



The Brazilian Journal of INFECTIOUS DISEASES

www.elsevier.com/locate/bjid



Brief Communication

Whole genome characterization, and geographical distribution of *M. tuberculosis* in central region of Veracruz, Mexico

Esdras Antonio Fernández-Morales^{a,b}, Gustavo Bermudez^{id c},
Hilda Montero^b, Manuel Luzania-Valerio^b, Roberto Zenteno-Cuevas^{id b,d,*}

^a University of Veracruz, Health Sciences Institute, Health Sciences Master Program, Mexico

^b University of Veracruz, Public Health Institute, Mexico

^c University of Veracruz, Health Sciences Institute, Health Sciences Doctoral Program, Mexico

^d Multidisciplinary Network of Tuberculosis Research, Mexico

ARTICLE INFO

Article history:

Received 26 January 2022

Accepted 14 April 2022

Available online 6 May 2022

Keywords:

Tuberculosis

WGS

Geolocation

ABSTRACT

The purpose of this work was to perform by Whole Genomic Sequencing (WGS) a characterization of tuberculosis isolates circulating in the central region of Veracruz, Mexico, and to determine its geographical distribution. The genome of 25 clinical isolates of tuberculosis patients, recovered from central zone of Veracruz, Mexico, were sequenced and the information obtained was used to characterize lineage, prediction of drug resistance, identification of clonal complexes, and finally correlated with the geolocation data. Isolates analyzed were included into seven L4 sublineages, most frequent was X3; X1 (4.1.1.3) in 35%. rpoBSer450Leu polymorphism was the most frequently found variant. Sublineage Haarlem (4.1.2) had the widest distribution, found in five municipalities. Of the of two clonal complexes found, the most abundant included eight isolates, with X3/X1 lineage, placed in two municipalities. Combination of WGS and geographic information system was very useful for the identification of sublineages, clonal complexes, and their geographical dispersion with important implications in the epidemiological surveillance and clinical control of TB.

© 2022 Sociedade Brasileira de Infectologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Tuberculosis (TB) remains one of the most important infectious diseases in the world. According to the WHO global tuberculosis report 2020, almost 10 million people became ill, and 1.2 million deaths were due to TB.¹

In the past decade, the standardization of whole genome sequencing (WGS) had an important impact on the characterization and understanding of TB, thanks to its higher

resolution power to predict drug resistance, identify lineage, transmission routes, and geographic distribution.^{2,3} For these reasons, WGS has been used as a tool for the development of genomic surveillance programs in TB in developed countries. However, the use of WGS in low and middle-income settings, usually those with the highest burden of disease, has been scarce.⁴

In Mexico, TB has an incidence of 22 cases per 100,000 inhabitants, from which, 3% had resistance to at least one drug.¹ Multidrug- (MDR-TB) and extensively drug resistant-

* Corresponding author at.

E-mail address: rzenteno@uv.mx (R. Zenteno-Cuevas).

<https://doi.org/10.1016/j.bjid.2022.102357>

1413-8670/© 2022 Sociedade Brasileira de Infectologia. Published by Elsevier España, S.L.U. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

tuberculosis (XDR-TB) have been increasing annually, placing Mexico among the five main TB contributors in the Americas.⁵ Unfortunately there are few studies related to the characterization of TB in Mexico, using WGS,^{6,7} thus limiting the evaluation of its effectiveness. Considering the above, this study aimed to use WGS to characterize the *M. tuberculosis* isolates circulating in the central region of the state of Veracruz, Mexico, and by use of geographic information system determine the geographic distribution of the lineages found.

Twenty-five sputum samples from patients with confirmed TB diagnosis from January 2018-July 2019 were randomly collected by the Mycobacteriosis Program from the Veracruz Ministry of Health in Mexico from 125 clinical cases detected in the period. This geographic area includes 25 municipalities, covers an area of 5,000 km² and gives attention to more than one million inhabitants (Fig. 1).

