Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Global Geographical Distribution of *Mycobacterium Tuberculosis* Complex and the Association Between GHSI and Drug-Resistance

Bilin Tao
Nanjing Medical University
Jizhou Wu
Nanjing Medical University
Beibei Qiu
Nanjing Medical University
Zhongqi Li
Nanjing Medical University
Mengyao Zhan
Nanjing Medical University
Zhuchao Wu
Nanjing Medical University
Jianming Wang (jmwang@njmu.edu.cn)
Nanjing Medical University

Research Article

Keywords: tuberculosis, whole-genome sequencing, mycobacterium tuberculosis complex lineage, Global Health Security index, interaction

Posted Date: October 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-927338/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background

Mycobacterium tuberculosis complex (MTBC) is the causative agent of tuberculosis (TB). This study aims to map the global distribution of MTBC lineage and to explore the correlation between Global Health Security Index (GHSI) and drug resistance.

Methods

We mapped the global geographic distribution of MTBC lineages using the whole-genome sequencing (WGS) and verified data from the TBProfiler. The hierarchical structure was visualized in different continents and sublineages. We also performed two-dimensional twisted surface and interaction plots to explore the interactions.

Results

Lineage 4 was widely distributed globally, while lineage 2 had the highest risk of developing DR-TB. We observed an interaction between GHSI and GDP on the prevalence of multidrug-resistant TB. In countries with ln(GDP per capita) ≥2.35, there was a negative association between GHSI and drug resistance.

Conclusions

There was a significantly different geographic distribution pattern of MTBC lineages in the world. The GHSI was related to the drug resistance of TB and was affected by the social-economic level.

Introduction

Tuberculosis (TB) is a contagious disease that is collectively a top infectious disease killer from a single infectious agent globally, causing an estimated 10.0 million new TB cases annually[1]. Tackling the transmission and drug resistance burden of Mycobacterium tuberculosis (*M.tb*) requires concerted global effort in the diagnosis, treatment, prevention, and surveillance. Molecular characterization of resistance from the *M.tb* circular genome (small size 4.4 Mb) offers an alternative to the traditional culture-based method, providing a highly accurate, time-saving diagnosis, drug susceptibility profiling, and a better understanding of *M.tb* transmission dynamics[2, 3].

The whole-genome sequencing (WGS) approach applies a DNA sequencing platform to reconstruct the complete DNA sequence of an organismal genome[4]. It offers an attractive option as it simultaneously examines all loci and provides information regarding both small and large changes in the genome[5], thus guiding all TB control components, including diagnosis, treatment, source investigation, and surveillance[6, 7].

The Global Health Security Index (GHSI) is a comprehensive assessment and benchmarking of health security and related capabilities across 195 countries. The GHSI is scored by examining published and publicly available evidence across six categories (prevention, detection, rapid response, health system, commitments to international norms, and risk environment) to encourage nations to document and publicize preparations. The GHSI is a broad assessment that points explicitly to very high-consequence events, robust health systems, commitments to international norms, and the risk environment[8, 9]. Whether the GHSI is related to the risk of drug-resistant TB (DR-TB) within different phylogenetic lineages of *M.tb* remains unclear.

In this study, we performed a descriptive and comparative analysis, aiming to describe the worldwide geographical distribution of MTBC and their lineages (sublineages) and analyze the risk of drug resistance by sublineages. Also, we made the first attempt to explore the GHSI in association with the epidemic situation of multidrug resistance (MDR).

Methods

Data collection

We extracted the data of MTBC from the TBProfiler webserver (https://tbdr.lshtm.ac.uk/sra) on February 25, 2021. Countries with *M.tb* isolates <10 were not included in the final analysis. TBProfiler webserver is a free academic platform that allows researchers to upload raw WGS data of MTBC to retrieve a concise report on lineage and drug resistance[5, 10]. To date, it has been expanded to incorporate 178 new markers across 16 anti-TB drugs [11]. The GHSI was collected from the Global Health Security Index website (https://www.ghsindex.org). We retrieved sociological information of each country (https://ourworldindata.org/), including continent, population and population density, median age, life expectancy, income levels, and gross domestic product (GDP) per capita. Data on DR-TB were extracted from the “Global Tuberculosis Report 2020” (https://www.who.int/teams/global-tuberculosis-programme/data)[1]. We can confirm that all methods were performed in accordance with the relevant guidelines and regulations.

Definition of drug resistance

MDR: strains resistant to isoniazid and rifampin; pre-multidrug resistance (Pre-MDR): Strains resistant to either isoniazid, rifampin, ethambutol, or streptomycin, or any combination of these drugs except isoniazid and rifampin[12]; extensive drug resistance (XDR): on the base of MDR, the additional resistance to the fluoroquinolones and injectable medications (amikacin, kanamycin, and capreomycin); pre-extensively drug resistance (Pre-XDR): resistance to isoniazid and rifampicin and additional resistance to either any second-line injectable drug (amikacin, kanamycin, and capreomycin) or any fluoroquinolone-resistant[13].

Data Analysis

Categorized variables were expressed as percentages and analyzed by Fisher's exact probability method. Continuous skewed distribution data were described as the median (interquartile range, IQR), and differences between groups were analyzed using the Kruskal-Wallis rank test. We used the R package of "maps" and "ggplot2" to draw a world geographic map. The hierarchical structure was visualized using R packages of "ggtree", "treeio", along with "ggplot2". The GHSI and GDP per capita were logarithmic transformed before analysis. The rate of MDR was estimated by the exact binomial test and described as mean (95% confidence interval [CI]), and its error bar plot was performed using the R package of "ggplot2". We also applied two-dimensional twisted surface and interaction plots[14] to explore the interaction between GHSI and GDP per capita using an R code script (http://www3.i-med.ac.at/genepi/). The correlation plots between MDR/RR and GHSI were analyzed by winsorization transformation with an R package of "ggstatsplot". The significant level was set at a two-tailed *P*-value of 0.05. We performed analyses using R software for Windows version 4.0.3 (https://www.r-project.org/).

Results

Sociological characteristics

A total of 60 countries with *M.tb* isolates >10 in the TBprofiler database was selected for analysis (Supplementary Figure 1). According to their geographical distribution, the countries were assigned to 6 continents. The sociological characteristics were described in Table 1. Northern America had the highest median population number (184.37 million, IQR: 111.06-257.69 million), GHSI (79.40, IQR: 77.35-81.45), GDP per capita (49.12 US$1000, IQR: 46.57-51.67 US$1000), and relatively longer median life expectancy (80.65 years, IQR: 79.75-81.54 years). Africa had the lowest level of median age (19.60 years, IQR: 18.10-21.50 years), life expectancy (64.07 years, IQR: 60.19-66.60 years), GHSI (35.50, IQR: 29.00-40.60), and GDP per capita (2.99 US$1000, IQR: 1.73-5.34 US$1000). Asia had the largest median population density with an estimated 230.00 per square kilometer (IQR: 120.38-376.51 per square kilometer).

