Infection, Genetics and Evolution 73 (2019) 337–341

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Short communication

Mycobacterium tuberculosis lineage 1 genetic diversity in Pará, Brazil, suggests common ancestry with east-African isolates potentially linked to historical slave trade

Emilyn Costa Conceição[a,b,*], Guislaine Refregier[a], Harrison Magdinier Gomes[a,c], Xavier Olessa-Daragon[a], Francesc Coll[d], Noël Harijaona Ratovonirina[a,e], Voahangy Rasolofo-Razanamparany[e], Maria Luiza Lopes[f], Dick van Soolingen[g,h], Liliana Rutaihwa[i,j], Sebastien Gagneux[i,j], Valdes Roberto Bollela[k], Philip Noel Suffys[c], Rafael Silva Duarte[b], Karla Valéria Batista Lima[f], Christophe Sola[a,k,*]

[a] Institut de Biologie Intégrative de la Cellule, I2BC, UMR9198, CEA, CNRS, Univ. Paris-Sud, Univ. Paris-Saclay, 91198 Gif-sur-Yvette cedex, France
[b] Pós-Graduação Instituto de Microbiologia Professor Paulo de Góes, Universidade Federal do Rio de Janeiro, Rio de Janeiro-RJ, Brazil
[c] Laboratório de Biologia Molecular Aplicada a Micobactéria, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro-RJ, Brazil
[d] Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, WC1E 7HT London, UK
[e] Unité des Mycobactéries, Institut Pasteur de Madagascar, Antananarivo, Madagascar
[f] Seção de Bacteriologia e Micologia, Instituto Evandro Chagas, Ananindeua-PA, Brazil
[g] National Institute for Public Health and the Environment (RIVM), Bilthoven, the Netherlands
[h] Department of Medical Microbiology, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands
[i] Swiss Tropical & Public Health Institute, Socinstrasse 57, 4051 Basel, Switzerland
[j] University of Basel, Basel, Switzerland
[k] Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto-SP, Brazil

A R T I C L E I N F O

Keywords:
Mycobacterium tuberculosis complex
Spoligotyping
Lineage 1
Whole genome sequencing
Single-nucleotide polymorphisms
Molecular evolution

A B S T R A C T

Lineage 1 (L1) is one of seven *Mycobacterium tuberculosis* complex (MTBC) lineages. The objective of this study was to improve the complex taxonomy of L1 using phylogenetic SNPs, and to look for the origin of the main L1 sublineage prevalent in Para, Brazil. We developed a high-throughput SNPs-typing assay based on 12-L1-specific SNPs. This assay allowed us to experimentally retrieve SNP patterns on nine of these twelve SNPs in 277 isolates previously tentatively assigned to L1 spoligotyping-based sublineages. Three collections were used: Pará-Brazil (71); RIVM, the Netherlands (102), Madagascar (104). One-hundred more results were generated *in Silico* using the PolyTB database. Based on the final SNPs combination, the samples were classified into 11 clusters (C1-C11). Most isolates within a SNP-based cluster shared a mutual spoligotyping-defined lineage. However, L1/EAI1-SOM (SIT48) and L1/EAI6-BGD1 (SIT591) showed a poor correlation with SNP data and are not monophyletic. L1/EAI8-MDG and L1/EAI3-IND belonged to C5; this result suggests that they share a common ancestor. L1.1.3/SIT129, a spoligotype pattern found in SNPs-cluster C6, was found to be shared between Pará/Brazil and Malawi. SIT129 was independently found to be highly prevalent in Mozambique, which suggests a migration history from East-Africa to Brazil during the 16th–18th slave trade period to Northern Brazil.

1. Introduction

Insights from Whole Genome Sequencing (WGS) provided a classification of the *Mycobacterium tuberculosis* complex (MTBC) into seven main human-adapted lineages (Lineages 1–7) and has demonstrated a larger genetic diversity than expected (Coscolla and Gagneux,

Abbreviations: DPO, dual-priming oligonucleotide; DR, Direct Repeat; EAI, East-African-Indian; Lineage 1, L1; MTBC, *Mycobacterium tuberculosis* complex; Mutant, Mut; SA-PE, Streptavidin-Phycoerythrin; SIT, Shared International Type; SNPs, single-nucleotide polymorphisms; TE, Tris-EDTA; WGS, whole-genome sequencing; Wild-Type, Wt

* Corresponding authors at: Institut de Biologie Intégrative de la Cellule, I2BC, UMR9198, CEA, CNRS, Univ. Paris-Sud, Univ. Paris-Saclay, 91198 Gif-sur-Yvette cedex, France.

