

Transmission of rifampicin-resistant tuberculosis in Ho Chi Minh City, Viet Nam: a prospective genomic epidemiology study



Ruan Spies,^{a,b,*} Nguyen Hong Hanh,^a Phan Trieu Phu,^a Luong Kim Lan,^a Kim Lan,^a Ngo Ngoc Hue,^a Nguyen Le Quang,^a Do Dang Anh Thu,^a Nguyen Thi Le Huong,^a Tran Le Thi Ngoc Thao,^a Trinh Thi Bich Tram,^a Vu Thi Ngoc Ha,^a Dang Thi Minh Ha,^c Nguyen Phuc Hai,^c Nguyen Hung Thuan,^c Tran Thi Kim Quy,^c Nguyen Huu Lan,^c Viola Dreyer,^{d,e} Stefan Niemann,^{d,e,f} Derrick Crook,^{g,h} Le Hong Van,^a Guy Thwaites,^{a,b} Nguyen Thuy Thuong Thuong,^{a,b} Marc Choisy,^{a,b} James A. Watson,^{b,i} and Timothy M. Walker^{a,b,g}



^aOxford University Clinical Research Unit, Ho Chi Minh City, Viet Nam

^bCentre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, United Kingdom

^cPham Ngoc Thach Hospital, Ho Chi Minh City, Viet Nam

^dGerman Center for Infection Research (DZIF), Partner Site Hamburg-Lübeck-Borstel-Riems, Borstel, Germany

^eResearch Center Borstel, Leibniz Lung Center, Molecular and Experimental Mycobacteriology, Borstel, Germany

^fResearch Center Borstel, Leibniz Lung Center, Supranational and National Reference Center for Mycobacteria, Borstel, Germany

^gModernising Medical Microbiology Unit, Nuffield Department of Medicine, University of Oxford, United Kingdom

^hShared Hospital Laboratory and Sunnybrook Research Institute, Sunnybrook Hospital, Toronto, Canada

ⁱInfectious Diseases Data Observatory, Oxford, United Kingdom

Summary

Background Rifampicin-resistant tuberculosis (RR-TB) is a major threat to public health in Viet Nam, with nearly 10,000 incident cases estimated annually. It is uncertain whether these cases are driven by transmission of resistant strains or *de novo* resistance acquisition during treatment.

Methods We undertook dense, city-wide sampling of adults newly diagnosed with pulmonary RR-TB in Ho Chi Minh City, Viet Nam's largest city, between March 2020 and April 2024. Participants provided sputum for culture and whole-genome sequencing (WGS), and demographic and clinical data were collected at enrolment. Phylogenetic analyses were combined with clinical histories to infer transmitted versus acquired rifampicin resistance. Estimates were corrected for sampling coverage using simulation-extrapolation (SIMEX). Temporal emergence of rifampicin resistance was reconstructed by lineage using Bayesian phylogenetic dating, and the geographic and demographic structure of transmission clades was assessed using geocoded residential data and commute time-based analyses.

Findings Among 2319 RR-TB cases diagnosed during the study period, 1491 (64%) isolates were successfully sequenced. Among these, 1320 (89%) were Lineage 2. After accounting for sampling and phylogenetic uncertainty, we estimated that 72–87% of RR-TB was attributable to transmission of already-resistant strains, with the remainder reflecting *de novo* acquired resistance. Bayesian dating analyses revealed that resistance emergence events occurred repeatedly from the 1980s to the present, with early events seeding long-lived, city-wide transmission clades. Transmission clades were geographically dispersed across the city, with limited household clustering, and only weakly structured by host demographics, consistent with diffuse, city-wide transmission rather than localised or assortative spread.

Interpretation RR-TB in Ho Chi Minh City is driven predominantly by ongoing transmission, but a substantial minority of cases arise from newly acquired resistance. Alongside promoting early diagnosis and treatment to interrupt transmission, the main drivers of acquired resistance need to be identified to control RR-TB.

Funding The Rhodes Trust, Wellcome.

Copyright © 2026 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The Lancet Regional Health - Western Pacific 2026;71: 101906

Published Online xxx
<https://doi.org/10.1016/j.lanwpc.2026.101906>

*Corresponding author. Oxford University Clinical Research Unit, Hospital for Tropical Diseases, 764 Vo Van Kiet Street, Ward 1, District 5, Ho Chi Minh City, 700000, Viet Nam.

E-mail address: rspies@oucru.org (R. Spies).

Keywords: Tuberculosis; Rifampicin-resistant tuberculosis; Multidrug-resistant tuberculosis; Whole-genome sequencing; Phylogenetic analysis; Genomic epidemiology; Transmission dynamics; Viet Nam; Ho Chi Minh City

Research in context

Evidence before this study

We searched PubMed from database inception up to November 2025 for studies describing the transmission and acquisition of multidrug/rifampicin-resistant tuberculosis (MDR/RR-TB), using terms including “tuberculosis”, “rifampicin resistance”, “multidrug resistance”, “genomic epidemiology”, “whole-genome sequencing”, “transmission” and “acquired resistance”. We additionally added terms “Ho Chi Minh City” and “Viet Nam” to identify results specific to our study setting. Previous studies from high burden settings typically reported that MDR/RR-TB isolates clustered with high proportions, suggesting transmission is the major contributor to MDR/RR-TB burden. However, estimates varied widely across settings, study designs and analytic approaches. Many studies relied on SNP-based clustering thresholds and did not account for incomplete sampling. Few studies integrated treatment history and homoplasmy-aware phylogenetic inference or accounted for uncertainty in their estimates of the population-level contribution of *de novo* rifampicin resistance acquisition. No studies from Viet Nam addressed this question and the extent to which transmitted and acquired resistance contribute to the burden of MDR/RR-TB remains unknown.

Added value of this study

We combined dense, city-wide sampling of RR-TB in a high-burden metropolitan setting with whole-genome sequencing, historic genomic context and phylogenetic

analyses to quantify the contributions of transmission and *de novo* acquisition to MDR/RR-TB burden. By integrating phylogenetic placement of resistance mutations with individual treatment history, we provide robust population-level estimates of transmission and independent resistance emergence while explicitly accounting for phylogenetic and sampling uncertainty. Acknowledging the potential effect of sampling coverage on these estimates, we further provide estimates extrapolated to complete coverage using a simulation-extrapolation framework. We show that the majority of RR-TB in Ho Chi Minh City is attributable to transmission, but that a substantial minority of cases reflect *de novo* acquisition of rifampicin resistance. Our findings demonstrate that *de novo* resistance acquisition continues to occur notwithstanding improvements in programmatic TB control, treatment success rates, and diagnostic capacity.

Implications of all the available evidence

Our findings emphasise that MDR/RR-TB in Ho Chi Minh City is overwhelmingly driven by transmission and that controlling it requires a transmission-focused response. Intensified efforts to interrupt transmission through earlier case detection and rapid treatment initiation should be programmatic priorities. Concurrently, further investigation into the precise drivers of acquired drug resistance, which represents a potentially preventable fraction of the RR-TB burden, remains warranted.

Introduction

The global incidence of tuberculosis (TB) is falling, yet an estimated 10.7 million people developed TB and 390,000 developed rifampicin-resistant TB (RR-TB) in 2024, and the pace of decline remains insufficient to meet the World Health Organization’s End TB targets.¹ In Viet Nam, TB incidence mirrors the global downward trend, but the burden of RR-TB has remained largely unchanged over the last decade, with nearly 10,000 new cases estimated annually.¹ Historical data from Ho Chi Minh City, the country’s largest city, indicate that most people diagnosed with RR-TB have had a previous episode of TB.² This may be because RR-TB is only more recently being correctly diagnosed, or because resistance has been selected through inadequate treatment.² Either way, the scale of the public health problem underscores the importance of understanding what is driving RR-TB in this setting.