Table 1
Characteristics of tuberculosis in sixty countries stratified by the six Continents in the World.

Characteristics	Continents						*P*-value
	Africa	Asia	Europe	Northern America	Oceania	South America	
Isolates (n=29126)	**7648 (26.3)**	**7047 (24.2)**	**10661 (36.6)**	**1941 (6.7)**	**224 (0.8)**	**1605 (5.4)**	
Lineages (%)*							<0.001
Lineage 1	384 (5.0)	1427 (20.2)	722 (6.8)	288 (14.8)	5 (2.2)	9 (0.6)	
Lineage 2	1439 (18.8)	3837 (54.4)	1875 (17.6)	381 (19.6)	158 (70.5)	94 (5.9)	
Lineage 3	396 (5.2)	583 (8.3)	1950 (18.3)	135 (7.0)	5 (2.2)	3 (0.2)	
Lineage 4	5068 (66.3)	1198 (17.0)	5879 (55.1)	1136 (58.5)	55 (24.6)	1337 (83.3)	
Lineage 5	207 (2.7)	0 (0.0)	31 (0.3)	1 (0.1)	0 (0.0)	0 (0.0)	
Lineage 6	95 (1.2)	0 (0.0)	23 (0.2)	0 (0.0)	0 (0.0)	0 (0.0)	
Lineage 7	47 (0.6)	0 (0.0)	1 (0.0)	0 (0.0)	1 (0.4)	0 (0.0)	
Lineage 9	1 (0.0)	0 (0.0)	1 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
M.bovis	8 (0.1)	2 (0.0)	149 (1.4)	0 (0.0)	0 (0.0)	158 (9.8)	
M.caprae	3 (0.0)	0 (0.0)	3 (0.0)	0 (0.0)	0 (0.0)	4 (0.2)	
M.orygis	0 (0.0)	0 (0.0)	27 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	
Drug Resistance types (%)*							<0.001
Sensitive	5469 (71.5)	3408 (48.4)	6964 (65.3)	1682 (86.7)	90 (40.2)	191 (11.9)	
MDR	992 (13.0)	1204 (17.1)	1109 (10.4)	25 (1.3)	93 (41.5)	542 (33.8)	
Pre-MDR	549 (7.2)	963 (13.7)	704 (6.6)	144 (7.4)	28 (12.5)	244 (15.2)	
XDR	249 (3.3)	212 (3.0)	429 (4.0)	2 (0.1)	5 (2.2)	80 (5.0)	
Pre-XDR	237 (3.1)	740 (10.5)	761 (7.1)	6 (0.3)	8 (3.6)	371 (23.1)	

Characteristics	Continents						*P*-value
	Africa	Asia	Europe	Northern America	Oceania	South America	
Others	152 (2.0)	520 (7.4)	694 (6.5)	82 (4.2)	0 (0.0)	177 (11.0)	
Population(million)**	26.38 (11.82-53.77)	76.90 (34.48-137.41)	10.10 (5.79-46.75)	184.37 (111.06-257.69)	17.22 (13.09-21.36)	48.04 (36.03-109.42)	0.009
Population density(kilometer^{-2})**	79.49 (42.73-104.96)	230.00 (120.38-376.51)	93.11 (46.86-205.86)	19.82 (11.93-27.72)	10.71 (6.96-14.47)	55.33 (29.90-66.44)	0.009
Median age(year)**	19.60 (18.10-21.50)	31.50 (27.45-34.12)	42.30 (40.80-43.20)	39.85 (39.07-40.62)	30.25 (26.43-34.07)	29.30 (29.15-31.25)	<0.001
Life expectancy(year)**	64.07 (60.19-66.60)	75.26 (71.12-76.97)	81.63 (76.05-82.40)	80.65 (79.75-81.54)	73.97 (69.24-78.70)	75.86 (75.05-76.72)	<0.001
GHSI (points)**	35.50 (29.00-40.60)	47.05 (37.15-49.15)	64.60 (52.30-68.70)	79.40 (77.35-81.45)	51.65 (39.73-63.58)	53.40 (45.45-58.35)	<0.001
GDP per capita (US$1000)**	2.99 (1.73-5.34)	10.48 (6.23-16.98)	39.75 (24.77-46.68)	49.12 (46.57-51.67)	24.24 (14.03-34.44)	15.30 (12.49-17.34)	<0.001

Geographical distribution of M.tb **isolates data**

We collected 29126 individual isolates from included 60 countries. Among them, 18 countries were ranked as the top 30 high TB-burden countries[15]. Countries with >1000 strains were South Africa (4209), United Kingdom (4052), Netherlands (2430), Vietnam (1981), China (1911), Canada (1814), Malawi (1747), Thailand (1483), and Russia (1074), respectively. South Africa in Africa, Vietnam in Asia, the United Kingdom in Europe, Canada in Northern America, Austria in Oceania, Peru in South America had the largest number of isolates strains in separated corresponding continents (Figure 1).

Global distribution of isolates lineages

Most *M.tb* isolates belonged to the Euro-American lineage 4. A striking feature of lineage 4 was identified as the widespread global distribution lineage, and it was the majority lineage in 4 of the 6 continents: Africa (66.3%), Europe (55.1%), Northern America (58.5%), and South America (83.3%). Globally, lineage 4 accounted for more than half of all lineages (50.38%). Our map also showed the widespread distribution of East Asian lineage 2, which was identified in 2 of the 6 continents and was the majority lineage in Asia (54.4%) and Oceania (70.5%). Lineage 2 accounted for more than one-quarter of all lineages (26.73%)

(Table 1 and Figure 2). South America had the highest proportion of *Mycobacterium Bovis* (*M.bovis*) with an estimated 9.8%, followed by Europe (1.4%).

Lineage 1 to lineage 6 accounted for 99.03% of all isolates. Among them, the largest proportion of each lineage were the sublineages of lineage 1.2.1.2.1 (635, 22.40%), lineage 2.2.1 (6458, 82.97%), lineage 3 (3072, 69.21%), lineage 4.8 (1975, 13.46%), lineage 5.1.1 (23.9, 38.08%), and lineage 6.3.1 (62, 52.54%), respectively (Figure 2). Our geographical map of major lineages also demonstrated that lineage 1 was more prevalent in South Asia (India and the Philippines). Lineage 2 was identified as the major lineage in Eastern Asia (China, South Korea, and the eastern part of Russia) and Oceania (Austria). Lineage 4 was identified as the major lineage in Europe, North American, South American, and Africa. We also found that lineage 3 was the major lineage in Norway (Europe), Saudi Arabia (Western Asia), and Iran (Southern Asia). Lineage 5 and lineage 7 were the major lineage only in Ghana (Africa) and Ethiopia (Africa), respectively (Figure 3).

Global drug-resistance and resistant lineages

The proportion of strains resistant to at least one anti-TB drug was 28.5% in Africa, 34.7% in Europe, 13.3% in Northern America, 88.1% in South America, 51.6% in Asia, and 59.8% in Oceania. The MDR and XDR strains accounted for 13.0% in Africa, 17.1% in Asia, 10.4% in Europe, 1.3% in Northern America, 41.5% in Oceania, and 33.8% in South America (Table 1).