E-mail addresses: emilyncosta@gmail.com (E.C. Conceição), christophe.sola@i2bc.paris-saclay.fr (C. Sola).

https://doi.org/10.1016/j.meegid.2019.06.001
Received 14 April 2019; Received in revised form 30 May 2019; Accepted 1 June 2019
Available online 03 June 2019

2014). PolyTB (Coll et al., 2014b), a WGS web-database, is a powerful tool to investigate deep phylogenetic correlation among MTBC lineages based on 1627 *M. tuberculosis* isolates, containing 74,039 (63% non-synonymous) high quality SNPs (Coll et al., 2014a). These markers are linked to spoligotypes, among which 62 allow to define a barcode typing system. Spoligotyping patterns, based on CRISPR-*locus* (Direct Repeat region) polymorphism, has previously revealed strong phylogeographical specificity (Brudey et al., 2006). The latest database release (SITVIT2) contains 111,635 isolates from 169 countries with a total of 9658 patterns (Couvin et al., 2018).

The East-African-Indian (EAI) lineage was defined by a combination of characteristic spoligotype and VNTR patterns and was later encompassed within Lineage 1 (L1) (Sola et al., 2001), as it is considered an "*ancient*" lineage, phylogeographically restricted to East-Africa, South-Asia, South-East-Asia and the Philippines (Palittapongarnpim et al., 2018). L1 is likely to be among the most genetically diverse lineages and some of its sublineages have been associated to spoligotyping-patterns as shown in TB-MINER and SITVIT2, whereas others are not monophyletic (Azé et al., 2015; Couvin et al., 2018; Palittapongarnpim et al., 2018).

The issue of TB origin in Americas is a subject of interest for already many years. In Brazil, TB is likely to have first been imported and secondly spread through centuries of migration from Europe and Africa (Stelzig, 2008). Rio de Janeiro was a main entry port for millions of Africans during the slave trade and a study confirmed that African slaves came along with tuberculosis infections from Africa (Jaeger et al., 2013). A recent study carried out in Pará, Brazil, and elsewhere, has shown the existence of L1 sublineages, which are rare in South America, including one named SIT129 (Conceição et al., 2017; Duarte et al., 2017). Since this signature was previously found in Malawi and Mozambique, it could suggest an East-African origin of these isolates However, the introduction of TB in Northern Brazil could also be related to other less ancient immigration events (Perdigão et al., 2018). To study the genetic diversity of L1 isolates in Pará, we developed a multiplex SNP assay as a first step to a more extended WGS-based characterization of the L1 lineage.

2. Materials and methods

2.1. Strains and study population

The study includes 377 isolates from four collections of L1 isolates: Pará-Brazil (71); RIVM, The Netherlands (102), Madagascar (104); and 100 *in Silico* spoligotypes from WGS obtained from PolyTB: Malawi (64); UK, Midlands (19); The Netherlands (5); Russia and Vietnam (4 each); Uganda (2); Ghana and Tanzania (1 each).

2.2. Spoligotyping

Spoligotyping patterns were determined using a Luminex®200 device (Luminex Corporation, Austin, USA) as previously described (Zhang et al., 2010). PolyTB spoligotypes were constructed *in silico* using SpolPred (Coll et al., 2012).