Central to this question is distinguishing between the two pathways by which rifampicin resistance can arise: transmission of resistant strains or *de novo* acquisition of

resistance mutations during treatment for rifampicin-susceptible disease. This distinction has direct public health relevance, as each pathway requires different interventions. Interrupting transmission requires interventions that facilitate rapid diagnosis and treatment, such as active case finding,³ whereas preventing acquired resistance depends on patients receiving the right number of the right drugs, at the right dose, for the right duration.⁴

When drug-resistant TB first emerged as a public health concern in the 1990s, rifampicin resistance was believed to arise mainly from inadequate treatment, and resistant strains were thought to carry fitness costs that limited their capacity for widespread transmission.⁵ The Directly Observed Treatment, short course (DOTS) strategy, standardised treatment regimens, and fixed-dose drug combinations were designed to prevent such acquired resistance.⁶ As these measures became established, our understanding of the epidemic evolved, with molecular epidemiology and modelling studies subsequently demonstrating that resistant *Mycobacterium*

tuberculosis strains transmit effectively, and that transmission of already-resistant strains likely accounts for most RR-TB globally.^{7,8}

Viet Nam's mature National TB Programme (NTP) includes long-standing safeguards against acquired resistance including DOTS, fixed-dose drug combinations and uninterrupted access to 1st line drugs,⁹ suggesting the contemporary RR-TB burden should be driven predominantly by transmission of established resistant clones. We conducted comprehensive sampling and whole-genome sequencing (WGS) of rifampicin-resistant *M. tuberculosis* isolates across Ho Chi Minh City over a four-year period to test this hypothesis and inform targeted control strategies by identifying risk-factors for transmission.

Methods

Study setting and design

Ho Chi Minh City has a population of approximately 10 million people and a high burden of TB and RR-TB.¹⁰ All individuals diagnosed with RR-TB in Ho Chi Minh City are referred to initiate treatment at Phạm Ngọc Thạch Hospital, the city's specialist centre for TB and lung disease, which provides treatment for approximately 80% of drug-resistant TB cases in Viet Nam.¹¹

We conducted a prospective cohort study aiming to enrol all adult residents (≥ 18 years) newly diagnosed with pulmonary RR-TB by Xpert MTB/RIF between March 2020 and May 2024. Eligible participants provided written informed consent at their first appointment at Phạm Ngọc Thạch Hospital. At enrolment, and prior to initiating treatment, participants submitted a sputum specimen for culture and WGS and completed a standardised case-report form recording demographic characteristics and clinical history. Not all eligible individuals were enrolled, primarily due to temporary recruitment interruptions during COVID-19 lockdowns.

This study received ethical approval from the Institutional Review Board at Phạm Ngọc Thạch Hospital (643/PNT-HDDD) and the Oxford Tropical Research Ethics Committee (reference no. 51–19).

Sample processing

Sputum sample processing and culture have been described previously.¹² Genomic DNA was extracted from cultured isolates using the cetyl trimethylammonium bromide (CTAB) method^{13,14} and prepared for sequencing using a modified Nextera XT library preparation protocol,¹⁵ generating paired-end reads on Illumina NextSeq 1000/2000 platforms (Illumina, San Diego, CA, USA) with a target mean depth of $\geq 50\times$ (Supplementary Methods). Bioinformatic processing was performed using the GPAS *Mycobacterium* pipeline v2.2.1 to determine species and lineage, predict drug resistance and generate intermediate files for phylogenetic analyses.¹⁶ Resistance

classification was based on genotypic prediction with phenotypic drug susceptibility testing not performed.

Phylogenetic analysis

To place the newly sequenced isolates in evolutionary context, we combined them with previously published *M. tuberculosis* genomes from Ho Chi Minh City and additional genomes we generated in a linked contemporary study (Supplementary Methods).^{17–19} This combined dataset included both rifampicin-resistant and rifampicin-susceptible isolates, providing the context required to identify the emergence of resistance mutations. Maximum-likelihood phylogenies were inferred using IQ-TREE v3.0.1 (Supplementary Methods).²⁰ Samples with evidence of mixed-lineage infection, defined by the detection of lineage-specific SNPs from more than one *M. tuberculosis* sub-lineage, were excluded because their phylogenetic placement could not be determined with confidence.

We classified rifampicin resistance as arising through transmission (primary resistance) or *de novo* acquisition (secondary resistance) by first looking at the clinical history and then the phylogenetic placement of resistance-conferring mutations. Participants with no prior TB were classified as having developed RR-TB through transmission, regardless of phylogenetic position, as they had never received anti-tuberculosis treatment and therefore had no opportunity for treatment-driven resistance selection. Among participants with previous TB, classification was determined by how resistance mutations mapped to branches in the phylogeny (Fig. 1). To this end, we identified and mapped mutations in the *rpoB* gene that had arisen independently across multiple branches of the phylogeny using SNPPar²¹ and annotated these according to the World Health Organisation mutation catalogue.²² Where a resistance mutation mapped to the branch leading to a single isolate, we assumed the most parsimonious interpretation that resistance emerged in that individual *de novo* rather than being inherited through transmission from an ancestor.

When a resistance mutation mapped to an internal node in the phylogeny there was no means of determining whether all sampled descendant strains inherited it through transmission or whether it actually emerged *de novo* in one of those strains with subsequent transmission to the other descendants. We therefore undertook two alternative analyses: the first (hereafter 'N') classified all N descendants of an internal resistance node as having developed RR-TB through transmission, providing a lower bound on the number of *de novo* acquired resistance events. The second approach (hereafter 'N -1') attributed *de novo* acquired resistance to one descendant strain so long as at least one of the patients from whom a descendant strain was obtained had a history of prior TB. This approach makes the parsimonious assumption that the remaining N -1 descendants