Due to the limited samples of rare isolates of lineage 7, *M.bovis*, *Mycobacterium caprae* (*M.caprae*), and *Mycobacterium orygis* (*M.orygis*), we only selected the isolates of lineage 1 - lineage 6 for further drug resistance analysis. The lineage 2 had the highest prevalence of DR-TB (0.564, 95% CI: 0.553-0.576), MDR-TB (0.388, 95% CI: 0.378-0.399), and XDR-TB (0.072, 95% CI: 0.066-0.078). The lineage 4 had the second-highest prevalence of DR-TB (0.326, 95% CI: 0.319-0.334) which was similar to that of lineage 3 in DR-TB (0.326, 95% CI: 0.310-0.343), MDR-TB (0.219, 95% CI: 0.212-0.225), and XDR-TB (0.025, 95% CI: 0.022-0.027). On the contrary, lineage 6 had the lowest risk of developing DR-TB, MDR-TB, and XDR-TB with the estimated proportion of 0.119 (95% CI: 0.066-0.191), 0.042 (95% CI: 0.014-0.096), and 0, respectively. In addition, lineage 3 had the second-lowest risk of being resistant strains with the prevalence of MDR-TB (0.172, 95% CI: 0.158-0.185). Lineage 5 was also identified as isolates with the second-lowest prevalence of DR-TB (0.318, 95% CI: 0.259-0.381).

Interaction between GHSI and GDP

An overview of the interaction between ln(GHSI) and ln(GDP per capita) can be given by the twisted surface in Figure 4A. The surface was twisted, indicating that the direction of the ln(GDP per capita) effect changed with the varying values of ln(GHSI). The highest prevalence of DR-TB can be found at consecutive low values of both ln(GHSI) and ln(GDP per capita), and the ln(GHSI) did have a positive impact on the prevalence of DR-TB for low values of ln(GDP per capita). With an increase of ln(GDP per capita), the ln(GHSI) effect was attenuated. Furthermore, for higher values of ln(GDP per capita), the ln(GHSI) was negatively associated with the percentage of isolates with DR-TB.

The interaction plots in Figure 4B demonstrated that the effect estimator of ln(GHSI) on the prevalence of DR-TB was significantly lower than 0 for exceeding the threshold value of ln(GDP per capita) at 2.35. The non-significance region could also be read from this plot at the range of ln(GDP per capita) <2.35. In brief, the prevalence of DR-TB had a significantly negative association with ln(GHSI) in countries with ln(GDP per capita) ≥2.35.

Association between GHSI and DR-TB

Overall analysis on 60 countries suggested no significant relationship between the prevalence of DR-TB and ln(GHSI) (r =-0.24, 95% CI: -0.46, 0.02) (Figure 5). Subgroup analysis found that 24 countries with ln(GDP per capita) <2.35 (low income: 8, 33.33%; lower middle income: 15, 62.50%; upper middle income: 1, 4.17%) had no significant association (r =-0.13, 95% CI: -0.51, 0.28) between ln(GHSI) and DR-TB. A significantly negative association (r =-0.58, 95% CI: -0.76, -0.31) was observed in 36 countries with ln(GDP per capita) ≥2.35 (lower middle income: 1, 2.78%; upper middle income: 14, 38.89; high income: 21, 58.33%) (Figure 5) .

External verification of association between GHSI and DR-TB

To verify the representativeness, extrapolation, and extensive suitability of above conclusion, we utilized external data extracted from the top 30 high TB-burden countries[15]. No overall association was found between ln(GHSI) and the prevalence of MDR/RR in these countries among new cases (r =-0.03, 95% CI: -0.38, 0.34) or previous treated ones (r =-0.06, 95% CI: -0.41, 0.31). Subgroup analysis found that the coefficient of the prevalence of MDR/RR on ln(GHSI) turned from non-significant (ln[GDP per capita] <2.35) to significant negative (ln[GDP per capita] ≥2.35). For new cases, there was no significant association (r =0.07, 95% CI: -0.37, 0.49) in countries whose ln(GDP per capita) <2.35, but we observed a significantly negative association (r =0.70, 95% CI: -0.93, -0.08) in countries with ln(GDP per capita) ≥2.35. Similar results were also found in the previous treated cases in countries with ln(GDP per capita) <2.35 (r =0.06, 95% CI: -0.38, 0.48) and ≥2.35 (r =-0.71, 95% CI: -0.93, 0.08), respectively (Figure 6).

Discussion

The Euro-American lineage 4 was the primary type in the world, while lineage 2 had the highest risk of DR-TB (56.4%), MDR-TB (38.8%), and XDR-TB (7.2%). DR-TB was significantly associated with GHSI in countries with a higher level of GDP per capita. To our knowledge, this is the first study to explore the association between DR-TB and GHSI by considering local social-economic status.

The MTBC lineages varied with geographic areas and were endemic in specific sites by adapting to local human populations[16]. They are associated with the emergence of drug resistance, transmissibility, virulence, host response, vaccine efficacy, disease site, and severity[17–19]. Matthias Merker et al. reported that lineage 2 strains increased capacity to acquire drug resistance[20]. This may be due to the

hypermutability or compensatory mutations, which can reduce the fitness cost of resistance-conferring mutations[21, 22].

A meta-analysis including 206 studies, representing over 200,000 bacterial isolates in 85 countries, was regarded as the most comprehensive dataset on MTBC lineages but did not include WGS data[23]. Gary Napier et al.[24] used WGS data to map the global geographical distribution of the lineages, but they did not describe the global geographical distribution of DR and its relationship with different lineages. Our results indicate that MTBC strains (lineage 2, lineage 3, and lineage 4) that evolved more recently in human history tend to be more widely distributed around the world, where Africa is the only region that contains all 7 species of MTBC strains, which is inconsistent with the previous study[25]. Hershberg and Wirth pointed out that 40,000 to 70,000 years ago[26, 27], TB went out of Africa to expand and evolve, and so different TB types appeared.

Despite substantial increases since 2000, funding for TB is still far short of global financing targets, and out-of-pocket spending remained high in resource-constrained countries, posing a barrier to the care and treatment adherence of patients[28]. The GHSI has been widely applied to identify areas of weakness and opportunities to collaborate across sectors, strengthen health systems, and achieve public health goals [9, 29]. In the current study, we explored the correlation between GHSI and DR-TB and observed an interaction between GHSI and GDP per capita, and subsequently uncovered a significantly negative association between GHSI and DR in countries with a higher social-economic level. Low-income countries were inadequately resourced in the health system, which induced a lower level of GHSI.

A significant strength of this study is that we applied WGS data of *M.tb* from 60 countries on six continents, which can guarantee good global representativeness. In addition, data on results of WGS were reliable for identifying comprehensive drug resistance types and lineages with high sensitivity and specificity. Our study, to our knowledge, is the first to assess the relationship between drug resistance and GHSI and found an interaction between GHSI by considering the local social-economic level. Furthermore, we also applied an external dataset to verify our findings, and the results were consistent.

