HHS Public Access

Author manuscript

Tuberculosis (Edinb). Author manuscript; available in PMC 2015 December 08.

Published in final edited form as:

Tuberculosis (Edinb). 2015 December ; 95(6): 810–816. doi:10.1016/j.tube.2015.09.001.

Whole genome sequencing identifies circulating Beijing-lineage *Mycobacterium tuberculosis* strains in Guatemala and an associated urban outbreak

Joseph W. Saelens[a,1], **Dalia Lau-Bonilla**[b,1], **Anneliese Moller**[b], **Narda Medina**[b], **Brenda Guzmán**[b], **Maylena Calderón**[b], **Raúl Herrera**[b], **Dana M. Sisk**[a], **Ana M. Xet-Mull**[a], **Jason E. Stout**[c], **Eduardo Arathoon**[d,e], **Blanca Samayoa**[b,f,**], and **David M. Tobin**[a,*]

[a]Department of Molecular Genetics and Microbiology and Center for Host-Microbial Interactions, Duke University Medical Center, Durham, NC, USA

[b]Asociacion de Salud Integral, Guatemala City, Guatemala

[c]Department of Medicine, Division of Infectious Diseases and International Health, Duke University Medical Center, Durham, NC, USA

[d]Hospital General San Juan de Dios, Guatemala City, Guatemala

[e]Clínica Familiar Luis Angel García, Guatemala City, Guatemala

[f]Facultad de Ciencias Químicas y Farmacia, Universidad de San Carlos de Guatemala, Guatemala City, Guatemala

Summary

Limited data are available regarding the molecular epidemiology of *Mycobacterium tuberculosis* (*Mtb*) strains circulating in Guatemala. Beijing-lineage *Mtb* strains have gained prevalence worldwide and are associated with increased virulence and drug resistance, but there have been only a few cases reported in Central America. Here we report the first whole genome sequencing of Central American Beijing-lineage strains of *Mtb*. We find that multiple Beijing-lineage strains, derived from independent founding events, are currently circulating in Guatemala, but overall still represent a relatively small proportion of disease burden. Finally, we identify a specific Beijing-lineage outbreak centered on a poor neighborhood in Guatemala City.

Keywords

Tuberculosis; Molecular epidemiology; Whole genome sequencing; Beijing strains; Guatemala; Central America

*Corresponding author. david.tobin@duke.edu. **Corresponding author. investigacion@asi.org.gt.

[1]These authors contributed equally.

Competing interests: The authors report no conflicts of interest.

Ethical approval: This work was approved by the Institutional Ethics Committee of the Hospital San Juan de Dios (Guatemala). Mycobacterial culture, DNA extraction and spoligo-typing were performed in Guatemala, and,epidemiologic and patient information were maintained in Guatemala. Genome sequencing and analysis of deidentified samples were performed at Duke.

1. Introduction

Tuberculosis (TB) poses a major challenge to global health. The causative agent, *Mycobacterium tuberculosis* (*Mtb*), was responsible for around 1.5 million deaths in 2013 [1]. The seven major lineages of *Mtb* have historically been associated with specific geographical regions, and substantial evidence exists that genetic differences between lineages influence disease presentation and outcome [2–6]. Beijing lineage strains, also known as Lineage 2 strains, have emerged as important drivers of global *Mtb* burden [7–9]. Notably, outside of East Asia, "modern" Beijing strains account for the majority of this burden [10]. These "modern" Beijing strains display elevated rates of drug-resistance, rapid progression of disease and increased transmission [11–16]. While there is still debate as to precisely when the Beijing lineage originated, it appears to have arisen in East Asia and, through successive waves of human migration, spread throughout this region and beyond [10,17]. Positive selection on virulence-associated genes as well as compensatory mutations negating the fitness-cost of drug-resistance are key factors in making Beijing the most successful contemporary lineage [17–20].

The emergence of "modern" Beijing strains to new regions has been particularly pronounced in Eastern Europe and Africa, regions in which Euro-American strains are also present [21–23]. South Africa serves as one striking example, where data suggest that Beijing strains have continually expanded since their arrival to account for an increasing proportion of total burden [20]. Beijing strains in South Africa have developed resistance to standard drug therapies and contributed to poor outcomes for patients [24,25].

In the Western Hemisphere, the emergence of Beijing strains has thus far been more complex. Beijing strains have spread across the United States, with notable outbreaks occurring in New York City, NY, and Houston, TX [26–28]. In contrast, South and Central America have heretofore been thought to be relatively isolated from this trend, and Euro-American/Lineage 4 strains predominate [29–36]. Current evidence suggests Beijing strains have failed to gain a foothold in South America outside of Peru, where a substantial number of Chinese and Japanese immigrants settled throughout the 19th and 20th centuries [30]. However, there is very little information on *Mtb* strain diversity in Central America. Thus far, the sparse evidence that does exist indicates that Beijing strains contribute only minimally to the burden of tuberculosis in the region [17,37].

Here we investigate the status of Beijing-lineage strains in Guatemala, the most populous country in Central America. We take a whole genome sequencing (WGS) approach to assess recent TB cases at an HIV clinic and associated public hospital. We combine our results with global genome sequences as well as epidemiologic patient data to assess patterns of Beijing-lineage *Mtb* in Guatemala.

2. Materials and methods

2.1. Patient population

The Clínica Familiar "Luis Angel García" (CFLAG) is an HIV-specialized clinic located at the Hospital General San Juan de Dios, one of Guatemala's two public teaching hospitals in

Guatemala City. It provides comprehensive treatment and follow-up for people living with HIV for both outpatient and inpatient care and has provided medical care for more than 10,000 patients over the last thirty years. From 2010 to 2014, routine spoligotyping was performed at CFLAG on all *Mtb* isolates from CFLAG and the associated hospital during that period, derived from a total of 514 independent patients. Spoligotyping identified 11 Beijing-lineage strains; five of the 11 strains could be regrown for WGS.

2.2. DNA extraction and genotyping

Bacterial culturing, spoligotyping, and DNA extraction were performed according to standard practices (detailed in Supplemental Material) [38,39].

