
	rpoB Ser450Leu							1		35	36	60
	rpoB His445Arg									2	2	3
	rpoB Ser450Leu, *rpoC* Leu527Val									1	1	2
	rpoB His445Leu					1					1	2
	rpoB His445Tyr									2	2	3
	rpoB Asp435Phe						1	1			2	3
	rpoB Asp435Tyr				1						1	2
	rpoB Asp435Val							1		1	2	3
	rpoB Leu452Pro		1	1		1					3	5
	rpoB Ser441Leu		1						1		2	3
	rpoB Ser450Trp		1								1	2
	rpoB Val170Phe					1					1	2
	rpoB 1295_1303del				1						1	2
	rpoB 1295_1303del, *rpoB* Ser450Leu						1				1	2
	rpoB Leu430Arg, *rpoB* Asp435Tyr									1	1	2
Ethambutol		**≤0.5**	**1**	**2**	**4**	**8**	**16**	**32**	**>32**			
(CC = 4 µg/ml)	No EMB mutations	1	7	7	3						18	30
	embB Met306Ile			1	6	6	3				16	27
	embB Gly406Asp			1	5						6	10
	embB Met306Val				1	2	1				4	7
	embB Tyr319Ser				1	3					4	7
	embB Gln497Arg				1	1					2	3

Supplementary Table 11 Frequency and MIC-distribution of isolates with drug resistance-conferring mutations (Cont.)

Drug	Mutations [a]	MIC (µg/ml) distribution for relevant isolate									No. of isolates	%
Ethambutol		**≤0.5**	**1**	**2**	**4**	**8**	**16**	**32**	**>32**			
	embB Asp328Tyr					1					1	2
	embB Asp354Ala			1							1	2
	embB Gly406Cys					1					1	2
	embB Met306Leu				1						1	2
	embB Asp1024Asn, *embB* Met306Ile					2					2	3
	embB Asp1024Asn, *embB* Gly406Ser				1						1	2
	embA -12 c/t			2							2	3
	embB Asp328Tyr, *embA* -16 c/t					1					1	2
Streptomycin		**≤0.25**	**0.5**	**1**	**2**	**4**	**8**	**16**	**32**	**>32**		
(CC = 2 µg/ml)	No STR mutations	10	5		1	1					17	28
	rpsL Lys43Arg									28	28	47
	rpsL Lys88Arg					2		2	**1**	1	6	10
	gid Gly73Ala			1	1	3	**1**				6	10
	gid 115_115del				1						1	2
	*gid*_Chromosome:g.4407954_4408172del		1								1	2
	rpsL Lys88Arg, *gid* Gly73Ala					1					1	2
Kanamycin		**≤0.6**	**1.2**	**2.5**	**5**	**10**	**20**	**40**	**>40**			
(CC = 5 µg/ml)	No KAN mutations	11	23	6					1		41	68
	rrs A1401G								17		17	28
	eis -14 c/t					1					1	2
	eis -8 c/a					1					1	2
Amikacin		**≤0.12**	**0.25**	**0.5**	**1**	**2**	**4**	**8**	**16**	**>16**		
(CC = 4 µg/ml)	No AMK mutations	4	26	**9**	3					1	43	72
	rrs A1401G							2	**1**	14	17	28
Ofloxacin		**≤0.25**	**0.5**	**1**	**2**	**4**	**8**	**16**	**32**	**>32**		
(CC = 2 µg/ml)	No FQ mutations	1	12	15	3		1				32	53
	gyrA Asp94Gly					1	7	6	1		15	25
	gyrA Ala90Val					3	2		1		6	10
	gyrA Asp94Asn						1	1			2	3
	gyrA Asp94His						2				2	3
	gyrA Asp94Ala						1				1	2
	gyrA Ala90Val, *gyrA* Asp94Tyr					1					1	2
	gyrA Asp94Asn, *gyrA* Ala90Val					1					1	2

Supplementary Table 11 Frequency and MIC-distribution of isolates with drug resistance-conferring mutations (Cont.)