Several limitations should not be ignored. First, although we have made a hierarchical structure tree to map the lineage of isolates in each continent and the composition of sublineages in each lineage, a phylogenetic tree was not available due to lacking of WGS original fastq or fasta data for pairwise blasting to calculate the single nucleotide polymorphism (SNP) distance of each isolate[30]. Second, we observed that the major lineage in the United States was lineage 2, not the previously reported lineage 4. One possible reason is that the sample size in the United States was relatively small. Another reason is that the United States is a diverse multi-ethnic country, and strain types are specifically adapted to different human populations[16, 17].

In conclusion, a striking feature of global MTBC lineages was that lineage 4 was widely distributed globally, followed by East Asian lineage 2, which was related to an increased risk of MDR and XDR-TB. The GHSI was associated with the drug resistance of TB and affected by the social-economic level. More funding is needed to achieve the “End TB” target, especially in low-income countries.

Declarations

Supplementary Data

The datasets and R codes used and/or analyzed during this study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate

This study was approved by the ethics committee of Nanjing Medical University. We applied the TBProfier to analyze *M.tb* WGS data to predict lineage and drug resistance. Data were accessed from the public database. This study didn't involve any individual information. So there was no requirement for informed consent.

Consent for publication

Not applicable.

Availability of data and materials

The datasets supporting the conclusions of this study are publicly available from TB Profiler (https://tbdr.lshtm.ac.uk/) upon request.

Competing interests

The authors declare no conflict of interest.

Funding

This study was funded by the National Natural Science Foundation of China (81973103), National Key R&D Program of China (2017YFC0907000), Key Project of Philosophy and Social Science Research in Colleges and Universities in Jiangsu Province (2020SJZDA096), and Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD). The funding agencies had no role in the study design, data collection, analysis, decision to publish, or preparation of the manuscript.

Authors' contribution

The authors thank the study participants. Bilin Tao, Beibei Qiu and Jianming Wang conceived, initiated, and led the study. Bilin Tao, Jizhou Wu, Zhongqi Li, Mengyao Zhan and Zhuchao Wu collected the data. Bilin Tao, Jizhou Wu, and Beibei Qiu analyzed the data with input from all the authors. Bilin Tao, Beibei Qiu prepared the manuscript. All authors reviewed and approved the manuscript.

Acknowledgments

We wish to acknowledged the platform TB Profiler (https://tbdr.lshtm.ac.uk/) and the entire staff using this website.

Authors' information

[1]Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, 211166, China

[a]These authors contributed equally to this work.

*Correspondence: Jianming Wang. Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, 101 Longmian Ave. Nanjing, 211166, China. Tel: +86-25-86868438. Email: jmwang@njmu.edu.cn

References

1. Organization WH: **Global tuberculosis report 2020**. In.: Geneva; 2020.
2. Gardy JL, Johnston JC, Ho Sui SJ, Cook VJ, Shah L, Brodkin E, Rempel S, Moore R, Zhao Y, Holt R *et al*: **Whole-genome sequencing and social-network analysis of a tuberculosis outbreak**. *N Engl J Med* 2011, **364**(8):730–739.
3. Consortium CR, the GP, Allix-Beguec C, Arandjelovic I, Bi L, Beckert P, Bonnet M, Bradley P, Cabibbe AM, Cancino-Munoz I *et al*: **Prediction of Susceptibility to First-Line Tuberculosis Drugs by DNA Sequencing**. *N Engl J Med* 2018, **379**(15):1403-1415.
4. Cole ST, Brosch R, Parkhill J, Garnier T, Churcher C, Harris D, Gordon SV, Eiglmeier K, Gas S, Barry CE, 3rd *et al*: **Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence**. *Nature* 1998, **393**(6685):537–544.
5. Coll F, McNerney R, Preston MD, Guerra-Assunção JA, Warry A, Hill-Cawthorne G, Mallard K, Nair M, Miranda A, Alves A *et al*: **Rapid determination of anti-tuberculosis drug resistance from whole-genome sequences**. *Genome medicine* 2015, **7**(1):51.
6. Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, Farhat MR, Guthrie JL, Laukens K, Miotto P *et al*: **Whole genome sequencing of Mycobacterium tuberculosis: current standards and open issues**. *Nature reviews Microbiology* 2019, **17**(9):533–545.
7. Cohen KA, Manson AL, Desjardins CA, Abeel T, Earl AM: **Deciphering drug resistance in Mycobacterium tuberculosis using whole-genome sequencing: progress, promise, and challenges**. *Genome medicine* 2019, **11**(1):45.
8. Gostin LO, Katz R: **The International Health Regulations: The Governing Framework for Global Health Security**. *Milbank Q* 2016, **94**(2):264–313.
9. Boyd MJ, Wilson N, Nelson C: **Validation analysis of Global Health Security Index (GHSI) scores 2019**. *BMJ Glob Health* 2020, **5**(10).

10. Phelan J, O'Sullivan DM, Machado D, Ramos J, Whale AS, O'Grady J, Dheda K, Campino S, McNerney R, Viveiros M *et al*: **The variability and reproducibility of whole genome sequencing technology for detecting resistance to anti-tuberculous drugs**. *Genome Med* 2016, **8**(1):132.
11. Phelan JE, O'Sullivan DM, Machado D, Ramos J, Oppong YEA, Campino S, O'Grady J, McNerney R, Hibberd ML, Viveiros M *et al*: **Integrating informatics tools and portable sequencing technology for rapid detection of resistance to anti-tuberculous drugs**. *Genome Med* 2019, **11**(1):41.
12. Blower SM, Chou T: **Modeling the emergence of the 'hot zones': tuberculosis and the amplification dynamics of drug resistance**. *Nat Med* 2004, **10**(10):1111–1116.
13. Vasilyeva I, Mariandyshev A, Kazennyy B, Davidaviciene E, Lounis N, Keim S, on behalf of the TMCTBCsg: **Early access to bedaquiline for extensively drug-resistant (XDR) and pre-XDR tuberculosis**. *Eur Respir J* 2019, **54**(1).
14. Lamina C, Sturm G, Kollerits B, Kronenberg F: **Visualizing interaction effects: a proposal for presentation and interpretation**. *J Clin Epidemiol* 2012, **65**(8):855–862.
15. Babitsch B, Gohl D, von Lengerke T: **Re-revisiting Andersen's Behavioral Model of Health Services Use: a systematic review of studies from 1998-2011**. *Psychosoc Med* 2012, **9**(11):Doc11.
16. Coscolla M, Gagneux S: **Consequences of genomic diversity in Mycobacterium tuberculosis**. *Semin Immunol* 2014, **26**(6):431–444.
17. Ford CB, Shah RR, Maeda MK, Gagneux S, Murray MB, Cohen T, Johnston JC, Gardy J, Lipsitch M, Fortune SM: **Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis**. *Nat Genet* 2013, **45**(7):784–790.
18. Reiling N, Homolka S, Walter K, Brandenburg J, Niwinski L, Ernst M, Herzmann C, Lange C, Diel R, Ehlers S *et al*: **Clade-specific virulence patterns of Mycobacterium tuberculosis complex strains in human primary macrophages and aerogenically infected mice**. *mBio* 2013, **4**(4).
19. Castro RAD, Borrell S, Gagneux S: **The within-host evolution of antimicrobial resistance in Mycobacterium tuberculosis**. *FEMS Microbiol Rev* 2020.
20. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, Blum M, Rusch-Gerdes S, Mokrousov I, Aleksic E *et al*: **Evolutionary history and global spread of the Mycobacterium tuberculosis Beijing lineage** (**vol 47**, **pg 242**, 2015). *Nature Genetics* 2015, **47**(3).
21. Ford CB, Shah RR, Maeda MK, Gagneux S, Murray MB, Cohen T, Johnston JC, Gardy J, Lipsitch M, Fortune SM: **Mycobacterium tuberculosis mutation rate estimates from different lineages predict substantial differences in the emergence of drug-resistant tuberculosis**. *Nat Genet* 2013, **45**(7):784-+.
22. Comas I, Gagneux S: **A role for systems epidemiology in tuberculosis research**. *Trends Microbiol* 2011, **19**(10):492–500.
23. Wiens KE, Woyczynski LP, Ledesma JR, Ross JM, Zenteno-Cuevas R, Goodridge A, Ullah I, Mathema B, Djoba Siawaya JF, Biehl MH *et al*: **Global variation in bacterial strains that cause tuberculosis disease: a systematic review and meta-analysis**. *BMC Med* 2018, **16**(1):196.

