



Published in final edited form as:

*Lancet Reg Health Am.* 2022 May ; 9: . doi:10.1016/j.lana.2022.100186.

## The role of prisons in disseminating tuberculosis in Brazil: A genomic epidemiology study

Katharine S. Walter<sup>a,\*</sup>, Paulo César Pereira dos Santos<sup>b</sup>, Thais Oliveira Gonçalves<sup>c</sup>, Bruna Oliveira da Silva<sup>b</sup>, Andrea da Silva Santos<sup>b</sup>, Alessandra de Cássia Leite<sup>b</sup>, Alessandra Moura da Silva<sup>d</sup>, Flora Martinez Figueira Moreira<sup>b</sup>, Roberto Dias de Oliveira<sup>e</sup>, Everton Ferreira Lemos<sup>d</sup>, Eunice Cunha<sup>c</sup>, Yiran E. Liu<sup>a,f</sup>, Albert I. Ko<sup>g,k</sup>, Caroline Colijn<sup>h</sup>, Ted Cohen<sup>g</sup>, Barun Mathema<sup>i</sup>, Julio Croda<sup>d,g,j,1</sup>, Jason R. Andrews<sup>a,1</sup>

<sup>a</sup>Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine, Stanford, CA 94305, United States

<sup>b</sup>Health Sciences Research Laboratory, Federal University of Grande Dourados, Dourados, Brazil

<sup>c</sup>Laboratory of Bacteriology, Central Laboratory of Mato Grosso do Sul, Campo Grande, Brazil

<sup>d</sup>School of Medicine, Federal University of Mato Grosso do Sul, School of Medicine, Campo Grande, Brazil

<sup>e</sup>Nursing Course, State University of Mato Grosso do Sul, Dourados, Brazil

<sup>f</sup>Cancer Biology Graduate Program, Stanford University School of Medicine, Stanford, United States

<sup>g</sup>Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, United States

<sup>h</sup>Department of Mathematics, Simon Fraser University, Burnaby, Canada

<sup>i</sup>Department of Epidemiology, Columbia University Mailman School of Public Health, New York, United States

<sup>j</sup>Mato Grosso do Sul Office, Oswaldo Cruz Foundation, Campo Grande, Brazil

<sup>k</sup>Instituto Gonçalo, Moniz, Fundação Oswaldo Cruz, Salvador, BA, Brazil

### Summary

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

\*Corresponding author. [kwalter@stanford.edu](mailto:kwalter@stanford.edu) (K.S. Walter).

<sup>1</sup>These authors contributed equally to this work.

Authors' contributions

JRA and JC contributed to funding acquisition, project administration, resources, and supervision. JRA, JC, KSW, CC, TC, BM, and AIK contributed to conceptualization and methodology. PCPS, TOG, BOS, ASS, ACL, AMS, FMFM, RDO, EFL, EC, and YL contributed to investigation and data curation. KSW and JRA contributed to formal analysis, visualization, and writing the original draft. All authors contributed to review & editing.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.lana.2022.100186.

Declaration of interests

The authors declare no conflicts of interest.

**Background**—Globally, prisons are high-incidence settings for tuberculosis. Yet the role of prisons as reservoirs of *M. tuberculosis*, propagating epidemics through spillover to surrounding communities, has been difficult to measure directly.

**Methods**—To quantify the role of prisons in driving wider community *M. tuberculosis* transmission, we conducted prospective genomic surveillance in Central West Brazil from 2014 to 2019. We whole genome sequenced 1152 *M. tuberculosis* isolates collected during active and passive surveillance inside and outside prisons and linked genomes to detailed incarceration histories. We applied multiple phylogenetic and genomic clustering approaches and inferred timed transmission trees.

**Findings**—*M. tuberculosis* sequences from incarcerated and non-incarcerated people were closely related in a maximum likelihood phylogeny. The majority (70.8%; 46/65) of genomic clusters including people with no incarceration history also included individuals with a recent history of incarceration. Among cases in individuals with no incarceration history, 50.6% (162/320) were in clusters that included individuals with recent incarceration history, suggesting that transmission chains often span prisons and communities. We identified a minimum of 18 highly probable spillover events, *M. tuberculosis* transmission from people with a recent incarceration history to people with no prior history of incarceration, occurring in the state's four largest cities and across sampling years. We additionally found that frequent transfers of people between the state's prisons creates a highly connected prison network that likely disseminates *M. tuberculosis* across the state.

**Interpretation**—We developed a framework for measuring spillover from high-incidence environments to surrounding communities by integrating genomic and spatial information. Our findings indicate that, in this setting, prisons serve not only as disease reservoirs, but also disseminate *M. tuberculosis* across highly connected prison networks, both amplifying and propagating *M. tuberculosis* risk in surrounding communities.