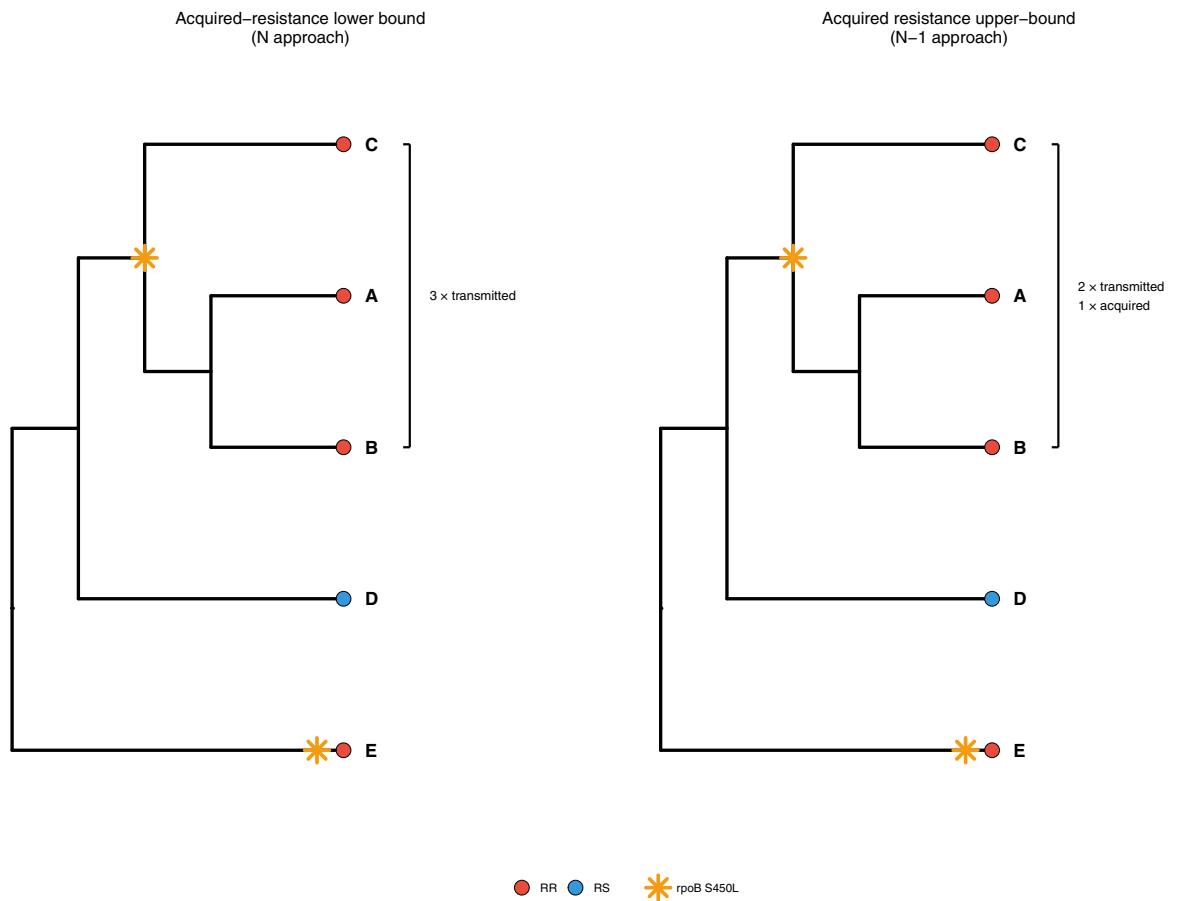


Fig. 1: Schematic illustration of N and N-1 classification approaches for acquired rifampicin resistance. Schematic phylogeny showing five sampled *Mycobacterium tuberculosis* isolates (A–E) with rifampicin-resistant (RR) and rifampicin-susceptible (RS) phenotypes indicated. The emergence of rifampicin resistance is indicated by a homoplasic rpoB S450L mutation, represented by the yellow star. When such a resistance-conferring mutation maps to an internal node of the phylogeny, its origin cannot be uniquely attributed to transmission or within-host acquisition. Under the N approach, all descendants of an internal resistance-associated node (A, B and C) are classified as having transmitted resistance, resulting in three transmitted RR cases in this example. Under the N-1 approach, the presence of at least one previously treated individual among the descendants permits one event to be classified as acquired resistance and the remaining as transmitted. These approaches therefore define plausible lower and upper bounds on the proportion of rifampicin-resistant TB attributable to acquired resistance in the presence of phylogenetic ambiguity. Isolate E represents a tip-level homoplasy classified as acquired resistance under both approaches.

inherited resistance through transmission (Fig. 1). Together, these approaches defined plausible ranges for the proportion of RR-TB arising through transmission and *de novo* acquisition.

Uncertainty and sensitivity analyses

We implemented a bootstrap procedure to account for uncertainty from both sampling variability and phylogenetic inference. In each of 100 iterations, the full dataset was resampled with replacement. Because phylogenetic inference requires unique sequences, trees were constructed using the unique set of resampled isolates, while the original resampling weights were retained for estimation. Classification of transmitted versus acquired

resistance was performed under both the N and N-1 approaches, with each sample's contribution weighted by its resampling frequency. Convergence was assessed at both population and sample levels by monitoring the running mean of the proportion of acquired resistance across replicates and the stability of individual sample classification probabilities. The proportion of acquired RR-TB was summarised as the plausible range spanned by the two classification approaches, with empirical 95% confidence intervals derived from the 2.5th and 97.5th percentiles of the bootstrap distributions.

Because apparent independent emergence could actually be the consequence of an infection with a resistant strain from outside Ho Chi Minh City

(importation) rather than true *de novo* acquisition, we conducted a sensitivity analysis with respect to birthplace. Among participants with a history of previous TB treatment, we used fractional logistic regression with the bootstrap-derived probability of acquired resistance (under the N approach) as the outcome to estimate the odds of acquired versus transmitted rifampicin resistance by birthplace. Using the same approach, we also conducted exploratory univariate analyses of potential risk factors for acquired resistance, including *M. tuberculosis* lineage, age, sex, HIV status, diabetes, BMI, smoking, alcohol use, and intravenous drug use.

Correction for incomplete sampling

Incomplete sampling may bias classification towards a label of *de novo* acquired resistance. When cases are missing from the dataset, true transmission networks may be fragmented, causing mutations that arose on shared internal branches to instead appear as independent events. To address this, we applied a simulation-extrapolation (SIMEX) framework.²³ We simulated progressively sparser sampling of the RR-TB isolates collected in this study (for which we were able to estimate sampling coverage) by subsampling the dataset at ten evenly-spaced fractions, retaining between 8% and 92% of the sampled isolates. We rebuilt phylogenies at each fraction and used generalised additive models (GAMs) to characterise how the estimated proportion of *de novo* acquired resistance changed as fewer cases were sampled. We then used the fitted GAMs to extrapolate to 100% sampling coverage. To further assess whether estimates were also sensitive to the temporal depth of sampling (i.e., if estimates would differ had sampling begun earlier than 2020), we repeated the analysis holding sample size constant while varying the range of years sample from one to five ([Supplementary Methods](#)).

Temporal emergence of rifampicin resistance

To estimate when rifampicin resistance emerged in the major *M. tuberculosis* lineages circulating in Viet Nam (Lineages 1, 2, and 4), we performed Bayesian phylogenetic dating using BactDating.²⁴ We analysed each lineage separately to allow for lineage-specific mutation rates. For Lineages 1 and 2, molecular clock rates were estimated from the data using a subsampling approach; for Lineage 4, we applied a previously published rate (0.237 substitutions/genome/year)²⁵ as there was insufficient temporal signal to reliably estimate a context-specific rate ([Supplementary Methods](#)). The timing of resistance emergence was inferred from the estimated age of the most recent common ancestor of each resistant clade.

Geographic and demographic structure of transmission clades

Transmission clades were defined as monophyletic clades descending from a shared internal node where a

rifampicin-resistance mutation had arisen. Residential addresses for participants from the present study were geocoded using the Google Maps Geocoding API via the *tidygeocoder* R package.²⁶ To assess whether household transmission was a major driver of RR-TB, we identified cases sharing the same residential address and examined whether co-resident cases belonged to the same or different transmission clades.

We characterised the structure of RR-TB transmission along three dimensions. To assess whether transmission was geographically localised or dispersed, we mapped transmission clades, measured the dispersion of each network using median pairwise commute times between cases derived from the Open-Source Routing Machine (OSRM),²⁷ and tested whether genetically related cases lived closer together than unrelated cases using the Mantel test. Dispersion was quantified for networks with ≥ 2 geocoded cases, and the association between dispersion and estimated year of emergence was assessed using Spearman's rank correlation. Within-versus between-network commute times were compared using the Wilcoxon rank-sum test. To assess whether transmission was occurring preferentially between individuals sharing demographic characteristics (assortative transmission), we used treeSeg,²⁸ a multiscale change-point approach that tests for dependence between a trait and the phylogenetic tree structure at all levels of the tree hierarchy simultaneously while controlling the false positive rate at a prespecified level ($\alpha = 0.05$). Binary traits (sex, HIV status, diabetes, birthplace [Ho Chi Minh City versus elsewhere] and drug use history) were tested directly and age was tested as a continuous variable under a Gaussian model. Demographic distributions were also compared across large transmission clades (≥ 10 members). To quantify the joint explanatory value of geographic proximity and demographic similarity for genetic relatedness, we fitted multiple regression on distance matrices (MRM) ([Supplementary Methods](#)).