Phenotypic drug sensitivity test (DST) against first-line drugs was performed by the staff of the Tuberculosis Program of the Health Secretary of the State of Veracruz using the fluorometric method (BACTEC, MGIT 960 Becton-Dickinson), according to standard conditions: isoniazid (H) > 0.1 µg/mL, rifampin (R) 1.0 µg/mL, ethambutol (E) 5.0 µg/mL and streptomycin (S) 1.0 µg/mL. Pyrazinamide sensitivity was determined using a BACTEC MGIT 960 PZA kit (Becton Dickinson). Clinical, and epidemiological variables were abstracted from medical records.

Mycobacterium tuberculosis (MTB) strains were isolated in Lowenstein-Jensen media and DNA was extracted following CTAB method.⁸ Libraries were prepared according to Nextera XT protocol (Illumina, CA., USA), using 1 ng of DNA quantified by Qubit fluorometer (Invitrogen, CA, USA), library quality was determined using Bioanalyzer 2100 (Agilent Genomics), and sequenced using MiSeq (Illumina, CA., USA) in a 2 × 300 paired-end format.

Quality control of reads and trimmed was performed by using FastQC and Trimmomatic software, respectively. Reads

were aligned using BWA,⁹ considering as reference, the genome of *M. tuberculosis* H37Rv (GenBank: NC_000962.3). Mapping, and variant calling was performed using a previously reported pipeline (http://tgu.ibv.csic.es/?page_id=1794).^{10,11} Variants in 20 reads, and at ≥ 90% of frequency within each isolate were called fixed-single nucleotide polymorphism (SNP), used to detect phylogenetic mutations. Variants in at least 10 reads at ≥10% to ≤90% frequency were called no fixed-SNPs and were used to detect antibiotic resistance. Analysis and selection of SNPs and INDELS related to antituberculosis drugs were done as describe by Madrazo-Moya et al.¹²

The data underlying the genome sequences presented in this study are available at ENA: <https://www.ebi.ac.uk/ena>. Accession number: PRJEB30933.

Strains were classified according to 62 phylogenetic variants associated with lineages as proposed by Coll et al.¹³ Phylogeny and genomic clusters were determined following recommendations by Jimenez-Ruano et al.⁷ and final tree visualization was done with Ito.¹⁴ Calculation of pairwise genetic distances between isolates considered a ≤ 12 SNPs threshold, as recommend Walker et al.²

Geographical distribution of the isolates (GIS) was carried out considering the geographic coordinates of the place of residence from the individuals bearing the isolates, and this was used to construct the spatial representation of the isolates using the geographic information system, QGIS (<https://www.qgis.org/en/site/>).

The results showed that individuals carrying isolates 88% (22) were male, mean age 43±13.2 years. The main comorbidity was type 2 diabetes mellitus at 32% (8), and 48% (12) reported a previous contact with a patient with TB in past year. Rifampicin resistance was observed in nine isolates (36%), followed by isoniazid in six (24%), pyrazinamide in three (12%), and five cases (20%) were MDR-TB.

According to the genomic resistance prediction, resistance to rifampicin was found in nine isolates (39%), followed by