24. Napier G, Campino S, Merid Y, Abebe M, Woldeamanuel Y, Aseffa A, Hibberd ML, Phelan J, Clark TG: **Robust barcoding and identification of Mycobacterium tuberculosis lineages for epidemiological and clinical studies**. *Genome Med* 2020, **12**(1):114.
25. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, Parkhill J, Malla B, Berg S, Thwaites G *et al*: **Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans**. *Nat Genet* 2013, **45**(10):1176–1182.
26. Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, Roach JC, Kremer K, Petrov DA, Feldman MW *et al*: **High Functional Diversity in Mycobacterium tuberculosis Driven by Genetic Drift and Human Demography**. *Plos Biol* 2008, **6**(12):2658–2671.
27. Wirth T, Hildebrand F, Allix-Beguec C, Wolbeling F, Kubica T, Kremer K, van Soolingen D, Rusch-Gerdes S, Locht C, Brisse S *et al*: **Origin, spread and demography of the Mycobacterium tuberculosis complex**. *Plos Pathog* 2008, **4**(9):e1000160.
28. Su Y, Garcia Baena I, Harle AC, Crosby SW, Micah AE, Siroka A, Sahu M, Tsakalos G, Murray CJL, Floyd K *et al*: **Tracking total spending on tuberculosis by source and function in 135 low-income and middle-income countries, 2000-17: a financial modelling study**. *Lancet Infect Dis* 2020, **20**(8):929–942.
29. Kandel N, Chungong S, Omaar A, Xing J: **Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries**. *Lancet* 2020, **395**(10229):1047-1053.
30. Coll F, McNerney R, Guerra-Assuncao JA, Glynn JR, Perdigao J, Viveiros M, Portugal I, Pain A, Martin N, Clark TG: **A robust SNP barcode for typing Mycobacterium tuberculosis complex strains**. *Nat Commun* 2014, **5**:4812.

Figures

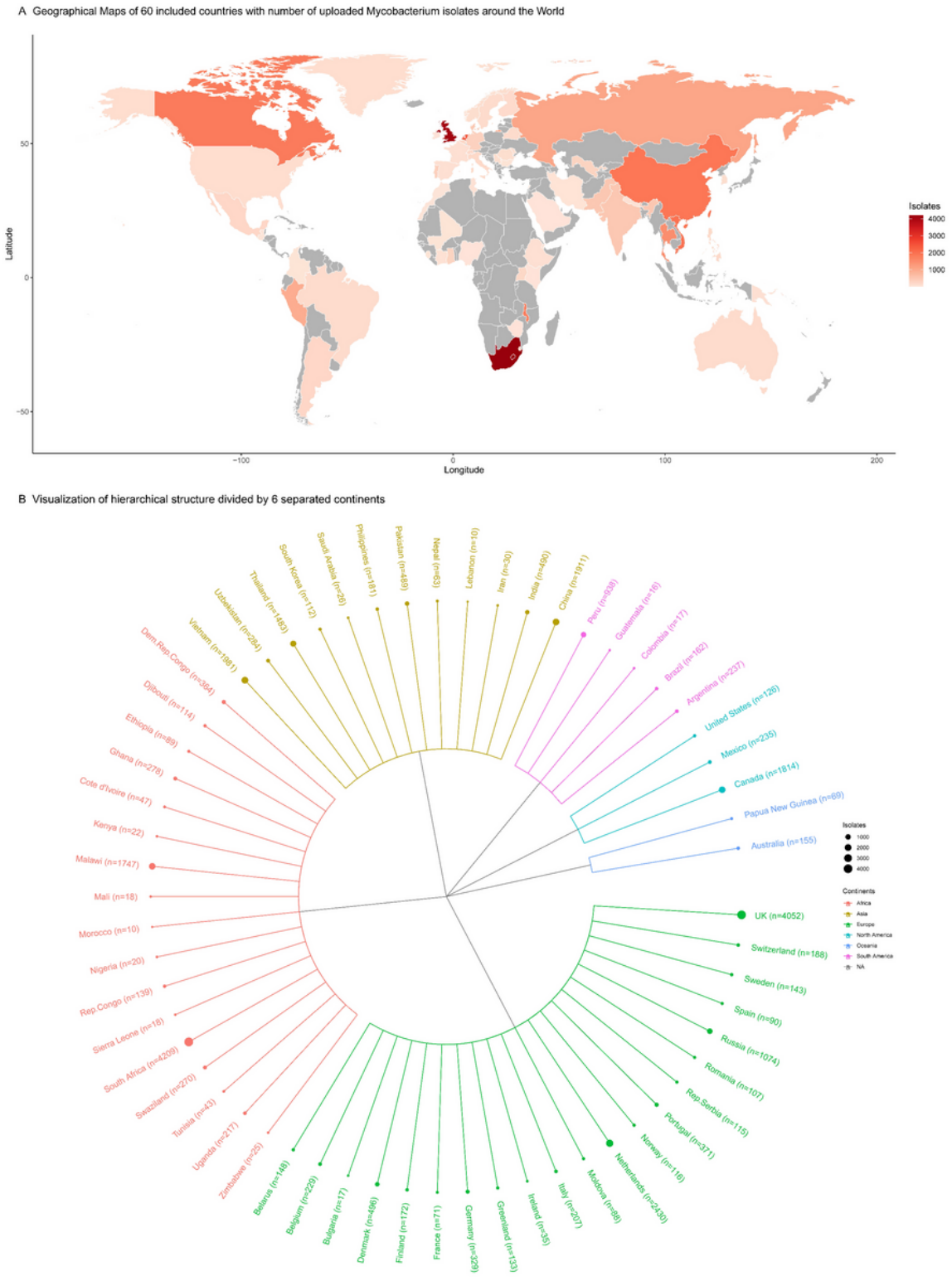

Figure 1

Countries with the number of uploaded Mycobacterium isolates and visualization of the hierarchical structure divided by 6 separated continents A: The number of isolated strains was used to scale the colors of different depths for each country. B: The number of isolated strains was used to scale the size of the circle point of each country.