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Between March 1st, 2020 and April 30th, 2024, 2596 Ho Chi Minh City residents were diagnosed with pulmonary RR-TB by routine services using Xpert MTB/RIF. We enrolled 2033 (78%) of these individuals into our prospective cohort study. Reasons for non-enrolment were not formally recorded but were largely attributed to temporary recruitment interruptions during COVID-19 lockdowns. Among enrolled participants, 1698 (84%) had sputum samples successfully cultured and whole-genome sequenced. Among these, WGS data indicated

that 108 (6%) were rifampicin-susceptible and that 74 (4%) were non-tuberculous mycobacteria (NTM). 25 (1%) had insufficient sequence coverage for resistance prediction (Fig. 2). We considered the rifampicin-susceptible and NTM isolates (182/1698) to represent programmatic false positives. Applying this 10.7% false positive rate to the entire RR-TB population, we estimated that we successfully sequenced 1491/2319 (64%) of all true RR-TB cases during the study period. Participants without a sequenced isolate did not differ from those with sequencing data, apart from a higher proportion with smear-negative disease (Supplementary Table S1). Most rifampicin resistance was due to the *rpoB* S450L mutation (Supplementary Table S2).

1070 (72%) participants were male, mean age was 46 and mean BMI 19.5. 802 (54%) participants had previous TB and 435 (29%) self-reported a diagnosis of diabetes. Participants lived in households with an average of 3.6 additional members (Table 1). Most isolates were Lineage 2 [1320 (89%)] (Fig. 3).

We incorporated 2922 previously published *M. tuberculosis* genomes from Ho Chi Minh City alongside 3050 contemporary sequences generated in this and a parallel study to increase phylogenetic resolution and provide the ancestral context needed to detect the emergence of resistance in our phylogenetic analysis. After excluding 254 mixed-lineage sequences, the combined dataset comprised 5718 genomes, half of which were rifampicin-susceptible (Supplementary Table S3, Supplementary Fig. S1). Model comparisons identified the TVM nucleotide substitution model as the best fit for reconstructing the phylogeny.

Among the data this study generated (64% sampling coverage), the N approach estimated that 19.5% (17.8–21.6%) of RR-TB was attributable to acquired resistance compared to 29.4% (27.7–31.6%) under N-1 approach. Among people with previous TB (under the N approach), 36% (34–39%) of RR-TB was estimated to arise through *de novo* acquisition and 64% (61–66%)

was attributed to transmission. Extrapolating to complete sampling coverage using SIMEX, we estimated a plausible range for *de novo* acquired rifampicin resistance of 13–28%. In the sensitivity analysis, temporal depth of RR-TB sampling had minimal independent effect on classification when sample size was held constant, suggesting that starting sampling earlier than 2020 would not have substantially changed our estimates (Supplementary Results).

Across the three major circulating lineages in Viet Nam (Lineages 1, 2, and 4), we identified 220 rifampicin-resistance emergence events, of which the *rpoB* S450L mutation was consistently the most frequent across lineages [122 (55%)]. The earliest inferred emergence occurred in Lineage 2 and was dated to 1980 (95% credible interval 1970–1989). Viet Nam's NTP introduced rifampicin nationwide from 1986.⁹ For Lineages 1 and 4, the estimates were later, although the credible intervals were similarly wide. Subsequent emergence of resistance events were estimated to have continued to the present day (Fig. 4).

When examining predictors of transmitted versus acquired rifampicin resistance (among participants with previous TB), we found no evidence that birthplace (Ho Chi Minh City versus elsewhere) influenced the inferred pathway to resistance (OR 0.92 [95% CI 0.69–1.23]). In contrast, infection with *M. tuberculosis* Lineage 2 was associated with lower odds of acquired (relative to transmitted) resistance (OR 0.23 [0.14–0.36]), whereas diabetes was associated with higher odds of acquired resistance (OR 1.58 [1.28–2.10]). No other host or pathogen covariates showed statistically significant associations (Supplementary Table S4).

We analysed transmission clades, defined as monophyletic groups of isolates descending from a shared internal resistance-emergence node, to estimate where in the city transmission was occurring. After excluding historic cases for whom residential addresses were unavailable, we identified 213 monophyletic RR-TB

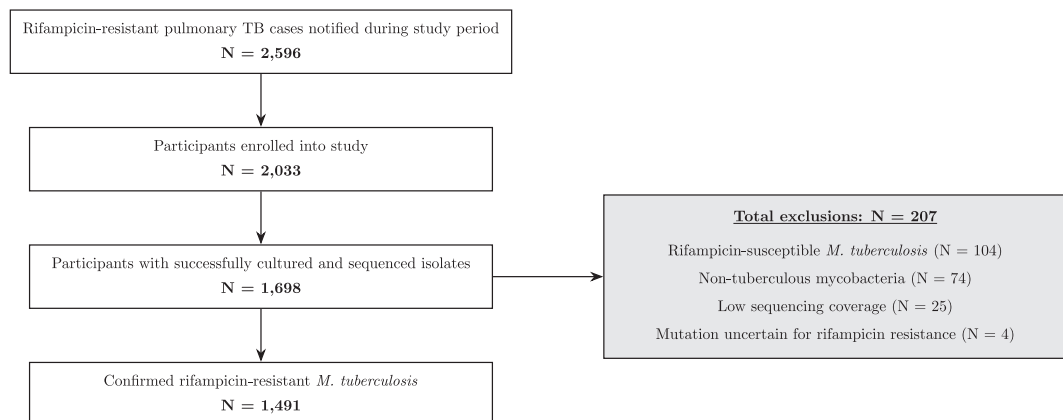


Fig. 2: Study flow diagram.

	Overall (N = 1491)	Previous TB (N = 802)	No Previous TB (N = 689)
Age, mean (SD)	46 (14)	48 (14)	43 (15)
Male sex, n	1070 (71.8%)	625 (77.9%)	445 (64.6%)
Body mass index, mean (SD)	19.5 (3.20)	19.2 (3.18)	19.7 (3.20)
Additional people in household, mean (SD)	3.60 (2.56)	3.62 (2.66)	3.59 (2.45)
HIV positive, n	90 (6.0%)	67 (8.4%)	23 (3.3%)
Diabetes, n	435 (29.2%)	249 (31.0%)	186 (27.0%)
Smoking status, n			
Never	669 (44.9%)	295 (36.8%)	374 (54.3%)
Previous	499 (33.5%)	311 (38.8%)	188 (27.3%)
Current	322 (21.6%)	195 (24.3%)	127 (18.4%)
Alcohol consumption, n			
Never	752 (50.4%)	375 (46.8%)	377 (54.7%)
<14 units/week	424 (28.4%)	235 (29.3%)	189 (27.4%)
>14 units/week	315 (21.1%)	192 (23.9%)	123 (17.9%)
Previous intravenous drug use, n	47 (3.2%)	39 (4.9%)	8 (1.2%)
Smear microscopy positive, n	1138 (76.3%)	625 (77.9%)	513 (74.5%)
Resistance classification, n			
Rifampicin mono-resistant	154 (10.3%)	80 (10.0%)	74 (10.7%)
Multidrug resistant	1110 (74.4%)	585 (72.9%)	525 (76.2%)
Pre-extensively drug resistant	221 (14.8%)	132 (16.5%)	89 (12.9%)
Multidrug and bedaquiline resistant	4 (0.3%)	3 (0.4%)	1 (0.1%)
Extensively drug resistance	2 (0.1%)	2 (0.2%)	0
<i>M. tuberculosis</i> lineage, n			
Lineage 1	69 (4.6%)	40 (5.0%)	29 (4.2%)
Lineage 2	1320 (88.5%)	704 (87.8%)	616 (89.4%)
Lineage 4	40 (2.7%)	25 (3.1%)	15 (2.2%)
Mixed or unknown	62 (4.2%)	33 (4.1%)	29 (4.2%)

Table 1: Participant characteristics stratified by TB history.

transmission clades, ranging in size from 2 to 37 cases. Large clades (>10 members) exhibited distinct spatial distributions across the city's urban core but were dispersed across multiple districts rather than concentrated within single geographic hotspots (Fig. 5). Geographic dispersion of transmission clades, measured as median pairwise commute time between cases, increased with time since network emergence (Spearman $\rho = -0.25$, $p = 0.002$). For comparison, a SNP-threshold analysis using 5 and 12 SNP cutoffs is presented in Supplementary Table S4 and Supplementary Figs. S2 and S3.