2.3. Sequencing, alignment and SNP calling

Using Illumina HiSeq 2500 and Illumina MiSeq platforms, we sequenced each strain at >500-fold coverage with a minimum read length of 50 base pairs. Additional sequence reads acquired from published WGS datasets also had minimum read lengths of 50 base pairs. We downloaded reads from the National Center for Biotechnology Information Sequence Read Archive and the European Nucleotide Archive, as described in [2,3,17,40]. We collected sequence reads for 24 globally extant strains of *Mtb* representing the seven major lineages [2,3,41], as well as 96 globally extant Beijing strains [17]. We aligned reads against the H37Rv reference genome (GenBank: AL123456.3) using BWA [42]. We called variants using SAMtools and filtered with VarScan for a minimum read depth of 10, a consensus quality score of 20, and a minimum variant frequency of 0.75 [43,44]. We discarded SNPs adjacent to indels and within repetitive regions of the genome. Additionally, as is standard practice, we discarded SNPs affecting genes commonly associated with drug-resistance to avoid homoplastic mutations among distantly related strains [10,17,45]. We visualized variants among the Guatemalan isolates associated with the outbreak with Circular Visualization for Microbial Genomes (URL: http://civi.cmbi.ru.nl/).

2.4. Nucleotide sequence accession numbers

The sequences determined in this study have been deposited in the NCBI Sequence Read Archive under accession numbers: GG-135-10 – **SRR1765871**, GG-120-10 – **SRR1765872**, GG-152-12 – **SRR1765877**, GG-131-11 – **SRR1765874**, GG-219-11 – **SRR1765879**, and GG-30-13 – **SRR1765888**. GG-219-11 and GG-30-13 were sequential isolates taken from the same patient 17 months apart.

2.5. Tree builds

All trees were based on genome-wide SNPs derived according to the parameters specified above. We constructed a superset of SNPs for each strain with reference alleles occupying sites for which no variants were detected using custom Perl scripts. These SNPs informed neighbor-joining and maximum-likelihood methods of phylogeny construction. We implemented neighbor-joining methodology with ClustalW2 using pairwise similarity scores of SNP supersets as a measure of genetic distance [46]. We generated a maximum-likelihood phylogeny with RAxML using a GTR model of nucleotide substitution [47]. For each method, 1000 bootstrap replicates provided support for nodes on the tree. The

phylogenies derived from each method were congruent (compare Figure 2 to Figure S1). Trees were visualized with FigTree (see URL tree.bio.ed.ac.uk/software/figgtree).

3. Results

3.1. Spoligotyping of Mtb strains in Guatemala city

Euro-American strains of *Mtb* have historically been thought to prevail in Central America, with very little contribution from Beijing strains. In a collection of 514 patient isolates collected from 2010 to 2014 in by the Hospital San Juan de Dios and CFLAG, we identified 11 (2.1%) with Beijing-lineage spoligotypes. Isolates during this time period came from both HIV-positive and HIV-negative patients. Patients with Beijing-lineage strains included individuals living in the capital, Guatemala City, as well as two patients from rural departments (Table 1).

Because no Beijing-lineage cases had been sequenced from Guatemala, we sought to further investigate the identity of these strains by WGS. Beyond identifying strain background, the resolution of WGS allows analysis of microevolution among closely related isolates, and offers insight into the transmission history of related strains [48]. We attempted to culture isolates retrospectively from all 11 patients, but only five of these isolates could be regrown in sufficient quantities for WGS (Table 1).

3.2. Phylogeny of Guatemalan Beijing-lineage strains

We used a sample of 24 globally extant strains of *Mtb*, representative of the seven major lineages, to initially place the Guatemalan Beijing strains in a broad phylogenetic context [3,40]. Neighbor-joining and maximum-likelihood methods of phylogeny construction based on 18,039 genome-wide SNPs resolved all 29 strains into the seven principal lineages of *Mtb*, and placed the Guatemalan isolates among known Beijing strains (see Methods) (Figure 1). To independently confirm the identity of these strains as Beijing-lineage, we searched for previously identified Beijing-specific variants. All five Guatemalan isolates harbored the Beijing lineage ACG- > AGG SNP in codon 20 of Rv2450c and the GGG-> AGG SNP in codon 176 of Rv2952 (Supplemental Table 1), con-firming their status as Beijing strains [41].

To understand the relationship of the Guatemalan Beijing strains to other globally extant Beijing strains, we generated phylogenetic trees based on 6584 SNPs among a total of 101 Beijing strains (see Methods) [17]. The Guatemalan strains were distributed on three independent branches in all constructed trees (Figure 2 and Figure S1). Three isolates (GG-131-11, GG-219-11, and GG-152-12) clustered on a single branch of the tree near strains previously identified in China and Thailand. A fourth isolate, GG-120-10, fell on an independent branch of the phylogenetic tree, and is most closely related to strains identified in Nepal. Finally, one of the Guatemalan patient isolates, GG-135-10, appeared to represent a more basal sublineage, and clustered with the so-called “ancient” Beijing strains. Notably, none of the isolates came from individuals who reported travel outside of Guatemala.

We next assessed by SNP analysis whether the phylogenetic distance was consistent with multiple introductions of Beijing strains into the country. Approximately 500 SNPs

separated GG-135-10 from the four other Guatemalan isolates, and GG-120-10 differed from GG-131-11, GG-219-11, and GG-152-12 by ~200 SNPs (Supplemental Table 1). The last three isolates were highly similar, differing by at most 31 SNPs, suggestive of a recently shared progenitor strain (detailed further below).

Phylogenetic analysis indicated GG-120-10, GG-131-11, GG-219-11 and GG-152-12 are evolutionary "modern" Beijing strains, while GG-135-10 is of an ancestral Beijing sublineage. Thus, we assessed each Guatemalan isolate for the insertion variant T1406760 + G in *Rv1258c*, as this variant is known to be a marker for the "modern" Beijing genotype [49]. As expected, four isolates displayed this polymorphism: GG-120-10, GG-131-11, GG-219-11, and GG-152-12. As additional confirmation, the four "modern" Beijing strains also harbored a GGA- > CGA mutation at codon position 58 in the DNA repair enzyme gene *mutT2*, yet GG-135-10 did not [8,50]. Together, these data reveal that both evolutionarily "modern" and "ancient" Beijing strains are present in Guatemala.