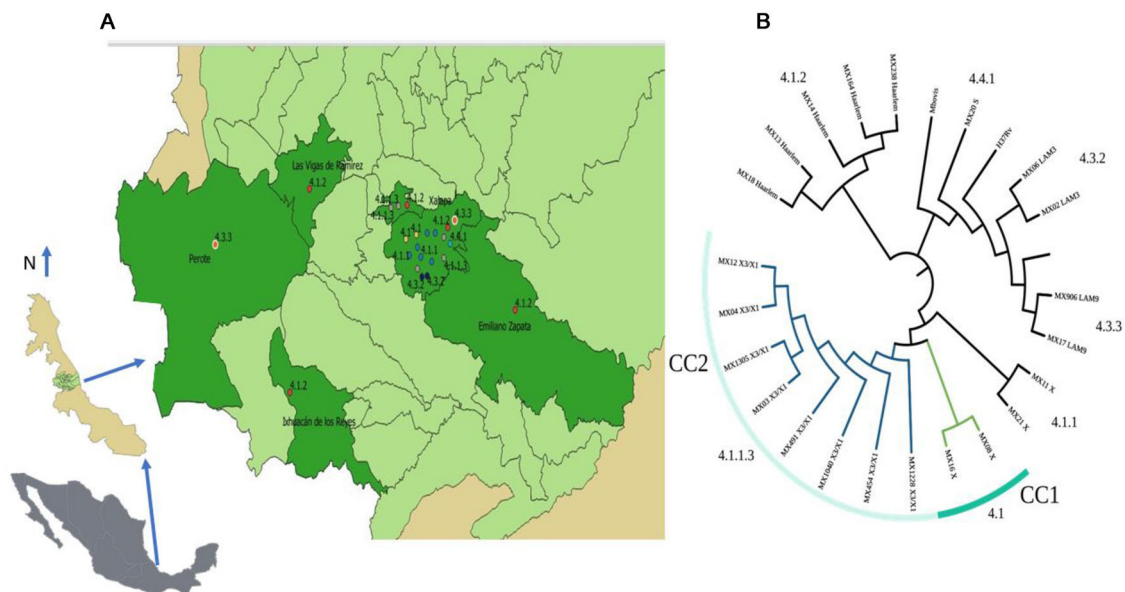


Fig. 1 – (A) Geographical distribution of *M. tuberculosis* lineages circulating in the center region of Veracruz, México. (B) Phylogenetic tree of isolates of circulating showing, lineages and clonal complexes found in the center region of Veracruz, Mexico.

the strong association with drug resistance. The municipality of Xalapa contained the highest number of cases, sublineages, and CCs, and most of them were placed in marginal colonies. This concentration could be explained by the fact that is the capital city of the state and have an important number of inhabitants living in poverty, marginalization, overcrowding, and precarious urban conditions.

To our knowledge, this is the first characterization of TB by WGS and the combined analysis with the geographic location and distribution of specific sublineages and CCs circulating in Mexico. These results confirm the diversity of sublineages circulating and the usefulness of GIS to generate a better definition of the epidemiological behavior of TB, similar to that described in China and Myanmar.^{20,21}

The fact that 43% of the isolates were grouped into two clonal complexes, evidence the high transmission level of some TB clones in the population, and the need to strengthen the contact studies in active cases to avoid major dispersion of these strains, and also confirm the need to include WGS analysis in epidemiological surveillance.

Despite the main limitation of the study being the low number of isolates analyzed, it is possible to conclude that combination of WGS and GIS have a deep impact in the characterization of the tuberculosis isolates circulating improving the description of distribution and dispersion of sublineages and clonal complexes. This combination will have important implications in terms of epidemiology, surveillance, and control of TB.

Statement of ethics

All aspects related with this research were approved by the Ethics Committee of the Public Health Institute of the Universidad Veracruzana (CEI-ISP-R18/2020). All subjects gave their written informed consent and all protocols were performed in accordance with the national guidelines and regulations.

Funding

EA F-M was funded by CONACyT, fellow number [850447](#). R Z-C was partially funded by CONACyT No. [A1-22956](#) and [ICOP-COOPB20344](#).

Authors' contributions

EAF-M, conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, writing original draft, Writing review & editing. GB, Data curation, Formal analysis, Methodology, Validation, Visualization. MLV, Methodology, Validation, Visualization, Writing original draft. HM, Data curation, Visualization, writing original draft, Writing review & editing. RZ-C, conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, writing original draft, Writing review & editing.