We identified 18 shared households comprising 38 participants (2.6% among 1486 cases with valid addresses). 17 households contained two cases and one household contained four. In only 8/18 (44.4%) households did all cases belong to the same transmission clade.

RR-TB transmission in Ho Chi Minh City was only weakly structured by host demographics or geography, with most genetic relatedness reflecting broad, city-wide transmission rather than assortative mixing within specific demographic or spatial subgroups. treeSeg detected no clades with distinct trait distributions for any binary demographic variable (sex, HIV status, diabetes, birthplace and drug use history). For age, treeSeg identified a single small clade of 13 isolates

(0.9% of the sample) with a lower mean age (21 versus 46 years), likely reflecting a localised transmission chain rather than systematic age-dependent structuring. There was no correlation between genetic and geographic distance (Mantel $r = 0.001$, $p = 0.94$), and pairwise commute times were similar for within-network and between-network pairs (median 14.0 versus 14.6 min). While individual transmission clades (≥ 10 members) differed in their demographic composition for several characteristics (Supplementary Fig. S4), this did not translate into assortative transmission at the population level where demographic and geographic similarity together explained little of the variance in genetic distance (MRM $R^2 = 0.62\%$).

Discussion

In this large, population-based genomic epidemiology study of RR-TB in Ho Chi Minh City, we show that transmission of already-resistant strains accounts for most of the current RR-TB burden, despite the fact that most patients with RR-TB have had a previous episode of TB. Nevertheless, a substantial and concerning proportion of RR-TB arises through *de novo* acquired resistance, despite long-standing programmatic efforts to prevent this.

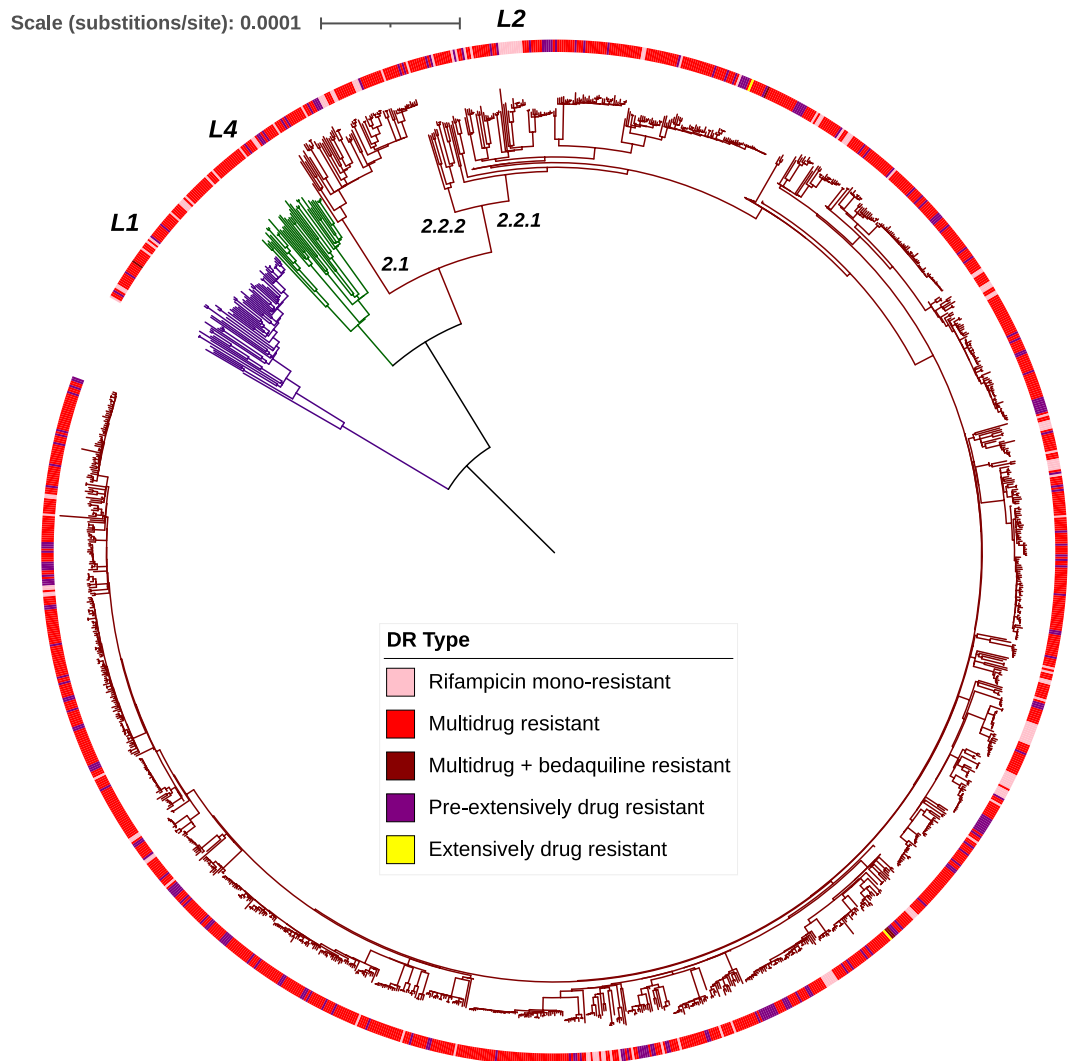


Fig. 3: Maximum-likelihood phylogeny for rifampicin-resistant *M. tuberculosis* isolates. *Excludes mixed lineage isolates. Outer ring represents drug resistance classification and branch colours represent lineage.

We further reconstructed the evolutionary history of RR-TB in Ho Chi Minh City and found that rifampicin resistance likely emerged not long after the introduction of the drug in the mid-1980s. We demonstrate that at least some of those early resistance events have resulted in networks that have persisted ever since, and that are now widespread across the city. Whilst most contemporary RR-TB in Ho Chi Minh City is the product of transmission, we find that the emergence of rifampicin-resistance has continued since the 1980s, despite the maturation of the NTP and the nationwide rollout of Xpert MTB/RIF by 2016.²⁹ Lineage 2 accounted for the majority of strains in our study, and was associated with higher odds of transmitted versus acquired resistance, consistent with prior evidence of its enhanced transmissibility and competitive dominance in Viet Nam.^{17,18}

The predominance of transmitted rifampicin resistance we observe aligns with findings from other high-burden settings. In a South African study, 70% of RR-TB isolates clustered within 5 SNPs and were attributed to transmission, while an analysis of more than 5000 strains from China estimated that 61% of RR-TB arose through transmission using similar methods.^{30,31} Estimates from SNP-based clustering may, however, be sensitive to sampling density and assumptions regarding the independence of resistance events. Incomplete sampling is likely to underestimate transmission while failure to account for homoplasy can inflate transmission estimates, particularly for rifampicin where a small number of resistance mutations occur frequently.^{22,32} Several smaller studies have therefore built upon SNP-based approaches by applying