3.3. Epidemiologic analysis of a Beijing-lineage outbreak in Guatemala city

We noted that three Guatemalan Beijing-lineage isolates (GG-131-11, GG-219-11, and GG-152-12) demonstrated a tight clustering on one branch of the tree, indicating transmission of a single founder strain (Figure 2). 125 SNPs separate this clade from its nearest neighbor on the tree (8662-05) (Figure 3). Of these, 94 SNPs are shared among all three strains. 18 SNPs are unique to GG-131-11, and 13 unique to GG-219-11 and GG-152-12. The patients from whom GG-219-11 and GG-152-12 were isolated had spent time within the same city jail, but their incarceration periods had not overlapped. Isolates from these two patients were the most closely related strains we identified, varying from each other by only 1 SNP. One variant shared between these two strains includes a GAC- > AAC mutation in codon 94 of the DNA gyrase gene *gyrA*, a mutation reported to associate with fluoroquinolone-resistance [51,52].

The closely related strain GG-131-11 differed from the jail isolates by 31 SNPs, and did not contain the *gyrA* variant. This patient lived in the same neighborhood as the jail where the patients had resided, but the third patient had no other known connection other than residence in the vicinity of the jail. Among the six strains identified as Beijing, but for which WGS could not be performed, two of them derived from HIV-negative patients who also lived in the same neighborhood as the jail (Table 1). In total, six of the 11 identified Beijing-lineage cases were centered on this same neighborhood.

Finally, the HIV-positive patient (Table 1) from whom strain GG-219-11 was isolated abandoned treatment completely and then returned to the clinic 17 months later, at which time a second isolate (GG-30-13) was obtained. A comparison of the sequential isolates revealed a single variant between the initial isolate and that taken after 17 months of untreated disease. These data reflect a relatively low mutation rate for this particular Beijing-lineage strain in the absence of antibiotic therapies and in the context of a compromised immune system.

4. Discussion

In this study, we use WGS to investigate Beijing *Mtb* strains circulating in Guatemala. The Beijing lineage, which has emerged globally, is considered generally more virulent than strains of other lineages, and has a higher frequency of drug-resistance [11–16]. Prior to this study, information on Beijing isolates in Guatemala has been sparse [5]. Here, we have for the first time conducted WGS on Beijing isolates from Central America for an in-depth characterization of the potential regional emergence of these strains.

We infer from WGS and phylogenetic analysis that multiple Beijing strains have been introduced into Guatemala and that these strains have transmitted within Guatemala itself; notably, none of the affected patients has any history of travel outside Guatemala, strongly suggesting in-country circulation of each strain. Patients infected with Beijing-lineage strains included both HIV-negative and HIV-positive individuals. Central and South America have traditionally been thought to harbor almost exclusively Euro-American *Mtb* strains [29–36]. Here we confirm the presence of circulating Beijing strains in Central America and link one set of strains to an outbreak in an urban setting.

We were able to connect patients through a combination of strain sequences and epidemiology. We identified an outbreak of one Beijing-lineage strain in a poor neighborhood in urban Guatemala centered on a city jail. Two of the cases were patients who had resided in the jail but not during overlapping time periods, suggesting the presence of other latently or undiagnosed actively infected inmates. Indeed, one of the unsequenced but spoligotype-identified Beijing strains (Table 1, strain GG-10-14) was from a patient who had spent time in the same jail in 2008. The Beijing-lineage outbreak strain does not represent the only *Mtb* strain within the jail. There were eight additional TB cases associated with the jail that, by spoligotyping, were identified as Euro-American/Lineage 4 strains (T1, X1, H2, LAM3, LAM9 families) (D.L.-B., N.M., E.A. and B.S., unpublished results).

Of the 11 Beijing isolates we identified, six were associated with a local jail or the neighborhood directly around it. Although we were unable to identify direct epidemiologic links between the neighborhood and jail-associated strains, the geographical and phylogenetic clustering is striking. Workers in the jail (guards and other staff), exiting inmates, and proximity to public bus lines within the neighborhood are possible links between the community and the jail. Consistent with previous results on outbreak strains of *Mtb* and rates of divergence, the clustering of these cases as well as substantial nucleotide divergence (31 SNPs) between the neighborhood and jail-associated strains suggests an outbreak that has lasted a number of years and the presence of other unidentified patients in the area [53–56]. The mutation rate of *Mtb* is estimated to be between 0.3 and 0.5 SNPs per year, and it has been proposed that a maximum of five SNPs separating isolates indicate ongoing transmission [54–56]. However, one study demonstrated an epidemiological link between strains that differed by 14 SNPs, and bursts of mutations have been shown to occur in clinical isolates of *Mtb* [57,58]. Together, these results suggest the founder strain arrived in Guatemala City and infected a citizen who entered the jail (either an employee or inmate), thereby establishing clonal transmission of this isolate within the jail's population. The

strain's entry into the jail likely initiated its divergence from related isolates circulating in the broader community.

In spite of no known anti-tuberculosis therapy for these patients prior to diagnosis, we identified a mutation that arose in the isolates and was associated with resistance to fluoroquinolones [51]. This mutation may have arisen in a progenitor strain due to the use of over-the-counter antibiotics, readily available in Guatemala, or through prescription from another practitioner. Regardless, this finding highlights the issue of emergent fluoroquinolone-resistant Beijing-lineage strains in countries with easy over-the-counter access to antibiotics.

Overall, we describe multiple circulating Beijing strains that have been introduced independently into Guatemala, including an “ancient” Beijing strain. However, the overall contribution to disease burden in Guatemala City still appears relatively low, with 11 Beijing isolates out of 514 reported cases (2.1%), in line with reports from other Central and South American countries [29–37]. Guatemalan immigrants to the United States with active TB show similar rates of infection with Beijing strains (2% of 609 genotyped cases reported 2009–2013, as compared with 15% prevalence of Beijing strains among all TB cases reported in the US in the same period) (Smita Ghosh, Centers for Disease Control and Prevention, personal communication).

The emergence of the Beijing lineage in Guatemala has important implications for public health in Central America, where no WGS on *Mtb* had been performed previously. More broadly, this work helps increase our understanding of the global spread of the Beijing lineage, which presents particular challenges in combating TB.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Smita Ghosh (C.D.C.) for sharing U.S. strain frequency data, Kim Dohlich and Tom Mitchell for comments on the manuscript, the personnel of the Clínica Familiar Luis Ángel García for their assistance in this study and the Instituto de Salud Carlos III (Madrid, Spain) for its support.

Funding: This work was supported by the Asociación de Salud Integral (Guatemala), a National Science Foundation Graduate Research Fellowship (J.W.S.), and a Whitehead Scholar Award (D.M.T.).