**Funding**—Brazil's National Council for Scientific and Technological Development and US National Institutes of Health.

## Keywords

Tuberculosis; Epidemiology; Genomics; Disease reservoir; Spillover; Prisons

## Introduction

Tuberculosis remains one of the leading causes of death by an infectious disease and led to 1.5 million deaths in 2020.<sup>1</sup> Locating where *M. tuberculosis* transmission occurs is critical for developing public health interventions to block transmission and prevent future infections.<sup>2,3</sup> Built environments often shape *M. tuberculosis* transmission risk, and high-transmission settings like prisons, hospitals, and mines may have epidemiological impacts that transcend their walls,<sup>4–6</sup> serving as institutional amplifiers of infection.<sup>7,8</sup> In particular, the conditions of detention, including severe overcrowding, poor ventilation, poor nutrition, and limited access to healthcare, put incarcerated people at elevated risk of many infectious diseases, including tuberculosis.<sup>4–6</sup>

Globally, incidence of tuberculosis in incarcerated populations is more than ten times that in the general population.<sup>5,9</sup> The role of prisons in driving community tuberculosis epidemics is of particular concern in Central and South America, where the incarcerated population has increased by over 200% since 2000 and tuberculosis cases among the incarcerated population have increased by over 269%.<sup>10</sup> Alarming, increasing tuberculosis notifications within prisons more than offset declines in tuberculosis outside prisons across the region.<sup>10</sup>

The full impact of increasing incarceration rates on the tuberculosis epidemic in the Americas and elsewhere is underestimated by total incident cases within prisons. Many people are infected in prison, but only diagnosed later.<sup>8,11,12</sup> Further, spillover may occur when incarcerated people, prison staff, or visitors are infected within high-incidence prison environments and transmit *M. tuberculosis* onwards in communities outside prison. Previous studies have found that the excess risk of tuberculosis within prisons extends to surrounding communities<sup>13,14</sup> and that *M. tuberculosis* genotypes responsible for jail and prison outbreaks are also found in the surrounding communities.<sup>11,13,15</sup> Because transmission is rarely observed, however, it is often difficult to link community cases to transmission from prisons and previous studies have largely been associative.

Pathogen genomes can be powerfully harnessed to infer high-resolution evolutionary and transmission histories, including who-infected-whom.<sup>16</sup> One previous study in Georgia used *M. tuberculosis* genomes to identify frequent transmission of multi-drug resistant tuberculosis from prisons to surrounding communities.<sup>17</sup> However, studies have not yet combined *M. tuberculosis* genomic, epidemiologic, and location data to reconstruct transmission linkages between potential institutional amplifiers and the community. To quantify the contribution of prison spillover to tuberculosis transmission in the community, we conducted prospective genomic surveillance for tuberculosis in the Central West Brazilian state of Mato Grosso do Sul. We integrated *M. tuberculosis* genomes with detailed incarceration histories and took multiple phylogenetic and genomic clustering approaches to reconstruct spillover events, transmission of tuberculosis from prisons to the community.

## Methods

### Study design

We conducted population-based tuberculosis surveillance in the state of Mato Grosso do Sul in Central West Brazil from January 2014 through May 2019. Surveillance included active screening in three of the largest prisons in the state as well as ongoing passive surveillance focused on the two largest cities in the state, and three cities at the state's border with Paraguay and Bolivia (Figure 1a,b; S1 Text). All participants provided written consent, and this study was conducted with the approval of the Research Ethics Committee from the Federal University of Grande Dourados, Federal University of Mato Grosso do Sul and National Research Ethics Committee (CONEP) (CAAE 37237814.4.0000.5160, 2676613.3.1001.5160, and 26620619.6.0000.0021) and Stanford University Institutional Review Board (IRB-40285) (S1 Text).

## Incarceration history

We obtained permission from the Mato Grosso do Sul state prison administration agency, Agência Estadual de Administração do Sistema Penitenciário, to access a database of all prison entries, exits, and transfers within the criminal justice record system from January 1, 2005 through December 31, 2018 (S1 Text).

## Whole genome sequencing

We sequenced whole *M. tuberculosis* genomes from cultured isolates (S1 Text) on an Illumina NextSeq ( $2 \times 151$ -bp). Sequence data is available on the Sequence Read Archive, in BioProject PRJNA671770. We applied variant calling methods closely following those described in Menardo et al.<sup>18</sup> to be consistent with the methods used for molecular clock estimation (S1 Text).

## Phylogenetic and Bayesian evolutionary analysis

We fit a maximum likelihood tree with RAxML-NG 1.0.1<sup>19</sup> and rooted the tree on two Lineage 1 isolates from this study. We clustered genomes using a 12-SNP threshold. For each cluster including three or more isolates, we fit Bayesian trees with BEAST 2.6.2<sup>20</sup> with a strict clock, constant coalescent population size model using tuberculosis notification dates to calibrate tips.