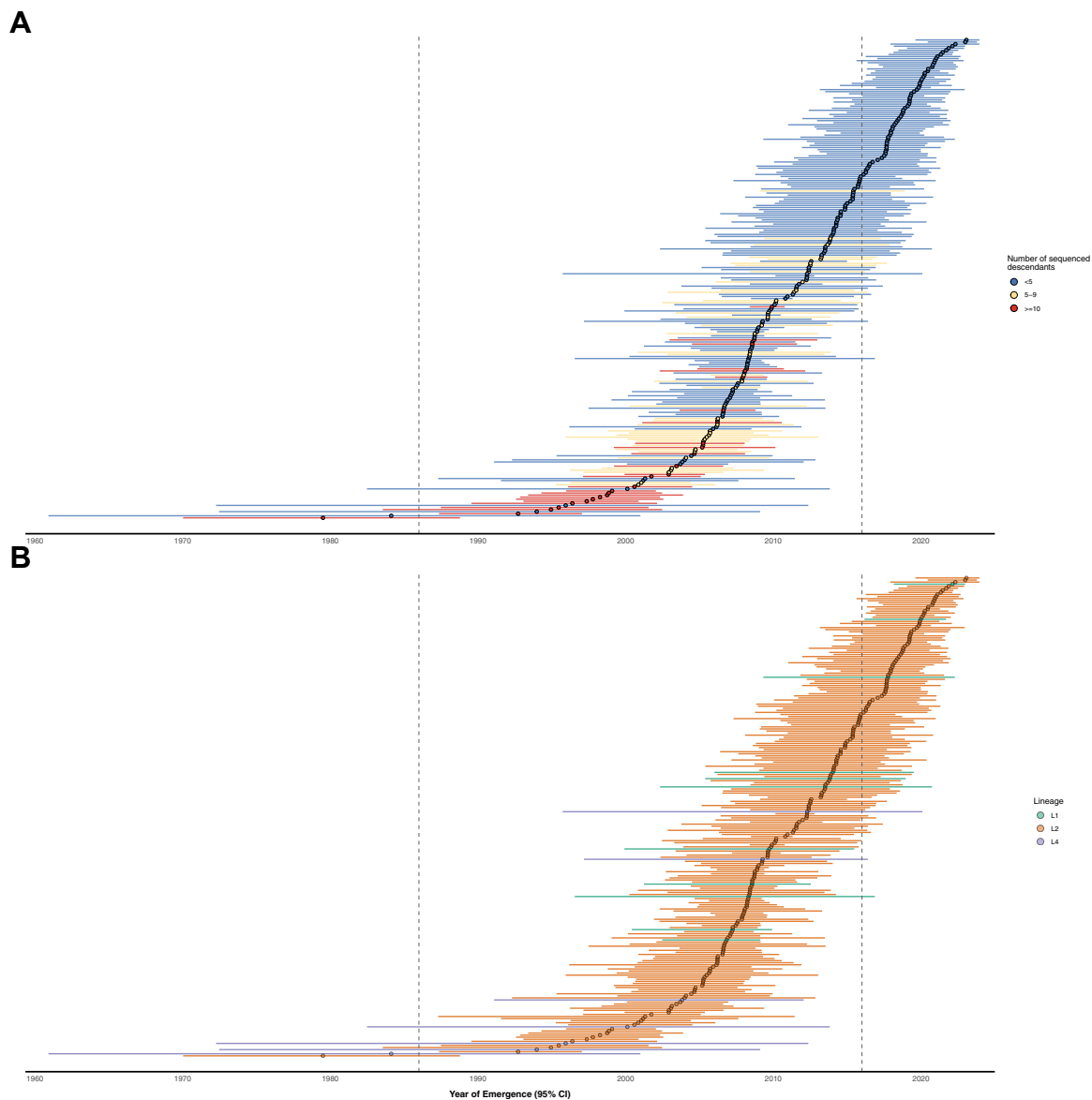


Fig. 4: Estimated dates of rifampicin-resistance emergence events for transmitted *M. tuberculosis* strains in Ho Chi Minh City. (A) Emergence events coloured by number of sequenced descendants. (B) Emergence events coloured by lineage. Dashed line at 1986 denotes the formation of the Viet Nam National Tuberculosis Programme and the nationwide rollout of rifampicin. Each point represents the median estimated date of emergence for a single resistance event; horizontal bars represent the corresponding 95% credible intervals. Dashed line at 2016 denotes nationwide roll of Xpert MTB/RIF assay.

phylogenetic methods, such as an analysis of 576 strains from Tibet which reported that 72% of MDR-TB was transmitted.³³ Our framework integrates additional sources of information including prior TB history, N and N –1 classification approaches and correction for sampling coverage which may improve the distinction of transmitted from acquired rifampicin resistance, particularly among individuals without previous TB. We explicitly characterise uncertainty arising from

sampling variability, phylogenetic inference, and classification boundaries, yielding estimates that more fully reflect the inherent limitations of inferring resistance origins from these data.

The predominance of transmitted resistance, even among people with previous TB, and the diffuse spatial structure of large RR-TB transmission clades may have important implications for TB control in urban Viet Nam. We previously showed, using routine case

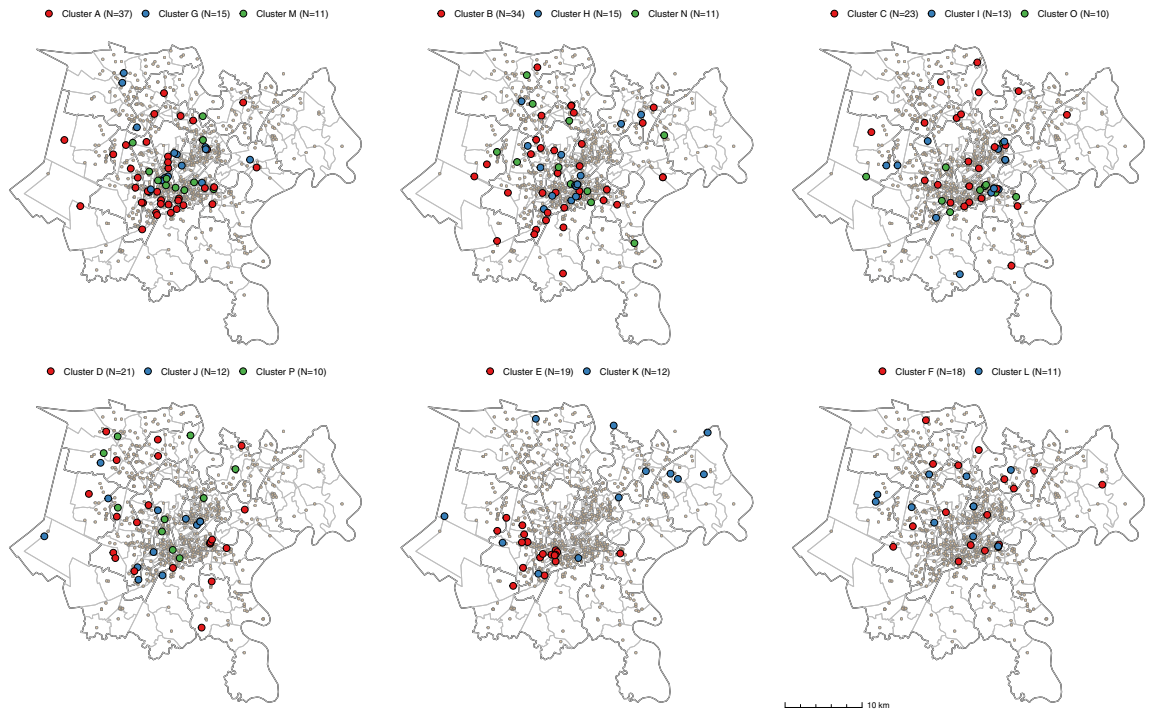


Fig. 5: Geographic distribution of rifampicin-resistant tuberculosis transmission clades including ≥ 10 cases in Ho Chi Minh City. Each coloured dot represents a participant's residential location, with colours distinguishing transmission clades within each panel. Tan coloured dots represent cases that are not members of any identified transmission clade. Transmission clades are grouped across multiple copies of the Ho Chi Minh City map to avoid overcrowding any single map. Maps exclude Cần Giờ and Củ Chi districts as no cases from large transmission clades located to these districts.