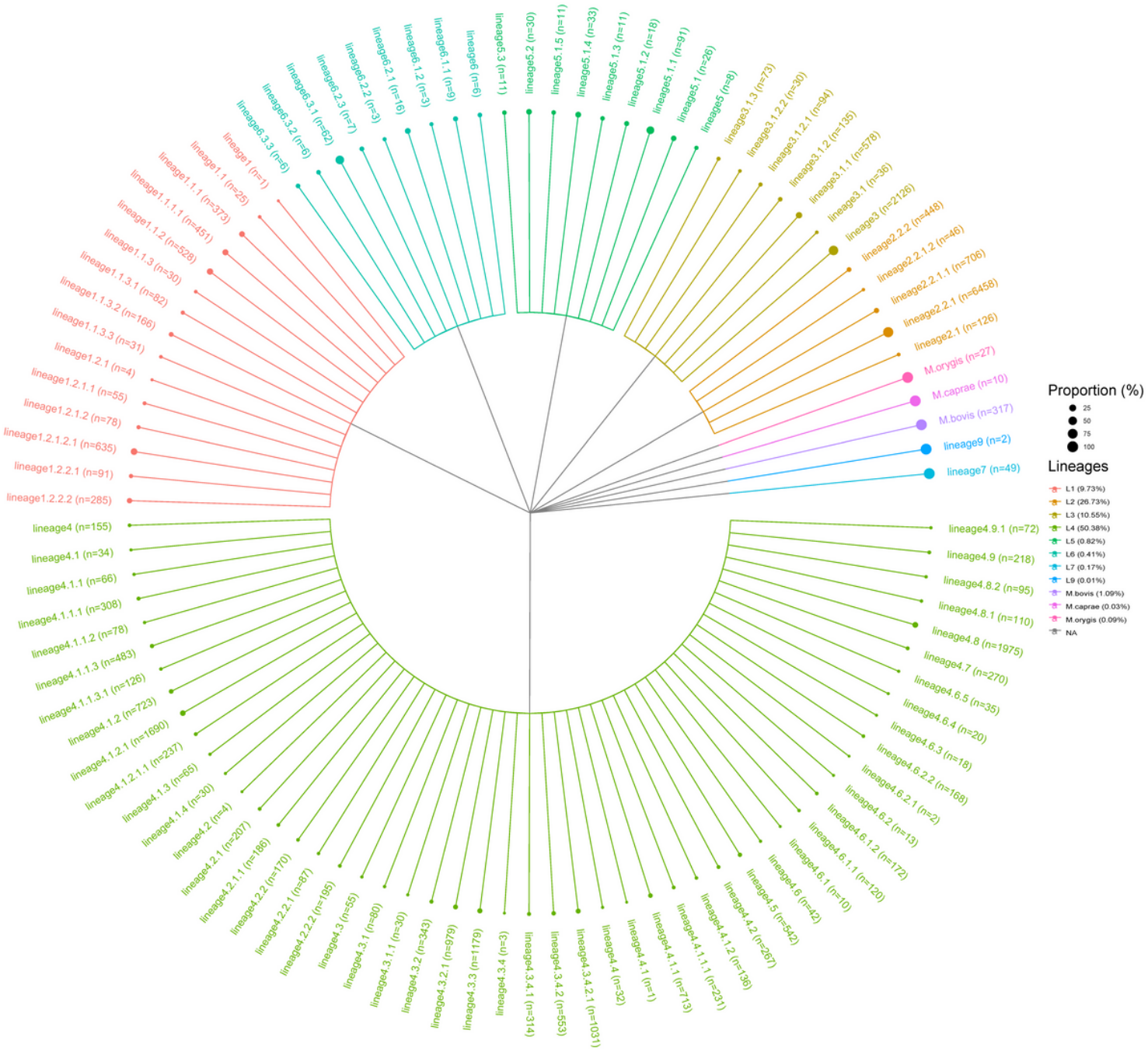

Figure 2

Geographical Maps of major mycobacterium lineages in 60 included countries around the world and distribution of different drug resistance types in each lineage DR: drug resistance, regarded as strains resistant to at least isoniazid and rifampin; MDR: multi-drug resistance, strains resistant to isoniazid and rifampin; XDR: extensive drug resistance, on the base of MDR, the additional resistance to the fluoroquinolones and injectable drugs (amikacin, kanamycin, and capreomycin).