2.3. SNP-typing

We included eight L1/EAI-specific SNPs as proposed previously. Four other SNPs linked to L1 also discovered using PolyTB, were added (Rv3056, Rv0944, Rv1317c and Rv1629) (Table S1). Dual-priming oligonucleotide (DPO) primers (Chun et al., 2007) were designed using Bioedit v7.0.5. The probes were designed manually, based on Wild-Type (Wt) and Mutant (Mut) alleles for each SNP, with a 5′-Amino-C12-linker moiety (Table S2). In total 24 oligonucleotides (12 Wt and 12 Mut) were coupled to a set of 24 MicroPlex® beads by an Ethyl-Diethyl-Carbodiimide (EDC)-based chemistry (Luminex® Cookbook). For the validation of the PCR, we first performed single-plex experiments for each target and optimized the multiplex assay using a touchdown-PCR principle (Korbie and Mattick, 2008; Mokrousov et al., 2004) (Table S3). We assessed the allelic status of the targeted SNPs on the PCR product by direct hybridization (Abadia et al., 2010). In brief, 2 µl of PCR product was added to a mix of 33 µl (31.5 µl TMAC 1,5X and 1.5 µl of the final 24-plex microspheres mix with a final bead concentration of 35–50 beads/µl concentration per target) and 15 µl Tris-EDTA (TE) pH = 8. The hybridizing conditions were 95 °C (10 min) and 52 °C (20 min) using a thermocycler (BioRad-Alpha-Unit-Block, Mexico). Relative fluorescence Intensities (RFI) were read after SA-PE addition (Streptavidin-Phycoerythrin Lumigrade®, Roche Biochemicals, Pennsberg, Germany) and allelic status inferred. All results are shown in Table S4.

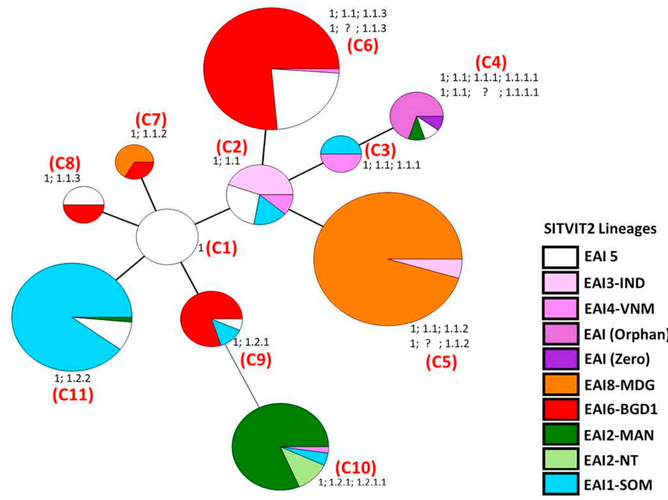

Fig. 1. A Minimum Spanning Tree (MST) based on nine SNPs that allow classification of the L1 lineage isolates from the present study in 11 clusters (C): C1 (n = 15), C2 (n = 18), C3 (n = 10), C5 (n = 107), C6 (n = 89), C7 (n = 3), C8 (n = 4), C9 (n = 15), C10 (n = 43) and C11 (n = 69). Clusters are colored according to spoligotyping-defined lineages.

2.4. Bioinformatical data-analysis

Data were exported from xPONENT® (version 3.1, Luminex Corp, Austin, TX) to Excel® spreadsheets (Microsoft, Redmond, USA) and imported into Bionumerics® v.6.0 (Biomérieux, Applied-Maths, Sint-Martens-Latem, Belgium). A Minimum Spanning Trees (MST) and dendrograms were produced (Figs. 1 and 2).

3. Results and discussion

Among the 12 SNPs detected, only nine were used in the final analysis (Table S1). SNPs for Rv3915 and Rv3056 genes worked only in single-plex, whereas the other targets (Rv0524, Rv2707, Rv2907c, Rv2343c, Rv1326c, Rv3111, Rv3101c) worked well in the multiplex assay.

Based on nine SNPs included in the analysis, 11 clusters were observed (Fig. 1 and Table S4), with most isolates gathered in a SNP-based sublineage being also clustered within one spoligotyping sublineage, showing a good correlation between SNP and spoligotype patterns (Figs. 1 and 2). The largest cluster, C5, gathered EAI8_MDG and EAI3_IND, which share a common ancestor within L1.1.2. Most EAI1_SOM were assigned to cluster C11; however, some were found in C2, C3, C9 and C10. C5 was shown to be linked to C7, and C6 to C8, defining L1.1.2 and L1.1.3, respectively. Clearly EAI1_SOM/SIT48, EAI6_BGD/SIT591 and EAI5/SIT236 (ancestor type) are not monophyletic, a result independently obtained by previous investigators (Palittapongarnpim et al., 2018; Rahim et al., 2007). The formerly