References

1. WHO. Global tuberculosis report 2014. Geneva, Switzerland: WHO; 2014.
2. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, Parkhill J, Malla B, Berg S, Thwaites G, Yeboah-Manu D, Bothamley G, Mei J, Wei L, Bentley S, Harris SR, Niemann S, Diel R, Aseffa A, Gao Q, Young D, Gagneux S. Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. Nat Genet. 2013; 45:1176–82. [PubMed: 23995134]
3. Firdessa R, Berg S, Hailu E, Schelling E, Gumi B, Erenso G, Gadisa E, Kiros T, Habtamu M, Hussein J, Zinsstag J, Robertson BD, Ameni G, Lohan AJ, Loftus B, Comas I, Gagneux S, Tschopp R, Yamuah L, Hewinson G, Gordon SV, Young DB, Aseffa A. Mycobacterial lineages causing

pulmonary and extrapulmonary tuberculosis, Ethiopia. Emerg Infect Dis. 2013; 19:460–3. [PubMed: 23622814]

4. Hershberg R, Lipatov M, Small PM, Sheffer H, Niemann S, Homolka S, Roach JC, Kremer K, Petrov DA, Feldman MW, Gagneux S. High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography. PLoS Biol. 2008; 6:e311. [PubMed: 19090620]
5. Mathema B, Kurepina N, Yang G, Shashkina E, Manca C, Mehaffy C, Bielefeldt-Ohmann H, Ahuja S, Fallows DA, Izzo A, Bifani P, Dobos K, Kaplan G, Kreiswirth BN. Epidemiologic consequences of microvariation in *Mycobacterium tuberculosis*. J Infect Dis. 2012; 205:964–74. [PubMed: 22315279]
6. Nicol MP, Wilkinson RJ. The clinical consequences of strain diversity in *Mycobacterium tuberculosis*. Trans R Soc Trop Med Hyg. 2008; 102:955–65. [PubMed: 18513773]
7. Caminero JA, Pena MJ, Campos-Herrero MI, Rodriguez JC, Garcia I, Cabrera P, Lafoz C, Samper S, Takiff H, Afonso O, Pavon JM, Torres MJ, van Soolingen D, Enarson DA, Martin C. Epidemiological evidence of the spread of a *Mycobacterium tuberculosis* strain of the Beijing genotype on Gran Canaria Island. Am J Respir Crit Care Med. 2001; 164:1165–70. [PubMed: 11673204]
8. Bifani PJ, Mathema B, Kurepina NE, Kreiswirth BN. Global dissemination of the *Mycobacterium tuberculosis* W-Beijing family strains. Trends Microbiol. 2002; 10:45–52. [PubMed: 11755085]
9. Cowley D, Govender D, February B, Wolfe M, Steyn L, Evans J, Wilkinson RJ, Nicol MP. Recent and rapid emergence of W-Beijing strains of *Mycobacterium tuberculosis* in Cape Town, South Africa. Clin Infect Dis. 2008; 47:1252–9. [PubMed: 18834315]
10. Luo T, Comas I, Luo D, Lu B, Wu J, Wei L, Yang C, Liu Q, Gan M, Sun G, Shen X, Liu F, Gagneux S, Mei J, Lan R, Wan K, Gao Q. Southern East Asian origin and coexpansion of *Mycobacterium tuberculosis* Beijing family with Han Chinese. Proc Natl Acad Sci U S A. 2015; 112(26):8136–41. [PubMed: 26080405]
11. de Jong BC, Hill PC, Aiken A, Awine T, Antonio M, Adetifa IM, Jackson-Sillah DJ, Fox A, Deriemer K, Gagneux S, Borgdorff MW, McAdam KP, Corrah T, Small PM, Adegbola RA. Progression to active tuberculosis, but not transmission, varies by *Mycobacterium tuberculosis* lineage in the Gambia. J Infect Dis. 2008; 198:1037–43. [PubMed: 18702608]
12. Hanekom M, Gey van Pittius NC, McEvoy C, Victor TC, Van Helden PD, Warren RM. *Mycobacterium tuberculosis* Beijing genotype: a template for success. Tuberc Edinb. 2011; 91:510–23.
13. Hanekom M, van der Spuy GD, Streicher E, Ndabambi SL, McEvoy CR, Kidd M, Beyers N, Victor TC, van Helden PD, Warren RM. A recently evolved sublineage of the *Mycobacterium tuberculosis* Beijing strain family is associated with an increased ability to spread and cause disease. J Clin Microbiol. 2007; 45:1483–90. [PubMed: 17360841]
14. Kato-Maeda M, Shanley CA, Ackart D, Jarlsberg LG, Shang S, Obregon-Henao A, Harton M, Basaraba RJ, Henao-Tamayo M, Barrozo JC, Rose J, Kawamura LM, Coscolla M, Fofanov VY, Koshinsky H, Gagneux S, Hopewell PC, Ordway DJ, Orme IM. Beijing sublineages of *Mycobacterium tuberculosis* differ in pathogenicity in the guinea pig. Clin Vaccine Immunol. 2012; 19:1227–37. [PubMed: 22718126]
15. Parwati I, van Crevel R, van Soolingen D. Possible underlying mechanisms for successful emergence of the *Mycobacterium tuberculosis* Beijing genotype strains. Lancet Infect Dis. 2010; 10:103–11. [PubMed: 20113979]
16. Ribeiro SC, Gomes LL, Amaral EP, Andrade MR, Almeida FM, Rezende AL, Lanes VR, Carvalho EC, Suffys PN, Mokrousov I, Lasunskaia EB. *Mycobacterium tuberculosis* strains of the modern sublineage of the Beijing family are more likely to display increased virulence than strains of the ancient sublineage. J Clin Microbiol. 2014; 52:2615–24. [PubMed: 24829250]
17. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, Blum MG, Rusch-Gerdes S, Mokrousov I, Aleksic E, Allix-Beguec C, Antierens A, Augus-tynowicz-Kopec E, Ballif M, Barletta F, Beck HP, Barry CE 3rd, Bonnet M, Borroni E, Campos-Herrero I, Cirillo D, Cox H, Crowe S, Crudu V, Diel R, Drobniewski F, Fauville-Dufaux M, Gagneux S, Ghebremichael S, Hanekom M, Hoffner S, Jiao WW, Kalon S, Kohl TA, Kontsevaya I, Lillebaek T, Maeda S, Nikolayevskyy V, Rasmussen M, Rastogi N, Samper S, Sanchez-Padilla E, Savic B, Shamputa IC, Shen A, Sng LH, Stakenas P, Toit K, Varaine F, Vukovic D, et al. Evolutionary history and global

spread of the *Mycobacterium tuberculosis* Beijing lineage. Nat Genet. 2015; 47:242–9. [PubMed: 25599400]