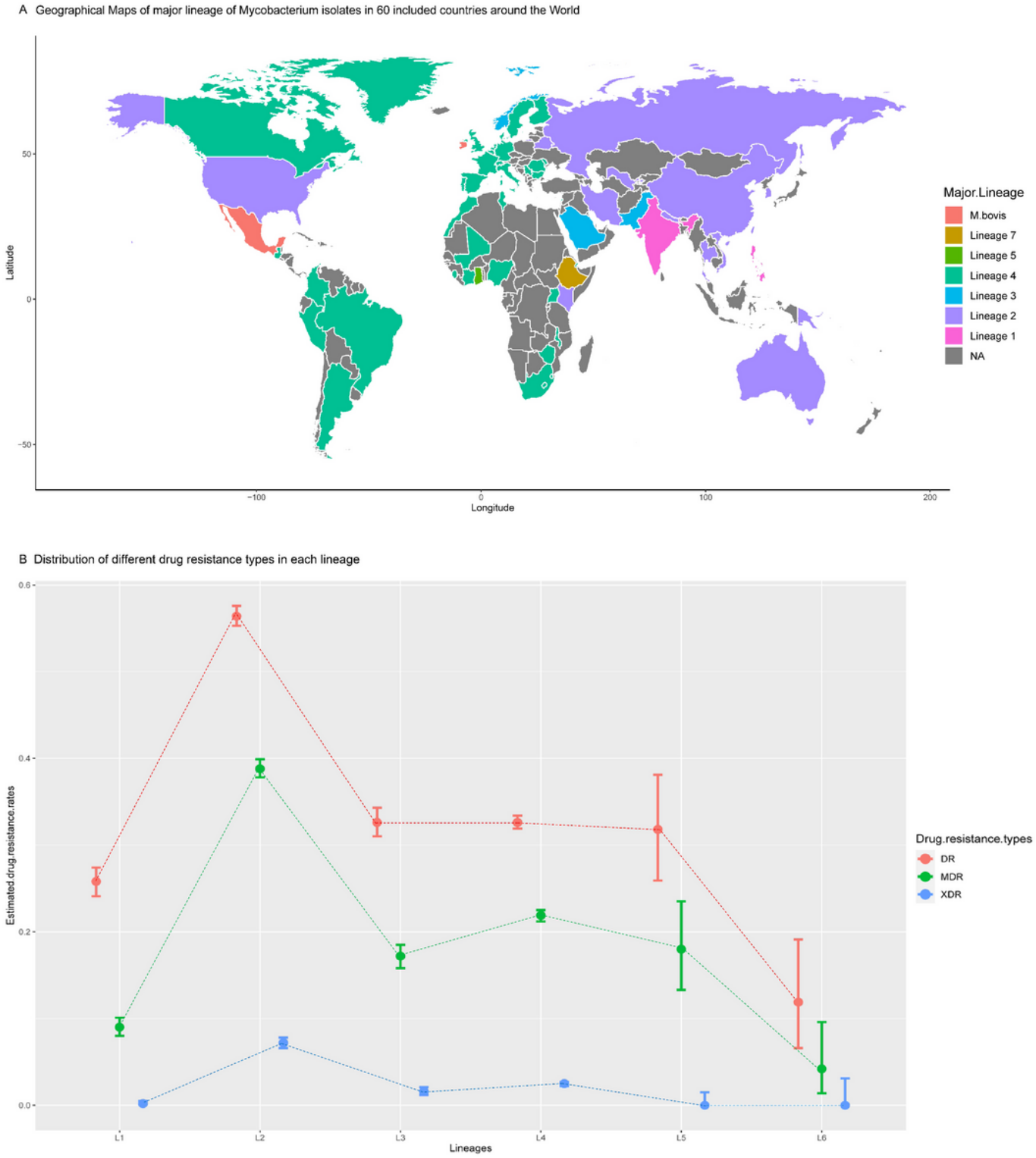

Figure 3

Visualization of the hierarchical structure divided by 11 lineages consisting of 101 sublineages The internal constituent ratio was used to scale the size of circle points for each lineage.

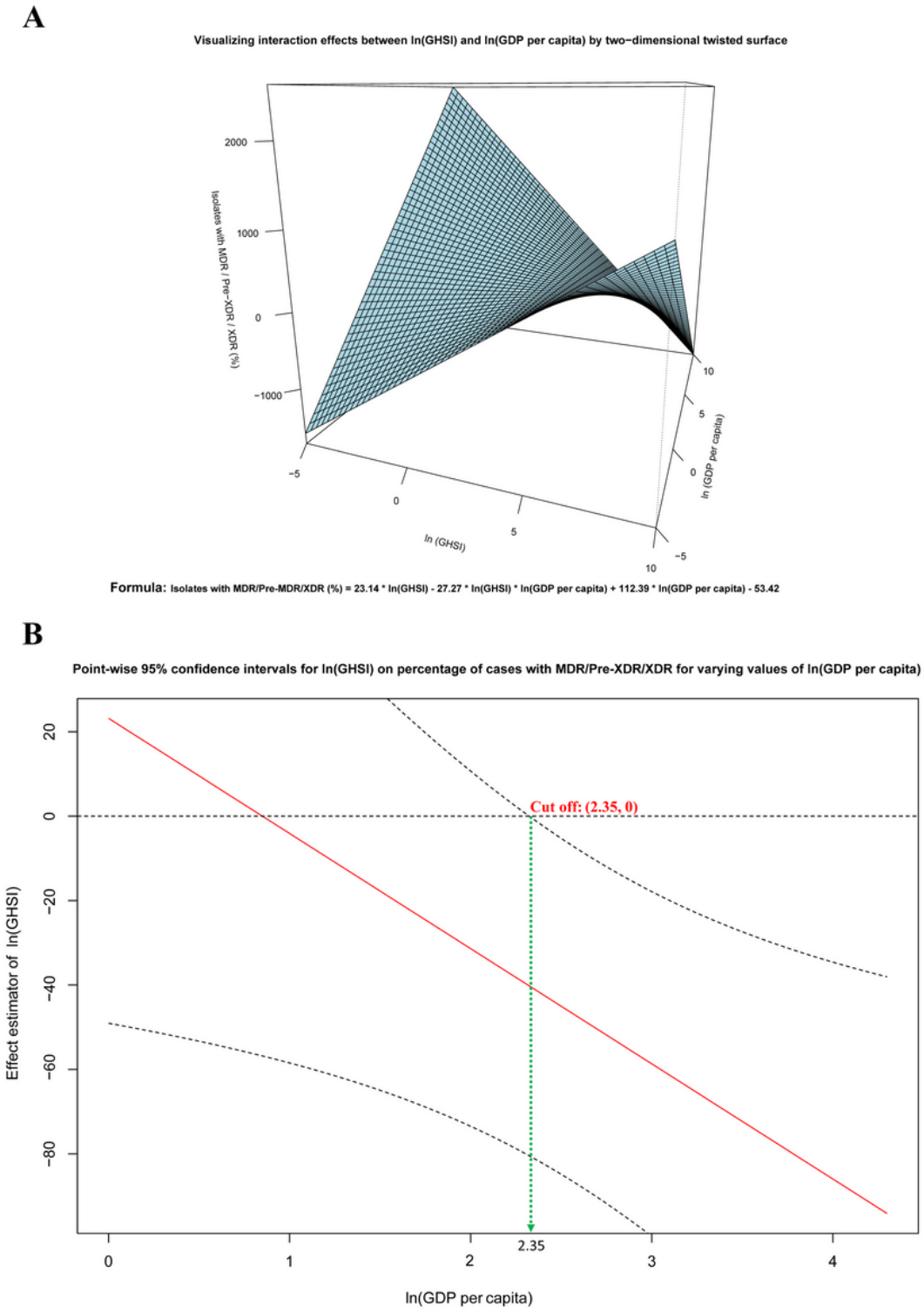

Figure 4

Visualizing interaction effects on the prevalence of MDR between ln(GHSI) and ln(GDP per capita) by two-dimensional twisted surface GHSI: global health score index; GDP: gross domestic product; The GHSI and GDP per capita were logarithmic transformed before analysis.