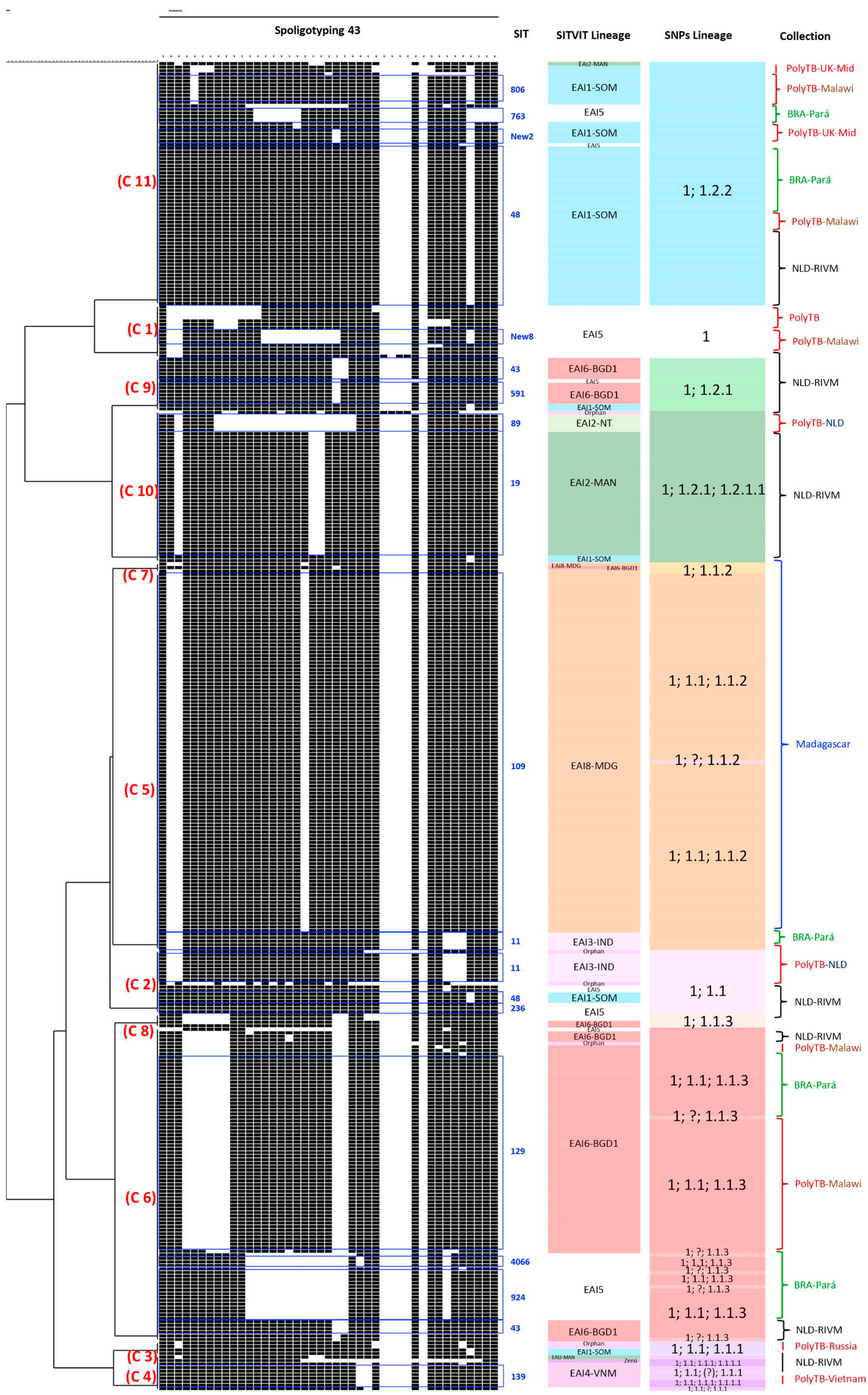

Fig. 2. Bionumerics® based phylogenetic tree computed on SNP-typing in association with spoligotyping and corresponding country of origin and WGS-based lineage names.