18. Niemann S, Diel R, Khechinashvili G, Gegia M, Mdivani N, Tang YW. *Mycobacterium tuberculosis* Beijing lineage favors the spread of multidrug-resistant tuberculosis in the Republic of Georgia. J Clin Microbiol. 2010; 48:3544–50. [PubMed: 20702677]
19. Casali N, Nikolayevskyy V, Balabanova Y, Harris SR, Ignatyeva O, Kontsevaya I, Corander J, Bryant J, Parkhill J, Nejentsev S, Horstmann RD, Brown T, Drobniewski F. Evolution and transmission of drug-resistant tuberculosis in a Russian population. Nat Genet. 2014; 46:279–86. [PubMed: 24464101]
20. Couvin D, Rastogi N. Tuberculosis – a global emergency: tools and methods to monitor, understand, and control the epidemic with specific example of the Beijing lineage. Tuberc Edinb. 2015; 95(Suppl 1):S177–89. http://dx.doi.org/10.1016/j.tube.2015.02.023.
21. Kvasnovsky CL, Cegielski JP, Erasmus R, Siwisa NO, Thomas K, der Walt ML. Extensively drug-resistant TB in Eastern Cape, South Africa: high mortality in HIV-negative and HIV-positive patients. J Acquir Immune Defic Syndr. 2011; 57:146–52. [PubMed: 21297482]
22. Ogarkov O, Mokrousov I, Sinkov V, Zhdanova S, Antipina S, Savilov E. 'Lethal' combination of *Mycobacterium tuberculosis* Beijing genotype and human CD209 -336G allele in Russian male population. Infect Genet Evol. 2012; 12:732–6. [PubMed: 22027159]
23. Balabanova Y, Radiulyte B, Davidaviciene E, Hooper R, Ignatyeva O, Nikolayevskyy V, Drobniewski FA. Survival of drug resistant tuberculosis patients in Lithuania: retrospective national cohort study. BMJ Open. 2011; 1:e000351.
24. Klopper M, Warren RM, Hayes C, Gey van Pittius NC, Streicher EM, Muller B, Sirgel FA, Chabula-Nxiweni M, Hoosain E, Coetzee G, David van Helden P, Victor TC, Trollip AP. Emergence and spread of extensively and totally drug-resistant tuberculosis, South Africa. Emerg Infect Dis. 2013; 19:449–55. [PubMed: 23622714]
25. Pietersen E, Ignatius E, Streicher EM, Mastrapa B, Padanilam X, Pooran A, Badri M, Lesosky M, van Helden P, Sirgel FA, Warren R, Dheda K. Long-term outcomes of patients with extensively drug-resistant tuberculosis in South Africa: a cohort study. Lancet. 2014; 383:1230–9. [PubMed: 24439237]
26. Agerton TB, Valway SE, Blinkhorn RJ, Shilkret KL, Reves R, Schluter WW, Gore B, Pozsik CJ, Plikaytis BB, Woodley C, Onorato IM. Spread of strain W, a highly drug-resistant strain of *Mycobacterium tuberculosis*, across the United States. Clin Infect Dis. 1999; 29:85–92. discussion 93–85. [PubMed: 10433569]
27. Soini H, Pan X, Amin A, Graviss EA, Siddiqui A, Musser JM. Characterization of *Mycobacterium tuberculosis* isolates from patients in Houston, Texas, by spoligotyping. J Clin Microbiol. 2000; 38:669–76. [PubMed: 10655365]
28. Munsiff SS, Nivin B, Sacajiu G, Mathema B, Bifani P, Kreiswirth BN. Persistence of a highly resistant strain of tuberculosis in New York City during 1990–1999. J Infect Dis. 2003; 188:356–63. [PubMed: 12870116]
29. Ritacco V, Lopez B, Cafrune PI, Ferrazoli L, Suffys PN, Candia N, Vasquez L, Realpe T, Fernandez J, Lima KV, Zurita J, Robledo J, Rossetti ML, Kritski AL, Telles MA, Palomino JC, Heersma H, van Soolingen D, Kremer K, Barrera L. *Mycobacterium tuberculosis* strains of the Beijing genotype are rarely observed in tuberculosis patients in South America. Mem Inst Oswaldo Cruz. 2008; 103:489–92. [PubMed: 18797764]
30. Iwamoto T, Grandjean L, Arikawa K, Nakanishi N, Caviedes L, Coronel J, Sheen P, Wada T, Taype CA, Shaw MA, Moore DA, Gilman RH. Genetic diversity and transmission characteristics of Beijing family strains of *Mycobacterium tuberculosis* in Peru. PLoS One. 2012; 7:e49651. [PubMed: 23185395]
31. Gomes HM, Elias AR, Oelemann MA, Pereira MA, Montes FF, Marsico AG, Kritski AL, dos Filho LA, Caldas PC, Possuelo LG, Cafrune P, Rossetti ML, Lucena N, Saad MH, Cavalcanti HR, Leite CQ, de Brito RC, Lopes ML, Lima K, Souza M, de Trindade RC, Zozio T, Sola C, Rastogi N, Suffys PN. Spoligotypes of *Mycobacterium tuberculosis* complex isolates from patients residents of 11 states of Brazil. Infect Genet Evol. 2012; 12:649–56. [PubMed: 21907830]
32. Ritacco V, Iglesias MJ, Ferrazoli L, Monteserin J, Dalla Costa ER, Cebollada A, Morcillo N, Robledo J, de Waard JH, Araya P, Aristimuno L, Diaz R, Gavin P, Imperiale B, Simonsen V,

Zapata EM, Jimenez MS, Rossetti ML, Martin C, Barrera L, Samper S. Conspicuous multidrug-resistant *Mycobacterium tuberculosis* cluster strains do not trespass country borders in Latin America and Spain. Infect Genet Evol. 2012; 12:711–7. [PubMed: 21718805]