Data availability statement

The data underlying the genome sequences presented in this study are available at ENA: <https://www.ebi.ac.uk/ena>. Accession number: PRJEB30933.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgement

Special thanks to the Dr. Everest de Igartua, Dr. Javier Fuentes, Dra. Leila Escobar and the personal of the tuberculosis program of the health secretary of the state of Veracruz

REFERENCES

1. WHO. WHO | Global tuberculosis report 2019. Geneva: World Health Organization; 2020.
2. Walker TM, Ip CLC, Harrell RH, et al. Whole-genome sequencing to delineate Mycobacterium tuberculosis outbreaks: a retrospective observational study. *Lancet Infect Dis.* 2013;13:137–46.
3. Meehan CJ, Goig GA, Kohl TA, et al. Whole genome sequencing of Mycobacterium tuberculosis: current standards and open issues. *Nat Rev Microbiol.* 2019: 533–45.
4. Rivière E, Heupink TH, Ismail N, et al. Capacity building for whole genome sequencing of Mycobacterium tuberculosis and bioinformatics in high TB burden countries. *Brief Bioinform.* 2020.
5. Pan American Health Organisation. Tuberculosis in the Americas 2018. PAHO; 2018.
6. Rendón-Bautista LA, Álvarez-Maya I, Sandoval-Díaz M, Villanueva-Arias JC, Ayala-Chavira N, Zenteno-Cuevas R. Characterization of genetic diversity and clonal complexes by whole genome sequencing of Mycobacterium tuberculosis isolates from Jalisco, Mexico. *Tuberculosis.* 2021;129:102106.
7. Jiménez-Ruano AC, Madrazo-Moya CF, Cancino-Muñoz I, et al. Whole genomic sequencing based genotyping reveals a specific X3 sublineage restricted to Mexico and related with multidrug resistance. *Sci Rep.* 2021;11:1870.
8. Van Soolingen D, Hermans PWM, De Haas PEW, Soll DR, Van Embden JDA. Occurrence and stability of insertion sequences in Mycobacterium tuberculosis complex strains: Evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. *J Clin Microbiol.* 1991;29:2578–86.
9. Picard Tools - By Broad Institute. Available at: <https://broadinstitute.github.io/picard/>[accessed March 29, 2021].
10. Comas I, Coscolla M, Luo T, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. *Nat Genet.* 2013;45:1176–82.
11. Xu Y, Cancino-Munoz I, Torres-Puente 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:e1002961.
12. Madrazo-Moya CF, Cancino-Muñoz I, Cuevas-Córdoba B, 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;14:1–17.

13. Coll F, McNerney R, Guerra-Assunção JA, et al. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. *Nat Commun.* 2014;5:4–8.
14. Letunic I, Bork P. Interactive tree of life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47(W1):W256–9.
15. Munro-Rojas D, Fernandez-Morales E, Zarrabal-Meza J, et al. Genetic diversity of drug and multidrug-resistant *Mycobacterium tuberculosis* circulating in Veracruz, Mexico. *PLoS One.* 2018;13:e0193626.
16. Papaventsis D, Casali N, Kontsevaya I, Drobniowski F, Cirillo DM, Nikolayevskyy V. Whole genome sequencing of *Mycobacterium tuberculosis* for detection of drug resistance: a systematic review. *Clin Microbiol Infect.* 2017;23:61–8.
17. Bradley P, Gordon NC, Walker TM, et al. Rapid antibiotic-resistance predictions from genome sequence data for *Staphylococcus aureus* and *Mycobacterium tuberculosis*. *Nat Commun.* 2015;6:10063.
18. Wiens KE, Woyczynski LP, Ledesma JR, et al. Global variation in bacterial strains that cause tuberculosis disease: a systematic review and meta-analysis. *BMC Med.* 2018: 196.
19. Stucki D, Brites D, Jeljeli L, et al. *Mycobacterium tuberculosis* lineage 4 comprises globally distributed and geographically restricted sublineages. *Nat Genet.* 2016;48:1535–43.
20. Xu A-M, He C-J, Cheng X, et al. Distribution and identification of *Mycobacterium tuberculosis* lineage in Kashgar prefecture. *BMC Infect Dis.* 2022;22:312.
21. Maung HMW, Palittapongarnpim P, Aung HL, Surachat K, Nyunt WW, Chongsuvivatwong V. Geno-Spatial distribution of *mycobacterium tuberculosis* and drug resistance profiles in myanmar-thai border area. *Trop Med Infect Dis.* 2020;5:153.