defined EAI6_BGD1 was mainly included in C6, and these isolates were derived from Brazil, Malawi and the Netherlands (Table S4) but are also found in South East Asia. A recent WGS-based study found 18 sub-lineages among 480 L1 isolates in Northern Thailand, all included into three major sublineages: i) L1.1 which gathers L1.1.1, L1.1.2 and L1.1.3 (including EAI6_BGD1); ii) L1.2, that includes L1.2.1 (EAI2_PHL and EAI2_NTB) and iii) L1.2.2. This study also unraveled monophyletic SITs from polyphyletic ones (SIT236, SIT48, SIT591) (Palittapongarnpim et al., 2018). Based on the previous knowledge on spoligotypes, SNPs and the latest Thailand WGS study (Coll et al., 2014b; Comas et al., 2009, 2015; Ferdinand et al., 2005; Mallard et al., 2015; Palittapongarnpim et al., 2018; Stavrum et al., 2014; Viegas et al., 2015) we suggest an improved definition of L1 sublineages with at least 6 sublineages (Table S5).

Brazil's population was established through a 300 years slave trade history, in relation to the development of the first sugar cane industry in the Northern and North Eastern regions of Brazil, that lasted from the mid-sixteenth century to the mid-nineteenth century. Slaves were coming from territories known today as Guinea, Angola, Mozambique, Nigeria and more. We suggest that this slave trade has played a major role in shaping today's MTBC L1 population structure observed in North Brazil. Most likely this population structure should be quite different from that found in South Brazil, where a later migration from Europe occurred.

SIT129 is rarely detected in Portugal and Brazil but it is of a great epidemiological significance in two places: (1) the coastal region of Beira and Sofala/Mozambique, a region recently investigated (Anselmo et al. 2019, unpublished results), and (2) in Northern Malawi. Malawi is an English-speaking country that welcomed 250,000 refugees of Mozambican nationality in the 90s, a potential origin of the outbreak in Malawi (Mallard et al., 2015; Toole and Waldman, 1993).

WGS of isolates from Pará and Mozambique is now in progress, to ascertain our scenario, and to compute the estimated divergence time on various isolates harboring SIT129, whose origin, whether purely African or "*back to Africa*" (with an Asian origin) also remains to be unraveled.

4. Conclusions

Based on a 9-SNPs combination analysis, 377 EAI samples were classified into 11 clusters (C1-C11), showing a good correlation between SNP phylogeny and spoligotyping profiles. However, for EAI1-SOM (SIT48), EAI5 (SIT236) and EAI6-BGD1 (SIT591), there is a poorer correlation with SNP data, showing that these signatures cover polyphyletic sublineages. Our results also suggest that EAI8_MDG and EAI3_IND are evolutionary linked and belong to L1.1.2. Additionally, a common spoligotype among isolated from Pará/Brazil, Malawi and Mozambique (SIT129) and common SNPs in the respective isolates from Pará and Malawi suggests a shared evolutionary history, possibly due to transmission during the slave trade from Mozambique-Malawi region in the Northern region of Brazil, and more recently from Mozambique to Malawi, spread by refugees.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2019.06.001.

Acknowledgements

We are thankful to the Brazilian program "Ciência sem Fronteiras" (Science without border) by the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) for investment in this project. M. Jan van Gils, Luminex Corp. Austin, TX is acknowledged for technical support as well as the Beamedex® SAS company (Orsay, France), for coupled-bead supply.

Funding sources

This work was supported by the *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) [207,071/2014-4]; by the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brazil* (CAPES) [Finance Code 001], by the Centre National de la Recherche Scientifique through the financial support of the Infection Genetics Emerging Pathogen Evolution (IGEPE) research team and by the Beamedex® SAS company.

Authors' contributions

Conceived the project: ECC, HGM, CS, PNS, KVBL and RSD; designed the experiments: CS, ECC, GR and HGM; provided samples and informations: MLP, KVBL, FC, NHR, VRR, DS, LR, SG and VRB; designed probes and primers: XOD and ECC; provided reagents and equipments: CS and GR; performed experiments: ECC and HGM; performed data analysis: ECC, CS, GR and CS; supervising: CS, GR, RSD and KVBL; wrote the manuscript: ECC, CS and GR. All authors have revised the manuscript.