33. Martins MC, Giampaglia CM, Oliveira RS, Simonsen V, Latrilha FO, Moniz LL, Couvin D, Rastogi N, Ferrazoli L. Population structure and circulating genotypes of drug-sensitive and drug-resistant *Mycobacterium tuberculosis* clinical isolates in Sao Paulo state. Braz Infect Genet Evol. 2013; 14:39–45.
34. Vasconcellos SE, Acosta CC, Gomes LL, Conceicao EC, Lima KV, de Araujo MI, de Leite ML, Tannure F, Caldas PC, Gomes HM, Santos AR, Gomgnimbou MK, Sola C, Couvin D, Rastogi N, Boechat N, Suffys PN. Strain classification of *Mycobacterium tuberculosis* isolates in Brazil based on genotypes obtained by spoligotyping, mycobacterial interspersed repetitive unit typing and the presence of large sequence and single nucleotide polymorphism. PLoS One. 2014; 9:e107747. [PubMed: 25314118]
35. Balcells ME, Garcia P, Meza P, Pena C, Cifuentes M, Couvin D, Rastogi N. A first insight on the population structure of *Mycobacterium tuberculosi*s complex as studied by spoligotyping and MIRU-VNTRs in Santiago, Chile. PLoS One. 2015; 10:e0118007. [PubMed: 25671320]
36. Gomes LL, Vasconcellos SE, Gomes HM, Elias AR, da Silva Rocha A, Ribeiro SC, Panunto AC, Ferrazoli L, da Silva Telles MA, Ivens de AM, Kritski AL, Mokrousov I, Manicheva OA, Lasunskaia E, Suffys PN. Genetic diversity of the *Mycobacterium tuberculosis* Beijing family in Brazil and Mozambique and relation with infectivity and induction of necrosis in THP-1 cells. Tuberc Edinb. 2015; 95(Suppl 1):S190–6. http://dx.doi.org/10.1016/j.tube.2015.02.025.
37. Rosales S, Pineda-Garcia L, Ghebremichael S, Rastogi N, Hoffner SE. Molecular diversity of *Mycobacterium tuberculosis* isolates from patients with tuberculosis in Honduras. BMC Microbiol. 2010; 10:208. [PubMed: 20678242]
38. Brudey K, Driscoll JR, Rigouts L, Prodinger WM, Gori A, Al-Hajoj SA, Allix C, Aristimuno L, Arora J, Baumanis V, Binder L, Cafrune P, Cataldi A, Cheong S, Diel R, Ellermeier C, Evans JT, Fauville-Dufaux M, Ferdinand S, Garcia de Viedma D, Garzelli C, Gazzola L, Gomes HM, Guttierez MC, Hawkey PM, van Helden PD, Kadival GV, Kreiswirth BN, Kremer K, Kubin M, Kulkarni SP, Liens B, Lillebaek T, Ho ML, Martin C, Martin C, Mokrousov I, Narvskaia O, Ngeow YF, Naumann L, Niemann S, Parwati I, Rahim Z, Rasolofo-Razanamparany V, Rasolonavalona T, Rossetti ML, Rusch-Gerdes S, Sajduda A, Samper S, Shemyakin IG, et al. *Mycobacterium tuberculosis* complex genetic diversity: mining the fourth international spoligotyping database (SpolDB4) for classification, population genetics and epidemiology. BMC Microbiol. 2006; 6:23. [PubMed: 16519816]
39. Demay C, Liens B, Burguiere T, Hill V, Couvin D, Millet J, Mokrousov I, Sola C, Zozio T, Rastogi N. SITVITWEB – a publicly available international multi-marker database for studying *Mycobacterium tuberculosis* genetic diversity and molecular epidemiology. Infect Genet Evol. 2012; 12:755–66. [PubMed: 22365971]
40. Comas I, Chakravartti J, Small PM, Galagan J, Niemann S, Kremer K, Ernst JD, Gagneux S. Human T cell epitopes of *Mycobacterium tuberculosi*s are evolutionarily hyperconserved. Nat Genet. 2010; 42:498–503. [PubMed: 20495566]
41. Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosi*s highlights the limitations of current methodologies. PLoS One. 2009; 4:e7815. [PubMed: 19915672]
42. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics. 2010; 26:589–95. [PubMed: 20080505]
43. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. Genome Project Data Processing S. The sequence Alignment/Map format and SAMtools. Bioinformatics. 2009; 25:2078–9. [PubMed: 19505943]
44. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, Weinstock GM, Wilson RK, Ding L. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics. 2009; 25:2283–5. [PubMed: 19542151]
45. Sandgren A, Strong M, Muthukrishnan P, Weiner BK, Church GM, Murray MB. Tuberculosis drug resistance mutation database. PLoS Med. 2009; 6:e2. [PubMed: 19209951]

46. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. Curr Protoc Bioinforma. 2002; Chapter 2(Unit 2):3.

47. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014; 30:1312–3. [PubMed: 24451623]

48. Hasnain SE, O'Toole RF, Grover S, Ehtesham NZ. Whole genome sequencing: a new paradigm in the surveillance and control of human tuberculosis. Tuberc Edinb. 2015; 95:91–4.

49. Villellas C, Aristimuno L, Vitoria MA, Prat C, Blanco S, Garcia de Viedma D, Dominguez J, Samper S, Ainsa JA. Analysis of mutations in streptomycin-resistant strains reveals a simple and reliable genetic marker for identification of the *Mycobacterium tuberculosis* Beijing genotype. J Clin Microbiol. 2013; 51:2124–30. [PubMed: 23616454]

50. Rindi L, Lari N, Cuccu B, Garzelli C. Evolutionary pathway of the Beijing lineage of *Mycobacterium tuberculosis* based on genomic deletions and mutT genes polymorphisms. Infect Genet Evol. 2009; 9:48–53. [PubMed: 18977316]

51. Chen J, Chen Z, Li Y, Xia W, Chen X, Chen T, Zhou L, Xu B, Xu S. Characterization of gyrA and gyrB mutations and fluoroquinolone resistance in *Mycobacterium tuberculosis* clinical isolates from Hubei Province, China. Braz J Infect Dis. 2012; 16:136–41. [PubMed: 22552454]

52. Maruri F, Sterling TR, Kaiga AW, Blackman A, van der Heijden YF, Mayer C, Cambau E, Aubry A. A systematic review of gyrase mutations associated with fluoroquinolone-resistant *Mycobacterium tuberculosis* and a proposed gyrase numbering system. J Antimicrob Chemother. 2012; 67:819–31. [PubMed: 22279180]