Drug	Mutations [a]	MIC (µg/ml) distribution for relevant isolate									No. of isolates	%
Moxifloxacin		**≤0.06**	**0.12**	**0.25**	**0.5**	**1**	**2**	**4**	**8**	**>8**		
(CC = 1 µg/ml)	No FQ mutations	2	9	13	7		1				32	53
	gyrA Asp94Gly						4	**11**			15	25
	gyrA Ala90Val					3	2	**1**			6	10
	gyrA Asp94Asn							**1**	1		2	3
	gyrA Asp94His						1	**1**			2	3
	gyrA Asp94Ala						1				1	2
	gyrA Ala90Val, *gyrA* Asp94Tyr					1					1	2
	gyrA Asp94Asn, *gyrA* Ala90Val						1				1	2
Ethionamide		**≤0.3**	**0.6**	**1.2**	**2.5**	**5**	**10**	**20**	**40**	**>40**		
(CC = 5 µg/ml)	No ETO mutations	1	11	16	3						31	52
	ethA 639_640del		1	4	3	2	1				11	18
	ethA 704_707del					2					2	3
	inhA Ser94Ala									2	2	3
	inhA -15 c/t			1	1	2	2				6	10
	ethA 32_33insG				1						1	2
	ethA 456_456del				1						1	2
	ethA 489_531del, *ethA*_Chromosome:g.43269				1						1	2
	ethA 551_552insG			1							1	2
	ethA Thr232Ala		1								1	2
	inhA -15 c/t, *inhA* Ser94Ala						1				1	2
	inhA -8 t/c, *ethA* 1047_1047del					1					1	2
	inhA -8 t/c, *ethA* 639_640del				1						1	2
***Para*-aminosalicylic acid**		**≤0.5**	**1**	**2**	**4**	**8**	**16**	**32**	**64**	**>64**		
(CC = 1 µg/ml)	No PAS mutations	25	7	**2**				2		1	37	62
	folC Glu40Gly			**1**	4		1	2		1	9	15
	folC Ser150Gly				1	1	2			2	6	10
	folC Glu153Ala								1		1	2
	folC Glu153Gly					1					1	2
	folC Ile43Thr				1						1	2
	thyX -16 c/t		2	**1**							3	5
	thyA Thr22Ala		1								1	2
	*thyA*_Chromosome:g.3073680_3074470del, *thyX* -16 c/t			**1**							1	2

Supplementary Table 11 Frequency and MIC-distribution of isolates with drug resistance-conferring mutations (Cont.)

Drug	Mutations [a]	MIC (µg/ml) distribution for relevant isolate									No. of isolates	%
Rifabutin		**≤0.12**	**0.25**	**0.5**	**1**	**2**	**4**	**8**	**16**	**>16**		
(CC = 0.5 µg/ml)	No RFB mutations	1		**1**		1					3	5
	rpoB Ser450Leu		1	**6**	10	5	8	5	1		36	60
	rpoB His445Arg				1	1					2	3
	rpoB Ser450Leu, *rpoC* Leu527Val				1						1	2
	rpoB His445Leu	1									1	2
	rpoB His445Tyr				1		1				2	3
	rpoB Asp435Phe		1		1						2	3
	rpoB Asp435Tyr	1									1	2
	rpoB Asp435Val	1	1								2	3
	rpoB Leu452Pro	2				1					3	5
	rpoB Ser441Leu	2									2	3
	rpoB Ser450Trp						1				1	2
	rpoB Val170Phe						1				1	2
	rpoB 1295_1303del	1									1	2
	rpoB 1295_1303del, *rpoB* Ser450Leu		1								1	2
	rpoB Leu430Arg, *rpoB* Asp435Tyr					1					1	2

[a] Drug resistance-conferring mutations used in our study were based on the most recent database from TB-Profiler (https://github.com/jodyphelan/TBProfiler/blob/master/db/tbdb.dr.json).

Supplementary Table 12 Minimum inhibitory concentration values (µg/ml) of studies isolates using MYCOTBI Sensititre plate