References

Abadia, E., Zhang, J., dos Vultos, T., Ritacco, V., Kremer, K., Aktas, E., Matsumoto, T., Refregier, G., van Soolingen, D., Gicquel, B., Sola, C., 2010. Resolving lineage assignation on *Mycobacterium tuberculosis* clinical isolates classified by spoligotyping with a new high-throughput 3R SNPs based method. Infect. Genet. Evol. 10, 1066–1074. https://doi.org/10.1016/j.meegid.2010.07.006.

Azé, J., Sola, C., Zhang, J., Lafosse-Marin, F., Yasmin, M., Siddiqui, R., Kremer, K., Van Soolingen, D., Refrégier, G., 2015. Genomics and machine learning for taxonomy consensus: the *Mycobacterium tuberculosis* complex paradigm. PLoS One. https://doi.org/10.1371/journal.pone.0130912.

Brudey, K., Driscoll, J.R., Rigouts, L., Prodinger, W.M., Gori, A., Al-Hajoj, S. a, Allix, C., Aristimuño, L., Arora, J., Baumanis, V., Binder, L., Cafrune, P., Cataldi, A., Cheong, S., Diel, R., Ellermeier, C., Evans, J.T., Fauville-Dufaux, M., Ferdinand, S., Garcia de Viedma, D., Garzelli, C., Gazzola, L., Gomes, H.M., Guttierez, M.C., Hawkey, P.M., van Helden, P.D., Kadival, G.V., Kreiswirth, B.N., Kremer, K., Kubin, M., Kulkarni, S.P., Liens, B., Lillebaek, T., Ho, M.L., Martin, C., Martin, C., Mokrousov, I., Narvskaïa, O., Ngeow, Y.F., Naumann, L., Niemann, S., Parwati, I., Rahim, Z., Rasolofo-Razanamparany, V., Rasolonavalona, T., Rossetti, M.L., Rüsch-Gerdes, S., Sajduda, A., Samper, S., Shemyakin, I.G., Singh, U.B., Somoskovi, A., Skuce, R. a, van Soolingen, D., Streicher, E.M., Suffys, P.N., Tortoli, E., Tracevska, T., Vincent, V., Victor, T.C., Warren, R.M., Yap, S.F., Zaman, K., Portaels, F., Rastogi, N., Sola, C., 2006. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol. 6, 23. https://doi.org/10.1186/1471-2180-6-23.

Chun, J.-Y., Kim, K.-J., Hwang, I.-T., Kim, Y.-J., Lee, D.-H., Lee, I.-K., Kim, J.-K., 2007. Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. Nucleic Acids Res. 35, e40. https://doi.org/10.1093/nar/gkm051.

Coll, F., Mallard, K., Preston, M.D., Bentley, S., Parkhill, J., McNerney, R., Martin, N., Clark, T.G., 2012. SpolPred: rapid and accurate prediction of *Mycobacterium tuberculosis* spoligotypes from short genomic sequences. Bioinformatics. https://doi.org/10.1093/bioinformatics/bts544.

Coll, F., McNerney, R., Guerra-Assunção, J.A., Glynn, J.R., Perdigão, J., Viveiros, M., Portugal, I., Pain, A., Martin, N., Clark, T.G., 2014a. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. Under Rev. 5, 4812. https://doi.org/10.1038/ncomms5812.

Coll, F., Preston, M., Guerra-Assunção, J.A., Hill-Cawthorn, G., Harris, D., Perdigão, J., Viveiros, M., Portugal, I., Drobniewski, F., Gagneux, S., Glynn, J.R., Pain, A., Parkhill, J., McNerney, R., Martin, N., Clark, T.G., 2014b. PolyTB: a genomic variation map for *Mycobacterium tuberculosis*. Tuberculosis (Edinb) 94, 346–354. https://doi.org/10.1016/j.tube.2014.02.005.

Comas, I., Homolka, S., Niemann, S., Gagneux, S., 2009. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. PLoS One 4. https://doi.org/10.1371/journal.pone.0007815.