53. Stucki D, Ballif M, Bodmer T, Coscolla M, Maurer AM, Droz S, Butz C, Borrell S, Langle C, Feldmann J, Furrer H, Mordasini C, Helbling P, Rieder HL, Egger M, Gagneux S, Fenner L. Tracking a tuberculosis outbreak over 21 years: strain-specific single-nucleotide polymorphism typing combined with targeted whole-genome sequencing. J Infect Dis. 2015; 211:1306–16. [PubMed: 25362193]

54. Bryant JM, Schurch AC, van Deutekom H, Harris SR, de Beer JL, de Jager V, Kremer K, van Hijum SA, Siezen RJ, Borgdorff M, Bentley SD, Parkhill J, van Soolingen D. Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data. BMC Infect Dis. 2013; 13:110. [PubMed: 23446317]

55. Roetzer A, Diel R, Kohl TA, Ruckert C, Nubel U, Blom J, Wirth T, Jaenicke S, Schuback S, Rusch-Gerdes S, Supply P, Kalinowski J, Niemann S. Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. PLoS Med. 2013; 10:e1001387. [PubMed: 23424287]

56. Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, Eyre DW, Wilson DJ, Hawkey PM, Crook DW, Parkhill J, Harris D, Walker AS, Bowden R, Monk P, Smith EG, Peto TE. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. Lancet Infect Dis. 2013; 13:137–46. [PubMed: 23158499]

57. Perez-Lago L, Comas I, Navarro Y, Gonzalez-Candelas F, Herranz M, Bouza E, Garcia-de-Viedma D. Whole genome sequencing analysis of intrapatient microevolution in *Mycobacterium tuberculosis*: potential impact on the inference of tuberculosis transmission. J Infect Dis. 2014; 209:98–108. [PubMed: 23945373]

58. Schurch AC, Kremer K, Kiers A, Daviena O, Boeree MJ, Siezen RJ, Smith NH, van Soolingen D. The tempo and mode of molecular evolution of *Mycobacterium tuberculosis* at patient-to-patient scale. Infect Genet Evol. 2010; 10:108–14. [PubMed: 19835997]

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.tube.2015.09.001.

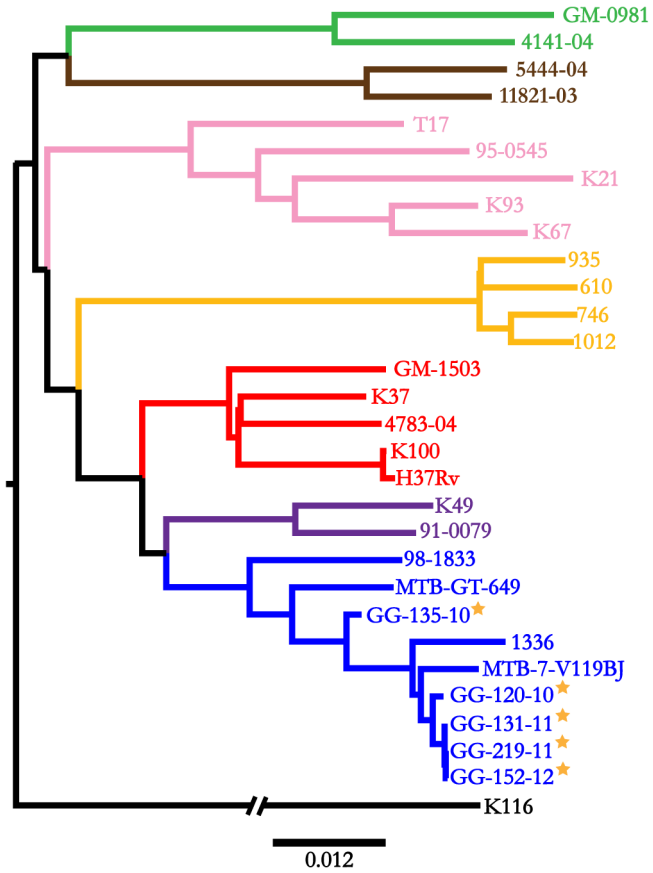

Figure 1.

Neighbor-joining phylogeny based on 18,039 SNPs among 29 global strains of *Mtb*. The tree is rooted by the outgroup *M. canetti*. Branches are colored by lineage: Pink – Lineage 1/ Indo-Oceanic; Blue – Lineage 2/East-Asian; Purple – Lineage 3/East-African-Indian; Red – Lineage 4/Euro-American; Brown – Lineage 5/West-Africa I; Green – Lineage 6/West Africa II; Yellow – Lineage 7/Ethiopia. Guatemalan isolates are marked with stars. Scale bar indicates substitutions per site.