notifications, that Ho Chi Minh City's central districts were RR-TB incidence hotspots.¹⁰ The present genomic analysis provides finer resolution. Although some large transmission clades are anchored in high-incidence districts, most are geographically dispersed, spanning multiple districts rather than forming tightly localised hotspots. This pattern may suggest that RR-TB transmission in Ho Chi Minh City is not primarily driven by household spread or small neighbourhood clusters, but instead reflects broader city-wide transmission. Indeed, despite most residents living in multi-person households, we identified only 18 affected households, less than half of which showed evidence of infection with the same strains. This suggests that transmission may occur predominantly in congregate settings outside the home, such as workplaces or social venues, where ventilation and contact patterns favour airborne spread. Nevertheless, inferences about transmission locations based solely on individuals with symptomatic disease are inherently uncertain given the complex natural history of TB, in which transmission may occur long before clinical presentation and many newly infected individuals may never progress to symptomatic disease.³⁴ Ultimately, our findings support the prioritisation of broad, population-level interventions aimed at

interrupting transmission to achieve the greatest reductions in RR-TB burden. While community-wide screening is inherently resource-intensive, it has been shown to be highly effective in other Vietnamese settings and is likely to be cost-effective over the long term.^{3,35}

Between 13 and 28% of RR-TB in our study was attributable to acquired resistance, a substantial figure given the long-standing DOTS implementation and high first-line treatment success rates in our setting.³⁶ This finding suggests that factors beyond incomplete or inadequate treatment may be contributing to the emergence of rifampicin resistance. One plausible mechanism is unrecognised isoniazid resistance among people with rifampicin-susceptible TB. In these patients, standard first-line therapy provides only three effective drugs during the intensive phase and may functionally reduce to rifampicin monotherapy during the continuation phase. In a large systematic review, the treatment of isoniazid-resistant, rifampicin-susceptible TB with standard first-line regimens was associated with a 25-fold increased risk of acquiring multidrug resistance compared to isoniazid-susceptible disease.³⁷ Isoniazid mono-resistance may affect up to 20% of new TB cases in Viet Nam,³⁸ more than double

the global average,³⁹ potentially creating a substantial reservoir for the amplification of rifampicin resistance. Host factors may also contribute, such as NAT2 fast-acetylator genotypes that reduce isoniazid exposure. In a small Vietnamese cohort, one-quarter of individuals carried fast-acetylator genotypes.⁴⁰ Finally, we observed higher odds of acquired resistance amongst people with diabetes, consistent with prior reports, although underlying mechanisms remain uncertain.⁴¹

Our study has several limitations. Although we achieved substantially higher sampling coverage than most population-based TB studies from high-burden settings, interruptions during the COVID-19 pandemic and culture or sequencing failures meant that one-third of RR-TB cases were unsampled. While we attempted to correct for this using SIMEX, the correction assumed unsampled cases were missing at random. Missing data in surveillance datasets are unlikely to be completely random; however, participants without a sequenced isolate did not differ from those with sequencing data on measured characteristics apart from smear status, providing some reassurance. Nevertheless, systematic differences in which patients underwent culture and sequencing could have introduced residual bias in the SIMEX-corrected estimates. Our classification of transmitted versus acquired resistance integrated phylogenetic evidence and clinical history, but there may still have been residual misclassification as treatment histories may have been incomplete, and infection outside Ho Chi Minh City could not be fully excluded. Birthplace, used as a proxy in our sensitivity analysis, is an imperfect measure of where infection was acquired in the absence of detailed migration and travel histories. Enrolment was restricted to cases detected through programmatic Xpert testing and initiating treatment at Phạm Ngọc Thạch Hospital excluding asymptomatic individuals and those who died before diagnosis. Finally, although our findings are context-specific and may not generalise, the analytic framework we describe should be applicable to similar investigations elsewhere.

RR-TB in Ho Chi Minh City is driven predominantly by ongoing transmission but a substantial contribution from *de novo* acquired resistance signals persistent vulnerabilities in current diagnostic and treatment pathways. Intensified efforts to interrupt transmission through earlier case detection and rapid treatment initiation will need to be coupled with further investigation into the precise drivers of *de novo* acquired drug-resistance.

Contributors

TMW conceptualised the study with input from GT and NTTT. Data generation and curation was performed by RS, NHH, PTP, LKL, KL, NNH, NLQ, DDAT, NTLH, TLNT, TTBT, VTNH, DTMH, NPH, NHT, TTKQ, NHL, VD, SN and LHV. Formal analysis and visualisation was performed by RS with input from TMW, JAW and MC. Funding was acquired by TMW. RS, JAW, MC and TMW contributed to methodology. Project administration was conducted by NHH, PTP, RS,

TMW, DTMH, NHL, NPH, NHT and TTKQ. Study resources were provided by TMW, GT, SN and DC. TMW, JAW and MC provided supervision. RS wrote the original draft with all authors reviewing and editing. RS and TMW verified the underlying data. All authors had full access to all the data in the study and had the final responsibility for the decision to submit for publication.

Data sharing statement

All newly generated sequences will be made available in the European Nucleotide Archive under study accession PRJEB107150 on final publication of the study. Accession numbers for previously published sequences are provided in the [Supplementary files](#). Code for the SIMEX analysis is available at <https://github.com/RuanSpies21/TB-SIMEX>.

Editor note

The Lancet Group takes a neutral position with respect to territorial claims in published maps and institutional affiliations.

Declaration of interests

We declare no competing interests.

Acknowledgements

RS is supported by the Rhodes Trust. This research was funded in whole, or in part, by the Wellcome Trust (214560/Z/18/Z). TMW is a Wellcome Trust Clinical Career Development Fellow (214560/Z/18/Z).

During the preparation of this work the author(s) used Claude AI in order to improve language and readability. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Appendix A. Supplementary data

Supplementary data related to this article can be found at <https://doi.org/10.1016/j.lanwpc.2026.101906>.

References

- World Health Organization. *Global Tuberculosis Report 2025*; 2025. Available from: <https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2025>. Accessed November 13, 2025.
- Van LH, Phu PT, Vinh DN, et al. Risk factors for poor treatment outcomes of 2266 multidrug-resistant tuberculosis cases in Ho Chi Minh City: a retrospective study. *BMC Infect Dis*. 2020;20:164.
- Marks GB, Nguyen NV, Nguyen PT, et al. Community-wide screening for tuberculosis in a high-prevalence setting. *N Engl J Med*. 2019;381:1347–1357.
- Mahomed S, Padayatchi N, Singh J, Naidoo K. Precision medicine in resistant tuberculosis: treat the correct patient, at the correct time, with the correct drug. *J Infect*. 2019;78:261–268.
- Keshavjee S, Farmer PE. Tuberculosis, drug resistance, and the history of modern medicine. *N Engl J Med*. 2012;367:931–936.
- Blomberg B, Spinaci S, Fourie B, Laing R. The rationale for recommending fixed-dose combination tablets for treatment of tuberculosis. *Bull World Health Organ*. 2001;79:61–68.
- Comas I, Borrell S, Roetzer A, et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nat Genet*. 2012;44:106–110.
- Kendall EA, Fofana MO, Dowdy DW. Burden of transmitted multidrug resistance in epidemics of tuberculosis: a transmission modelling analysis. *Lancet Respir Med*. 2015;3:963–972.
- Huong NT. *Tuberculosis control in Vietnam: does DOTS do it?* University of Amsterdam; 2007. Available from: <https://dare.uva.nl/id/4c5ec574-2c76-4601-a073-96a9167ef779>. Accessed October 22, 2025.
- Spies R, Hong HN, Trieu PP, et al. Spatial analysis of drug-susceptible and multidrug-resistant cases of tuberculosis, Ho Chi Minh City, Vietnam, 2020–2023. *Emerg Infect Dis*. 2024;30. <https://doi.org/10.3201/eid3003.231309>.
- Hoa NB, Khanh PH, Chinh NV, Hennig CM. Prescription patterns and treatment outcomes of MDR-TB patients treated within and outside the National Tuberculosis Programme in Pham Ngoc