Comas, I., Hailu, E., Kiros, T., Bekele, S., Mekonnen, W., Gumi, B., Tschopp, R., Ameni, G., Hewinson, R.G., Robertson, B.D., Goig, G.A., Stucki, D., Gagneux, S., Aseffa, A., Young, D., Berg, S., 2015. Population genomics of *Mycobacterium tuberculosis* in Ethiopia contradicts the virgin soil hypothesis for human tuberculosis in sub-Saharan Africa. Curr. Biol. https://doi.org/10.1016/j.cub.2015.10.061.

Conceição, E.C., Rastogi, N., Couvin, D., Lopes, M.L., Furlaneto, I.P., Gomes, H.M., Vasconcellos, S.E.G., Suffys, P.N., Schneider, M.P.C., de Sousa, M.S., Sola, C., de Paula Souza e Guimarães, R.J., Duarte, R.S., Batista Lima, K.V., 2017. Genetic diversity of *Mycobacterium tuberculosis* from Pará, Brazil, reveals a higher frequency of ancestral strains than previously reported in South America. Infect. Genet. Evol. 56. https://doi.org/10.1016/j.meegid.2017.10.021.

Coscolla, M., Gagneux, S., 2014. Consequences of genomic diversity in *Mycobacterium* *tuberculosis*. Semin. Immunol. https://doi.org/10.1016/j.smim.2014.09.012.

Couvin, D., David, A., Zozio, T., Rastogi, N., 2018. Macro-geographical specificities of the prevailing tuberculosis epidemic as seen through SITVIT2, an updated version of the *Mycobacterium tuberculosis* genotyping database. Infect. Genet. Evol. https://doi.org/10.1016/J.MEEGID.2018.12.030.

Duarte, T.A., Nery, J.S., Boechat, N., Pereira, S.M., Simonsen, V., Oliveira, M., Gomes, M.G.M., Penha-Gonçalves, C., Barreto, M.L., Barbosa, T., 2017. A systematic review of east African-Indian family of *Mycobacterium tuberculosis* in Brazil. Brazil. J. Infect. Dis. 21, 317–324. https://doi.org/10.1016/j.bjid.2017.01.005.

Ferdinand, S., Sola, C., Chanteau, S., Ramarokoto, H., Rasolonavalona, T., Rasolofo-Razanamparany, V., Rastogi, N., 2005. A study of spoligotyping-defined *Mycobacterium tuberculosis* clades in relation to the origin of peopling and the demographic history in Madagascar. Infect. Genet. Evol. 5, 340–348. https://doi.org/10.1016/j.meegid.2004.10.002.

Jaeger, L.H., de Souza, S.M.F.M., Dias, O.F., Iñiguez, A.M., 2013. *Mycobacterium tuberculosis* complex in remains of 18th–19th century slaves, Brazil. Emerg. Infect. Dis. https://doi.org/10.3201/eid1905.120193.

Korbie, D.J., Mattick, J.S., 2008. Touchdown PCR for increased specificity and sensitivity in PCR amplification. Nat. Protoc. 3, 13–15. https://doi.org/10.1038/nprot.2008.133.

Mallard, K., Khan, P., Chiwaya, A., McNerney, R., Parkhill, J., Clark, T., Houben, R., Pereira, R., Guerra-Assunção, J., Mzembe, T., Coll, F., Banda, L., Fine, P., Glynn, J., Crampin, A., 2015. Large-scale whole genome sequencing of *M. tuberculosis* provides insights into transmission in a high prevalence area. Elife. https://doi.org/10.7554/elife.05166.

Mokrousov, I., Bhanu, N.V., Suffys, P.N., Kadival, G.V., Yap, S.F., Cho, S.N., Jordaan, A.M., Narvskaya, O., Singh, U.B., Gomes, H.M., Lee, H., Kulkarni, S.P., Lim, K.C., Khan, B.K., Van Soolingen, D., Victor, T.C., Schouls, L.M., 2004. Multicenter evaluation of reverse line blot assay for detection of drug resistance in *Mycobacterium tuberculosis* clinical isolates. J. Microbiol. Methods 57, 323–335. https://doi.org/10.1016/j.mimet.2004.02.006.