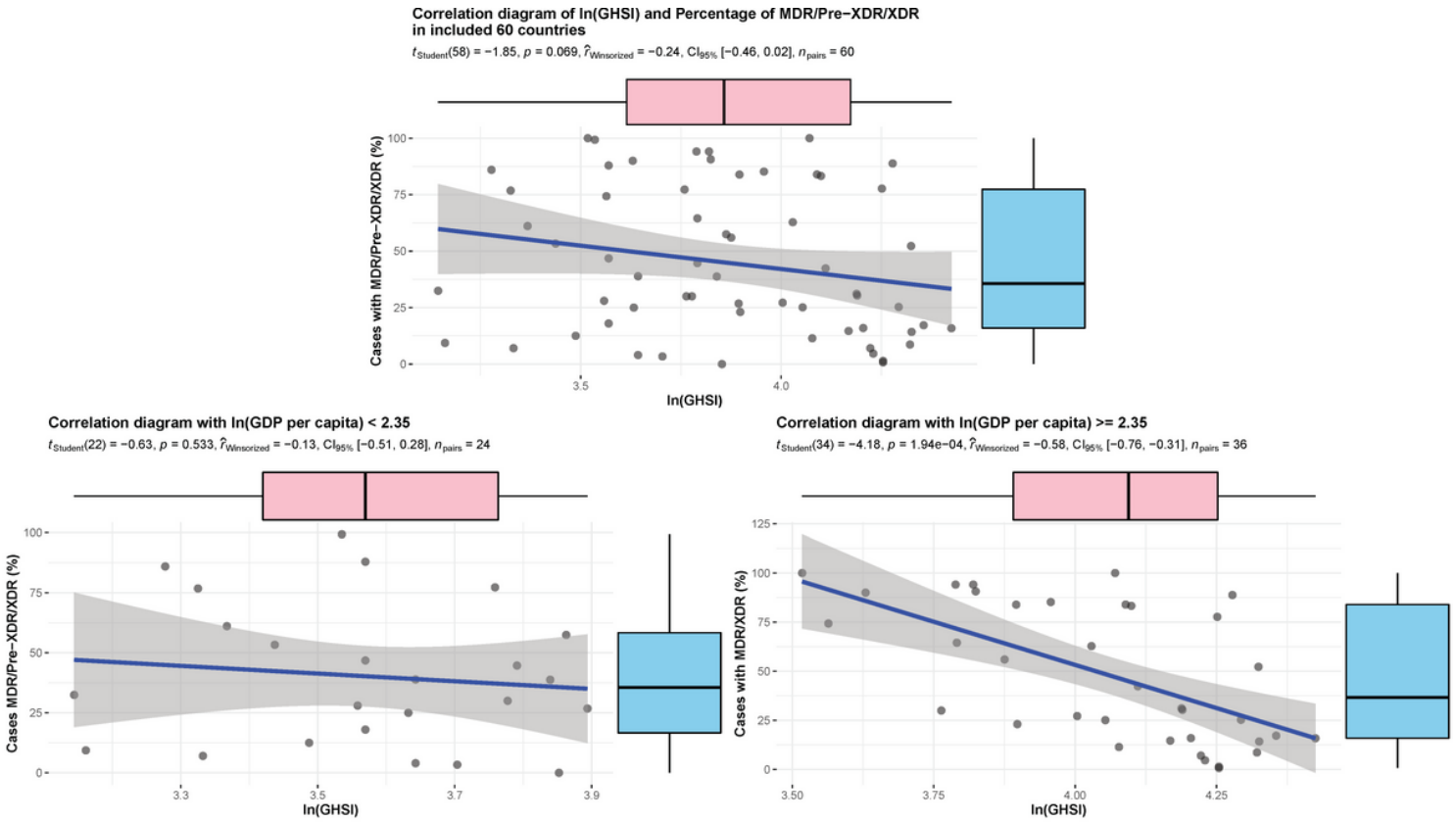

Figure 5

Correlation diagram of ln(GHSI) and prevalence of MDR GHSI: global health score index; GDP: gross domestic product; The GHSI and GDP per capita were logarithmic transformed before analysis.

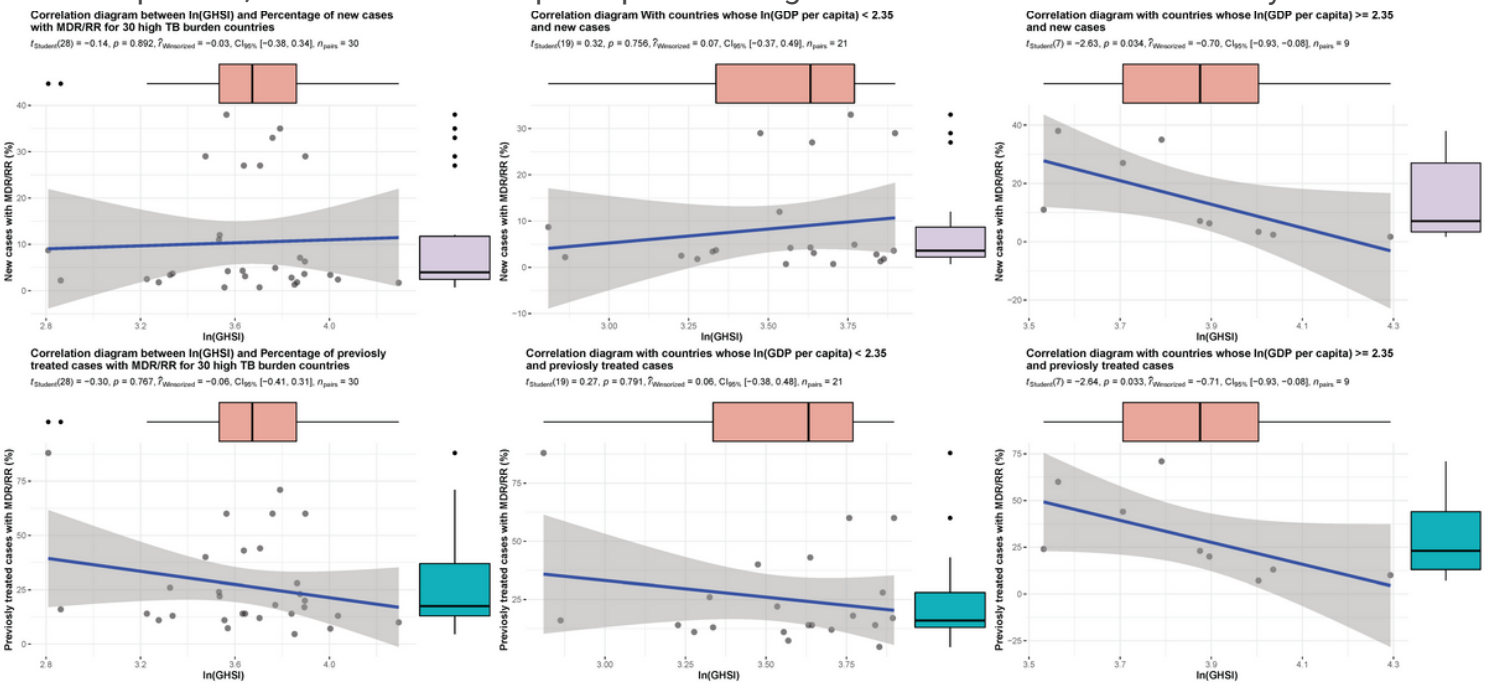

Figure 6

Verifying the correlation of ln(GHSI) and prevalence of MDR by using data of MDR/RR for 30 high TB burden countries

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfigure1.tiff