No	SampleID	LabID	Sex	Age	AFB	Lineage	INH	RIF	EMB	STR	KAN	AMK	OFX	MFX	ETO	PAS	RFB	DCS
1	DS-33474	WMB399	F	29	2+	4.4.2	0.50	>16.00	8.00	≤0.25	1.20	≤0.12	≤0.25	≤0.06	5.00	≤0.5	8.00	8.00
2	DS-6265	WBB255	F	70	3+	4.3.4.2	1.00	>16.00	4.00	2.00	2.50	0.50	8.00	2.00	5.00	≤0.5	1.00	8.00
3	DS-6882	WBB256	M	20	3+	2.2.1	2.00	>16.00	4.00	≤0.25	1.20	0.50	4.00	1.00	1.20	1.00	4.00	8.00
4	DS-15966	WBB257	M	30	2+	2.2.1	1.00	2.00	8.00	4.00	10.00	0.50	8.00	2.00	1.20	32.00	2.00	8.00
5	DS-16179	WBB258	M	32	3+	2.2.1	1.00	16.00	4.00	>32	>40	>16	8.00	4.00	2.50	8.00	≤0.12	8.00
6	DS-16220	WBB259	M	70	2+	2.1	2.00	>16.00	4.00	4.00	>40	>16	8.00	2.00	0.60	16.00	4.00	8.00
7	DS-16780	WBB260	M	73	2+	2.1	2.00	8.00	4.00	1.00	>40	8.00	4.00	2.00	0.60	2.00	1.00	16.00
8	DS-16825	WBB261	M	32	1+	2.2.1	2.00	1.00	8.00	>32	1.20	1.00	4.00	2.00	1.20	1.00	0.50	8.00
9	DS-17012	WBB262	M	36	3+	2.2.1	1.00	4.00	4.00	4.00	>40	8.00	8.00	2.00	0.60	4.00	0.25	8.00
10	DS-17016	WBB263	F	70	1+	2.2.1	1.00	1.00	4.00	16.00	>40	>16	16.00	8.00	0.60	>64	≤0.12	8.00
11	DS-17092	WBB264	M	59	3+	2.2.1.1	2.00	>16.00	8.00	>32	>40	>16	32.00	4.00	1.20	2.00	8.00	32.00
12	DS-17653	WBB265	M	NA	3+	2.2.1	1.00	>16.00	8.00	>32	1.20	0.25	8.00	4.00	1.20	1.00	0.50	8.00
13	DS-17688	WBB266	M	NA	3+	2.2.1	1.00	>16.00	4.00	>32	>40	>16	16.00	4.00	0.60	2.00	2.00	8.00
14	DS-17841	WBB267	F	37	3+	2.1	2.00	>16.00	16.00	2.00	>40	>16	8.00	4.00	1.20	32.00	1.00	32.00
15	DS-17984	WBB268	M	41	3+	2.2.1	4.00	>16.00	2.00	>32	10.00	1.00	8.00	4.00	2.50	1.00	1.00	4.00
16	DS-18810	WBB269	M	42	1+	2.2.1	2.00	>16.00	16.00	>32	>40	>16	4.00	1.00	5.00	4.00	0.50	32.00
17	DS-19109	WBB271	F	36	3+	2.1	2.00	>16.00	16.00	4.00	>40	>16	8.00	4.00	1.20	4.00	1.00	32.00
18	DS-20120	WBB272	M	32	3+	2.2.1	0.25	>16.00	8.00	4.00	2.50	0.25	4.00	1.00	2.50	≤0.5	1.00	16.00
19	DS-25474	WMB273	M	54	3+	2.2.1.1	>4	>16.00	4.00	>32	>40	>16	16.00	4.00	5.00	>64	2.00	8.00
20	DS-27535	WBB278	F	70	3+	2.2.1	2.00	>16.00	2.00	>32	>40	16.00	8.00	2.00	0.60	2.00	2.00	16.00
21	DS-28473	WMB279	M	41	3+	2.2.2	2.00	2.00	8.00	≤0.25	≤0.6	0.25	0.50	0.12	1.20	≤0.5	4.00	8.00
22	DS-29128	WBB279	M	36	3+	2.2.1	1.00	>16.00	16.00	>32	>40	>16	8.00	4.00	1.20	1.00	2.00	8.00
23	DS-29366	WBB281	M	31	2+	2.1	4.00	>16.00	8.00	8.00	>40	>16	16.00	4.00	1.20	4.00	1.00	64.00
24	DS-30056	WBB282	F	35	1+	2.2.2	1.00	>16.00	8.00	>32	1.20	0.50	32.00	4.00	10.00	1.00	4.00	8.00
25	DS-32449	WBB285	F	64	2+	2.1	4.00	>16.00	8.00	4.00	>40	>16	16.00	4.00	1.20	4.00	1.00	64.00
26	DS-32512	WBB286	F	36	3+	2.2.1	0.12	>16.00	4.00	4.00	>40	>16	4.00	1.00	1.20	1.00	0.50	8.00
27	***M. tuberculosis***	**H37Rv**	**ATCC27294**				0.06	≤0.12	2.00	1.00	2.50	0.50	1.00	0.25	2.50	≤0.5	≤0.12	16.00
28	DS-36687	M36687	M	27	1+	4.4.2	0.50	>16.00	8.00	0.50	1.20	0.25	1.00	0.25	10.00	≤0.5	8.00	4.00
29	DS-36881	M36881	M	26	3+	2.2.1	4.00	>16.00	2.00	>32	1.20	0.25	1.00	0.25	1.20	>64	4.00	8.00
30	DS-36914	M36914	M	28	2+	2.2.1	2.00	>16.00	4.00	>32	1.20	0.25	1.00	0.25	2.50	≤0.5	1.00	8.00
31	DS-36982	M36982	M	65	1+	4.3.3	4.00	>16.00	4.00	2.00	2.50	0.50	1.00	0.50	2.50	≤0.5	4.00	16.00
32	DS-37032	M37032	M	31	3+	2.2.1	2.00	>16.00	4.00	>32	1.20	0.25	8.00	2.00	2.50	≤0.5	1.00	16.00