Palittapongarnpim, P., Ajawatanawong, P., Viratyosin, W., Smittipat, N., Disratthakit, A., Mahasirimongkol, S., Yanai, H., Yamada, N., Nedsuwan, S., Imasanguan, W., Kantipong, P., Chaiyasirinroje, B., Wongyai, J., Toyo-oka, L., Phelan, J., Parkhill, J., Clark, T.G., Hibberd, M.L., Ruengchai, W., Palittapongarnpim, P., Juthayothin, T., Tongsima, S., Tokunaga, K., 2018. Evidence for host-bacterial co-evolution via genome sequence analysis of 480 Thai *Mycobacterium tuberculosis* lineage 1 isolates. Sci. Rep. https://doi.org/10.1038/s41598-018-29986-3.

Perdigão, J., Silva, C., Diniz, J., Pereira, C., Machado, D., Ramos, J., Silva, H., Abilleira, F., Brum, C., Reis, A.J., Macedo, M., Scaini, J.L., Silva, A.B., Esteves, L., Macedo, R., Maltez, F., Clemente, S., Coelho, E., Viegas, S., Rabna, P., Rodrigues, A., Taveira, N., Jordao, L., Kritski, A., Lapa e Silva, J.R., Mokrousov, I., Couvin, D., Rastogi, N., Couto, I., Pain, A., McNerney, R., Clark, T.G., von Groll, A., Dalla-Costa, E.R., Rossetti, M.L., Silva, P.E.A., Viveiros, M., Portugal, I., 2018. Clonal expansion across the seas as seen through CPLP-TB database: a joint effort in cataloguing *Mycobacterium tuberculosis* genetic diversity in Portuguese-speaking countries. Infect. Genet. Evol. https://doi.org/10.1016/j.meegid.2018.03.011.

Rahim, Z., Zaman, K., van der Zanden, A.G.M., Möllers, M.J., van Soolingen, D., Raqib, R., Zaman, K., Begum, V., Rigouts, L., Portaels, F., Rastogi, N., Sola, C., 2007. Assessment of population structure and major circulating phylogeographical clades of *Mycobacterium tuberculosis* complex in Bangladesh suggests a high prevalence of a specific subclade of ancient *M. tuberculosis* genotypes. J. Clin. Microbiol. 45, 3791–3794. https://doi.org/10.1128/JCM.01247-07.

Sola, C., Filliol, I., Legrand, E., Mokrousov, I., Rastogi, N., 2001. *Mycobacterium tuberculosis* phylogeny reconstruction based on combined numerical analysis with IS1081, IS6110, VNTR, and DR-based spoligotyping suggests the existence of two new phylogeographical clades. J. Mol. Evol. 53, 680–689. https://doi.org/10.1007/s002390010255.

Stavrum, R., PrayGod, G., Range, N., Faurholt-Jepsen, D., Jeremiah, K., Faurholt-Jepsen, M., Krarup, H., Aabye, M.G., Changalucha, J., Friis, H., Andersen, A.B., Grewal, H.M.S., 2014. Increased level of acute phase reactants in patients infected with modern *Mycobacterium tuberculosis* genotypes in Mwanza, Tanzania. BMC Infect. Dis. https://doi.org/10.1186/1471-2334-14-309.

Stelzig, S., 2008. Country Profile. (Brazil, Hamburg).

Toole, M.J., Waldman, R.J., 1993. Refugees and displaced persons. War, hunger, and public health. JAMA 270, 600–605.

Viegas, S.O., Ghebremichael, S., Massawo, L., Alberto, M., Fernandes, F.C., Monteiro, E., Couvin, D., Matavele, J.M., Rastogi, N., Correia-Neves, M., Machado, A., Carrilho, C., Groenheit, R., Källenius, G., Koivula, T., 2015. *Mycobacterium tuberculosis* causing tuberculous lymphadenitis in Maputo, Mozambique. BMC Microbiol. 15, 268. https://doi.org/10.1186/s12866-015-0603-5.

Zhang, J., Abadia, E., Refregier, G., Tafaj, S., Boschiroli, M.L., Guillard, B., Andremont, A., Ruimy, R., Sola, C., 2010. *Mycobacterium tuberculosis* complex CRISPR genotyping: improving efficiency, throughput and discriminative power of "spoligotyping" with new spacers and a microbead-based hybridization assay. J. Med. Microbiol. 59, 285–294. https://doi.org/10.1099/jmm.0.016949-0.