- Thach hospital, Viet Nam. *Trop Med Int Health*. 2014;19:1076–1081.
- 12 Le Quang N, Dang Anh Thu D, Pham Tien Trieu L, et al. A modified decontamination and storage method for sputum from patients with tuberculosis. *Wellcome Open Res*. 2023;8:166.
 - 13 Votintseva AA, Pankhurst LJ, Anson LW, et al. Mycobacterial DNA extraction for whole-genome sequencing from early positive liquid (MGIT) cultures. *J Clin Microbiol*. 2015;53:1137–1143.
 - 14 Vijay S, Vinh DN, Hai HT, et al. Influence of stress and antibiotic resistance on cell-length distribution in Mycobacterium tuberculosis clinical isolates. *Front Microbiol*. 2017;8:2296.
 - 15 Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R. Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One*. 2015;10:e0128036.
 - 16 Westhead J, Baker CS, Brouard M, et al. Characterizing the performance of an antibiotic resistance prediction tool, gnomonius, using a diverse test set of 2,663 Mycobacterium tuberculosis samples. *Microb Genom*. 2025;11:001592.
 - 17 Holt KE, McAdam P, Thai PVK, et al. Frequent transmission of the Mycobacterium tuberculosis Beijing lineage and positive selection for the EsxW Beijing variant in Vietnam. *Nat Genet*. 2018;50:849–856.
 - 18 Silcocks M, Chang X, Thuong Thuong NT, et al. Evolution and transmission of antibiotic resistance is driven by Beijing lineage Mycobacterium tuberculosis in Vietnam. *Microbiol Spectr*. 2023;11:e02562-23.
 - 19 The CRYPTIC Consortium. A data compendium associating the genomes of 12,289 Mycobacterium tuberculosis isolates with quantitative resistance phenotypes to 13 antibiotics. *PLoS Biol*. 2022;20:e3001721.
 - 20 Wong TKF, Ly-Trong N, Ren H, et al. IQ-TREE 3: phylogenomic inference software using complex evolutionary models. *EcoEvoRxiv*; 2025. Available from: <https://ecoevorxiv.org/repository/view/8916/>. Accessed July 17, 2025.
 - 21 Edwards DJ, Duchene S, Pope B, Holt KE. SNPPar: identifying convergent evolution and other homoplasies from microbial whole-genome alignments. *Microb Genom*. 2021;7:000694.
 - 22 World Health Organization. *Catalogue of mutations in Mycobacterium Tuberculosis complex and their Association with drug resistance*. 2nd ed.; 2023. Available from: <https://iris.who.int/bitstream/handle/10665/374061/9789240082410-eng.pdf?sequence=1>. Accessed April 19, 2024.
 - 23 Cook JR, Stefanski LA. Simulation-extrapolation estimation in parametric measurement error models. *J Am Stat Assoc*. 1994;89:1314–1328.
 - 24 Didelot X, Croucher NJ, Bentley SD, Harris SR, Wilson DJ. Bayesian inference of ancestral dates on bacterial phylogenetic trees. *Nucleic Acids Res*. 2018;46:e134.
 - 25 Menardo F, Duchêne S, Brites D, Gagneux S. The molecular clock of Mycobacterium tuberculosis. *PLoS Pathog*. 2019;15:e1008067.
 - 26 Cambon J, Hernangómez D, Belanger C, Possenriede D. tidygeocoder: An R package for geocoding. *J Open Source Softw*. 2021;6:3544.
 - 27 Giraud T. osrm: Interface between R and the OpenStreetMap-Based routing service OSRM. *J Open Source Softw*. 2022;7:4574.
 - 28 Behr M, Ansari MA, Munk A, Holmes C. Testing for dependence on tree structures. *Proc Natl Acad Sci USA*. 2020;117:9787–9792.
 - 29 Gumma V, DeGruy K, Bennett D, et al. Impact of external quality assurance on the quality of Xpert MTB/RIF testing in Viet Nam. *J Clin Microbiol*. 2019;57:e01669-18.
 - 30 Cox H, Salaam-Dreyer Z, Goig GA, et al. Potential contribution of HIV during first-line tuberculosis treatment to subsequent rifampicin-monoresistant tuberculosis and acquired tuberculosis drug resistance in South Africa: a retrospective molecular epidemiology study. *Lancet Microbe*. 2021;2:e584–e593.
 - 31 Liu D, Huang F, Li Y, et al. Transmission characteristics in Tuberculosis by WGS: nationwide cross-sectional surveillance in China. *Emerg Microb Infect*. 2024;13:2348505.
 - 32 Menardo F. Understanding drivers of phylogenetic clustering and terminal branch lengths distribution in epidemics of Mycobacterium tuberculosis. *eLife*. 2022;11:e76780.
 - 33 Jiang Q, Liu HC, Liu QY, et al. The evolution and transmission dynamics of multidrug-resistant tuberculosis in an isolated high-plateau population of Tibet, China. *Microbiol Spectr*. 2023;11:e03991-22.
 - 34 Coussens AK, Zaidi SMA, Allwood BW, et al. Classification of early tuberculosis states to guide research for improved care and prevention: an international Delphi consensus exercise. *Lancet Respir Med*. 2024;12:484–498.
 - 35 Schwalb A, Horton KC, Emery JC, et al. Potential impact, costs, and benefits of population-wide screening interventions for tuberculosis in Viet Nam: a mathematical modelling study. *PLoS Glob Publ Health*. 2025;5:e0005050.
 - 36 World Health Organization. Tuberculosis profile: Viet Nam. Available from: https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&entity_type=%22country%22&iso2=%22VN%22&lan=%22EN%22. Accessed December 24, 2025.
 - 37 Gegia M, Winters N, Benedetti A, van Sooling D, Menzies D. Treatment of isoniazid-resistant tuberculosis with first-line drugs: a systematic review and meta-analysis. *Lancet Infect Dis*. 2017;17:223–234.
 - 38 Callum J, Nguyen PT, Martinez E, et al. Prevalence and genetic basis of first-line drug resistance of Mycobacterium tuberculosis in Ca Mau, Vietnam. *ERJ Open Res*. 2022;8:00122–2022.
 - 39 Dean AS, Zignol M, Cabibbe AM, et al. Prevalence and genetic profiles of isoniazid resistance in tuberculosis patients: a multicountry analysis of cross-sectional data. *PLoS Med*. 2020;17:e1003008.
 - 40 Le TA, Pham NHT, Dinh LD, Le LT, Vu TT. NAT2 gene polymorphisms and Plasma Isoniazid Concentration in Vietnamese tuberculosis patients. *Walailak J Sci Tech*. 2021;18.
 - 41 Rehman AU, Khattak M, Mushtaq U, et al. The impact of diabetes mellitus on the emergence of multi-drug resistant tuberculosis and treatment failure in TB-diabetes comorbid patients: a systematic review and meta-analysis. *Front Public Health*. 2023;11:1244450.