Supplementary Table 12 Minimum inhibitory concentration values (µg/ml) of studies isolates using MYCOTBI Sensititre plate (Cont.)

No	SampleID	LabID	Sex	Age	AFB	Lineage	INH	RIF	EMB	STR	KAN	AMK	OFX	MFX	ETO	PAS	RFB	DCS
33	DS-37105	M37105	M	58	1+	2.2.1	2.00	>16.00	4.00	>32	1.20	0.25	0.50	0.12	1.20	≤0.5	0.50	8.00
34	DS-37195	M37195	M	35	2+	1.2.1	0.50	8.00	2.00	≤0.25	≤0.6	0.25	2.00	0.50	>40	1.00	0.25	8.00
35	DS-37242	M37242	M	34	3+	2.2.1	1.00	>16.00	2.00	>32	≤0.6	0.25	1.00	0.25	1.20	≤0.5	1.00	4.00
36	DS-37378	M37378	M	42	3+	2.2.1	2.00	>16.00	4.00	>32	≤0.6	0.25	0.50	0.25	0.60	4.00	1.00	8.00
37	DS-37446	M37446	M	39	3+	4.5	1.00	2.00	2.00	0.50	2.50	0.25	8.00	2.00	≤0.3	≤0.5	≤0.12	≤2
38	DS-40543	M40543	M	50	NA	2.2.1	4.00	8.00	1.00	>32	>40	>16	16.00	4.00	5.00	64.00	0.25	32.00
39	DS-34062	WMB299	M	NA	1+	2.2.1	2.00	>16.00	8.00	>32	1.20	0.25	16.00	4.00	1.20	16.00	4.00	16.00
40	DS-41960	M41960	M	36	3+	1.2.1	0.50	4.00	2.00	16.00	1.20	0.25	1.00	0.25	>40	2.00	1.00	8.00
41	DS-42070	M42070	M	32	3+	2.2.1	2.00	>16.00	4.00	32.00	1.20	0.25	0.50	0.12	2.50	≤0.5	8.00	8.00
42	DS-42084	M42084	M	46	1+	2.2.1	2.00	>16.00	2.00	>32	1.20	0.25	1.00	0.25	1.20	8.00	2.00	16.00
43	DS-42309	M42309	M	33	3+	2.2.1	2.00	>16.00	8.00	>32	≤0.6	0.25	1.00	0.25	0.60	1.00	0.25	8.00
44	DS-42412	M42412	M	56	2+	2.2.1	2.00	≤0.12	≤0.5	>32	≤0.6	0.25	1.00	0.25	1.20	≤0.5	≤0.12	8.00
45	DS-42539	M42539	M	25	2+	2.2.1	0.25	>16.00	2.00	>32	≤0.6	0.25	1.00	0.25	2.50	≤0.5	0.50	8.00
46	DS-41879	M41879	M	38	3+	2.2.1	0.12	0.50	8.00	>32	1.20	0.25	0.50	0.25	2.50	≤0.5	≤0.12	16.00
47	DS-41888	M41888	M	27	2+	2.2.1	2.00	0.25	1.00	0.50	1.20	0.25	0.50	0.12	1.20	≤0.5	2.00	4.00
48	DS-41955	M41955	M	78	3+	2.2.1	0.50	>16.00	1.00	≤0.25	1.20	0.25	0.50	0.12	0.60	≤0.5	2.00	4.00
49	DS-42002	M42002	M	64	1+	2.2.1	1.00	0.25	1.00	≤0.25	1.20	0.25	0.50	0.12	0.60	≤0.5	≤0.12	8.00
50	DS-42067	M42067	M	26	3+	2.2.1.1	1.00	>16.00	1.00	≤0.25	≤0.6	0.25	1.00	0.25	0.60	≤0.5	0.50	8.00
51	DS-42122	M42122	M	48	2+	1.1.3	2.00	0.25	1.00	≤0.25	≤0.6	≤0.12	0.50	0.12	1.20	≤0.5	≤0.12	≤2
52	DS-42137	M42137	F	39	1+	2.2.1	1.00	0.25	2.00	≤0.25	≤0.6	≤0.12	0.50	≤0.06	1.20	≤0.5	4.00	4.00
53	DS-42443	M42443	M	57	3+	1.2.1	1.00	1.00	1.00	≤0.25	≤0.6	≤0.12	1.00	0.50	0.60	≤0.5	≤0.12	4.00
54	DS-37947	M37947	M	42	3+	2.2.1	2.00	>16.00	8.00	>32	1.20	0.25	0.50	0.12	1.20	16.00	4.00	16.00
55	DS-39181	M39181	M	42	3+	2.1	4.00	>16.00	4.00	0.50	2.50	0.50	0.50	0.25	5.00	32.00	4.00	4.00
56	DS-39597	M39597	M	32	3+	2.1	>4	>16.00	8.00	0.50	1.20	0.50	1.00	0.50	10.00	>64	2.00	4.00
57	DS-39954	M39954	M	29	3+	2.2.1	4.00	>16.00	8.00	>32	1.20	0.50	2.00	0.50	0.60	≤0.5	1.00	64.00
58	DS-40579	M40579	F	42	3+	2.2.1	>4	>16.00	4.00	>32	>40	>16	1.00	0.12	2.50	≤0.5	≤0.12	64.00
59	DS-40949	M40949	M	42	3+	2.1	4.00	>16.00	4.00	0.50	2.50	0.50	8.00	2.00	5.00	32.00	8.00	4.00
60	DS-40994	M40994	M	81	1+	2.2.1	4.00	>16.00	2.00	>32	1.20	1.00	2.00	0.50	10.00	≤0.5	4.00	64.00
61	DS-40320	M40320	M	52	Negative	4.5	4.00	>16.00	2.00	>32	1.20	0.25	1.00	0.50	2.50	1.00	16.00	4.00