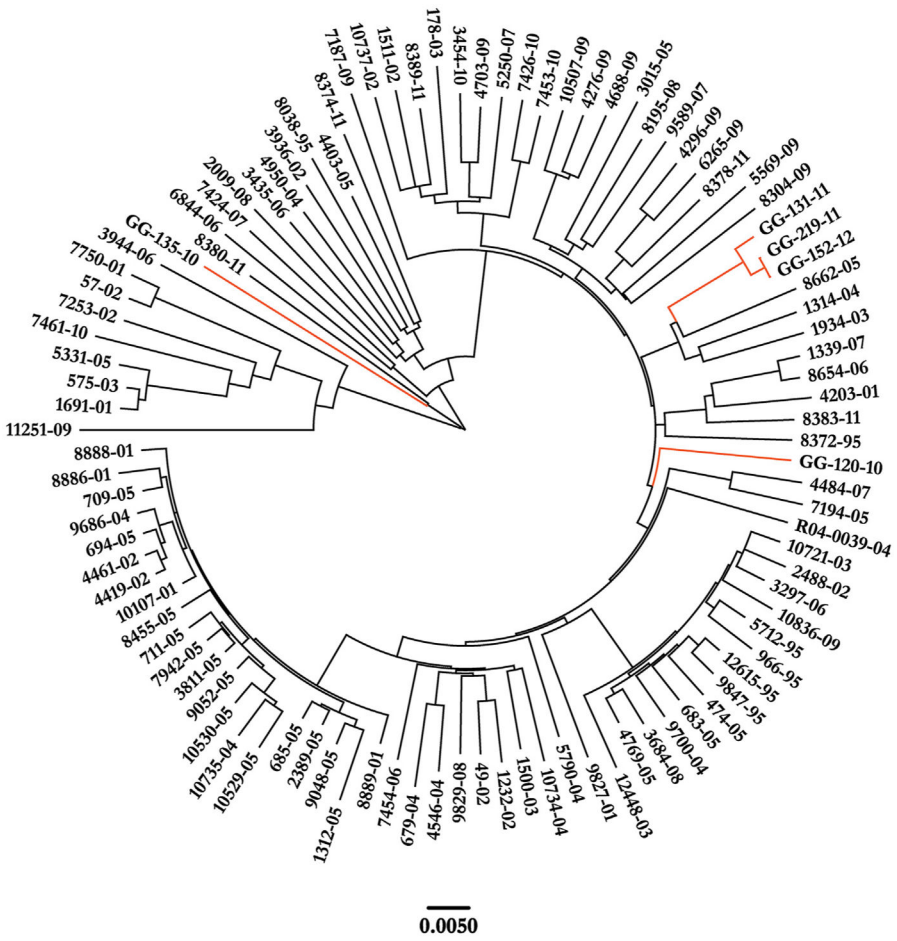

Figure 2.

Neighbor-joining phylogeny based on 6584 SNPs among 101 global Beijing isolates of *Mtb*. Guatemalan isolates are indicated in orange. Scale bar indicates substitutions per site.

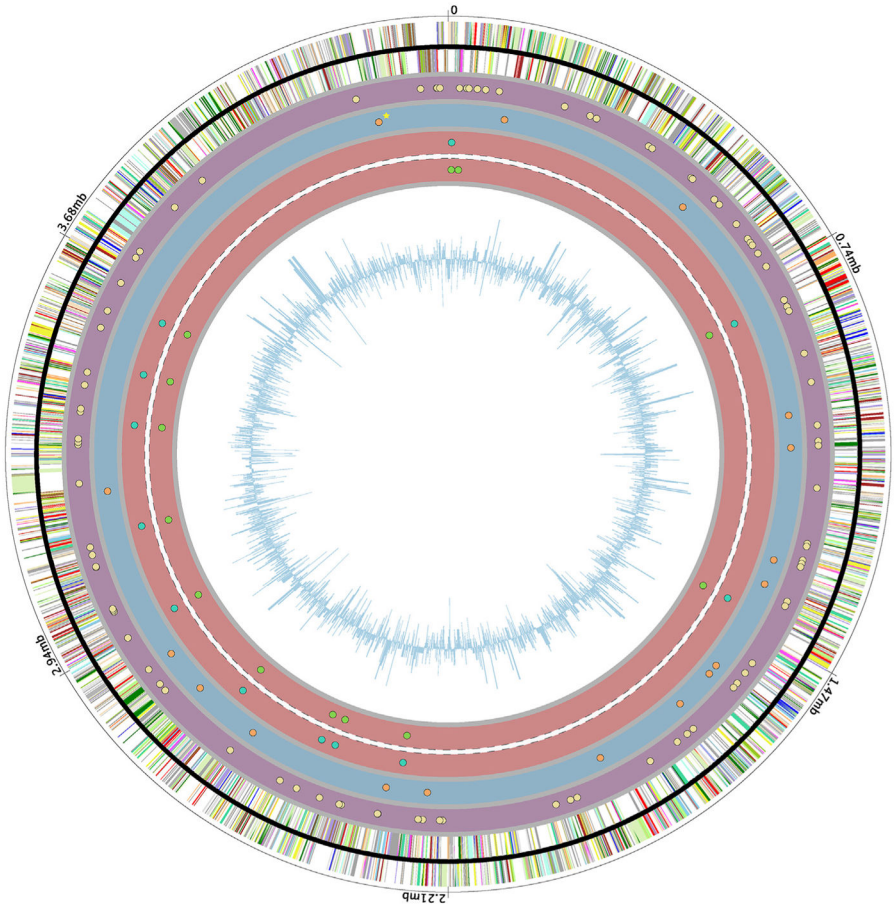

Figure 3.

Circular diagram of polymorphisms unique to the Guatemalan micro-outbreak strains plotted on the H37Rv reference genome. The outer wheel displays base pair coordinates. The second and third wheels represent positive- and negative-strand genes, respectively. Circles four through seven display genome positions of the 125 polymorphisms that separate the Guatemalan outbreak isolates from their nearest neighbor (8662-05) on the tree. The purple shaded wheel with yellow dots displays the 94 SNPs shared among all three isolates. The blue shaded wheel with orange dots displays the 18 SNPs specifically found in GG-131–11 (star denotes presence of two SNPs that are too close to resolve on the diagram). The red wheels display SNPs found in the jail-associated isolates GG-219-11 (outer, blue dots) and GG-152-12 (inner, green dots). The innermost circle represents GC-skew.

Table 1

Spoligotypes and patient risk factors for the 11 identified Beijing isolates from the Clínica Familiar “Luis Angel García” (CFLAG) and its associated public hospital.

Isolate	Spoligotype	Predicted lineage	WGS performed	Lineage by WGS	Risk factors	Residence	Year of Isolation
GG-120-10	000000000003770	Beijing	Yes	Beijing	HIV positive	Urban	2010
GG-135-10	000000000003770	Beijing	Yes	Beijing	HIV positive	Rural Department	2010
GG-22-10	000000000003771	Beijing	No	–	HIV negative	Urban	2010
GG-206-10	000000000003771	Beijing	No	–	HIV negative	Urban, near local jail	2010
GG-32-10	000000000003771	Beijing	No	–	HIV status unknown	Urban, near local jail	2010
GG-131-11	000000000003771	Beijing	Yes	Beijing	HIV positive	Urban, near local jail	2011
GG-219-11	000000000003770	Beijing	Yes	Beijing	HIV positive; history of incarceration	History of incarceration in local jail	2011
GG-152-12	000000000003770	Beijing	Yes	Beijing	HIV positive; history of incarceration	History of incarceration in local jail	2012
GG-10-14	000000000003771	Beijing	No	–	HIV positive; history of incarceration	Urban; history of incarceration in local jail	2014
GG-107-14	000000000003771	Beijing	No	–	HIV negative	Rural Department	2014
GG-109-14	000000000003771	Beijing	No	–	HIV negative	Urban	2014