Supplementary Table 13 Provinces of Thailand and abbreviations

Regions	Provinces	Abbreviations	Regions	Provinces	Abbreviations
Central	Bangkok	BKK	**Northeastern**	Amnat Charoen	ACR
	Chai Nat	CNT		Bungkan	BKN
	Kamphaeng Phet	KPT		Buri Ram	BRM
	Lop Buri	LRI		Chaiyaphum	CPM
	Nakhon Nayok	NYK		Kalasin	KSN
	Nakhon Pathom	NPT		Khon Kaen	KKN
	Nakhon Sawan	NSN		Loei	LEI
	Nonthaburi	NBI		Maha Sarakham	MKM
	P.Nakhon S.Ayutthaya	AYA		Mukdahan	MDH
	Pathum Thani	PTE		Nakhon Phanom	NPM
	Phetchabun	PNB		Nakhon Ratchasima	NMA
	Phichit	PCT		Nong Bua Lam Phu	NBP
	Phitsanulok	PLK		Nong Khai	NKI
	Samut Prakan	SPK		Roi Et	RET
	Samut Sakhon	SKN		Sakon Nakhon	SNK
	Samut Songkhram	SKM		Si Sa Ket	SSK
	Saraburi	SRI		Surin	SRN
	Sing Buri	SBR		Ubon Ratchathani	UBN
	Sukhothai	STI		Udon Thani	UDN
	Suphan Buri	SPB		Yasothon	YST
	Uthai Thani	UTI			
Eastern	Chachoengsao	CCO	**Northern**	Chiang Mai	CMI
	Chanthaburi	CTI		Chiang Rai	CRI
	Chon Buri	CBI		Lampang	LPG
	Prachin Buri	PRI		Lamphun	LPN
	Rayong	RYG		Phayao	PYO
	Sa Kaeo	SKW		Phrae	PRE
	Trat	TRT		Uttaradit	UTT

Supplementary Table 12 Provinces of Thailand and abbreviations (Cont.)

Regions	Provinces	Abbreviations	Regions	Provinces	Abbreviations
Southern	Chumphon	CPN	**Western**	Kanchanaburi	KRI
	Krabi	KBI		Phetchaburi	PBI
	Nakhon Si Thammarat	NST		Prachuap Khiri Khan	PKN
	Narathiwat	NWT		Ratchaburi	RBR
	Pattani	PTN		Tak	TAK
	Phang Nga	PNA			
	Phatthalung	PLG			
	Phuket	PKT			
	Ranong	RNG			
	Satun	STN			
	Songkhla	SKA			
	Surat Thani	SNI			
	Trang	TRG			
	Yala	YLA			

RESEARCH PUBLICATIONS

1. **Nonghanphithak D**, Reechaipichikul W, Namwat W, Wongwajana S, Lulitanond V, Faksri K. Comparison of CXCL9 polymorphism between pulmonary tuberculosis patients and healthy controls in northeast Thailand. *Srinagarind Med J* 2015; 30: 432-38.
2. **Nonghanphithak D,** Reechaipichitkul W, Chaiyasung T, Namwat W, Lulitanond V, Naranbhai V, Faksri K. Genetic polymorphisms of *CCL2* associated with susceptibility to latent tuberculous infection in Thailand. *Int J Tuberc Lung Dis* 2016; 20: 1242-1248.
3. **Nonghanphithak D**, Reechaipichitkul W,Chaiyasung T, Faksri K. Risk factors for latent tuberculosis infection among health-care workers in Northeastern Thailand. *Southeast Asian J Trop Med Public Health* 2016; 47: 1198-1208.
4. **Nonghanphithak D**, Reechaipichitkul W, Namwat W, Naranbhai V, Faksri K. Chemokines additional to IFN-gamma can be used to differentiate among *Mycobacterium tuberculosis* infection possibilities and provide evidence of an early clearance phenotype. *Tuberculosis (Edinb)* 2017; 105: 28-34.
5. Faksri K, Xia E, Ong RT, Tan JH, **Nonghanphithak D**, Makhao N, et al. Comparative whole-genome sequence analysis of *Mycobacterium tuberculosis* isolated from tuberculous meningitis and pulmonary tuberculosis patients. *Sci Rep* 2018; 8(1): 4910.
6. **Nonghanphithak D**, Kaewprasert O, Chaiyachat P, Reechaipichitkul W, Chaiprasert A, Faksri K. Whole-genome sequence analysis and comparisons between drug-resistance mutations and minimum inhibitory concentrations of *Mycobacterium tuberculosis* isolates causing M/XDR-TB. *PLoS One* 2020; 15(12): e0244829.
7. **Nonghanphithak D,** Chaiprasert A, Smithtikarn S, Kamolwat P, Pungrassami P, Chongsuvivatwong V, et al. Clusters of Drug-Resistant *Mycobacterium tuberculosis* Detected by Whole-Genome Sequence Analysis of Nationwide Sample, Thailand, 2014-2017. *Emerging Infectious Diseases* 2021; 27(3): 813-822.

CURRICULUM VITAE

Name:	Mr. Ditthawat Nonghanphithak
Day of Birth:	June 17th, 1991
Place of Birth:	Udon Thani, Thailand
Address:	372 Moo 15 Tambon Sam Phrao
	Amphoe Mueang Udon Thani Thailand 41000
For more information:	ditthawat@kkumail.com
	ditthawat_2534@hotmail.com

Education:

2013-2015	Master of Science (Medical Microbiology)
	Faculty of Medicine
	Khon Kaen University, Khon Kaen, Thailand
	Thesis: "Susceptible factors for *Mycobacterium tuberculosis* infection among tuberculosis close contacts and health care workers in Northeast Thailand"
2009 - 2012	Bachelor of Science (Medical Technology)
	Faculty of Associated Medical Science
	Khon Kaen University, Khon Kaen, Thailand.
	Project: "The Comparative Study of Thrombomodulin Gene in Epidermal Growth Factor 4-6 Like Domain Involving Protein C Function between Normal Healthy Control and Thai Patients with Deep Vein Thrombosis"