To focus on densely sampled trees that would not be dominated by unsampled hosts, transmission trees were inferred using timed trees with most recent common ancestors of 2012 and later. To speed computation and because our focus was on resolving phylogenetic structure between closely related isolates likely linked in transmission, we fit Bayesian timed phylogenies for genomic clusters in parallel rather than for the full collection of isolates.

We expected lineage 4 isolates from our study to share a common substitution rate. Because it was not possible to estimate substitution rates from small trees for our Bayesian timed tree inference, we estimated substitution rate for the largest cluster (170 tips) using a log-normal prior on substitution rate (mean =  $-17$ , SD = 2), consistent with previous estimates of the *M. tuberculosis* lineage 4 substitution rate (normally distributed with mean =  $5.8 \times 10^{-8}$ , SD =  $2.0 \times 10^{-8}$ ).<sup>18</sup> We then fit a log-normal distribution to the posterior samples of the substitution rate for the largest cluster and used this as a narrow log-normal prior on substitution rate (mean =  $-16.5$ , SD=0.17) for all other clusters. We used an HKY substitution rate model and estimated base frequencies. We ran sampling chains for 100 million iterations, discarding a burn-in of 10% of samples, and confirmed convergence using effective sample size estimates (confirming all parameters had ESS > 200) in Tracer 1.7.1.<sup>21</sup> We used TreeAnnotator v2.6.2 to summarize posterior trees as maximum clade credibility (MCC) trees, with median node heights. We corrected for ascertainment bias by specifying the number of invariant sites in the BEAST XML file (S1 Text).

## Transmission inference

*M. tuberculosis* phylogenetic trees represent patterns of evolutionary relatedness between the consensus bacterial genomes sampled from different individuals. Because most outbreaks

are incompletely sampled and individuals may be infected with diverse populations of *M. tuberculosis*, phylogenies do not represent the underlying history of transmission.<sup>16</sup> We used TransPhylo<sup>16</sup> to infer transmission trees—including unsampled hosts—consistent with the underlying timed phylogenies and epidemiology of tuberculosis (i.e. generation time and sampling intervals) (S1 Text).

We then combined transmission trees with incarceration histories to identify spillover events from prisons to surrounding communities. We defined spillover events as transmission from an individual with a history of current or former incarceration to the community, including individuals with no incarceration history and prior incarceration. We considered transmission probabilities >50% and included spillover events no more than two degrees removed from prisons. TransPhylo does not require that all transmission events be observed. Therefore, we are able to quantify the minimum number of transmission events, but not estimate the total role of spillover on community transmission.

### Role of the funding source

The study sponsors had no role in study design, data collection, data analysis, interpretation, writing of the report, or in the decision to submit the paper for publication.

## Results

### Tuberculosis and incarceration trends in Mato Grosso do Sul

Brazil's incarcerated population grew by 107% (from 361,402 to 748,009) from 2005 to 2019, an increase closely paralleled in the Central West state of Mato Grosso do Sul, where the incarcerated population grew by 115% (7891 to 16,976) over the same period (Figure 1a–c). In 2019, prisons in the state were at 203% occupancy. While the state's notification rate of new and retreatment tuberculosis in the general population (38 per 100,000) is similar to Brazil's national tuberculosis notification rate (40 per 100,000) in 2019,<sup>1</sup> the notification rate within prisons was 43 times as high (1666 per 100,000), again similar to Brazil's national notification rate within prisons (1596 per 100,000) in 2019 (Figure 1d,e).

### Population-based genomic surveillance across Mato Grosso do Sul state

From 2014 to 2019, 3491 new and retreatment cases of tuberculosis were reported in the Campo Grande and Dourados, the two largest cities in Mato Grosso do Sul, Brazil. Of these, 1249 had positive cultures and we sequenced 787 *M. tuberculosis* isolates. Other culture-positive isolates from these cities were not available, were contaminated, or had insufficient DNA for sequencing. To maximize capture of potential transmission linkages, we supplemented our collection by sequencing isolates from three other cities in the state and from earlier years in Dourados for a total of 1090 isolates from unique tuberculosis notifications (Figure 1b). Whole genome sequences (WGS) for a total of 1043 isolates met our coverage and quality criteria (Methods). We excluded 10% (108 of 1043) of isolates with mixed infection, resulting in 935 high-quality genomes from 935 unique tuberculosis episodes from 918 individuals.

Of the 935 isolates in our analyses, 50% (465 of 935) were from patients who were incarcerated at the time of tuberculosis notification; 16% (150 of 935) were formerly incarcerated; and 34% (320 of 935) of the study population did not have an incarceration history. Among those who did not have an incarceration history, 32 people reported contact with incarcerated individuals. Additional population characteristics are in Table S1. Isolates were largely from *M. tuberculosis* lineage 4 (932 of 935) and predominantly fell into three sub-lineages: lineage 4.1, 4.3, and 4.4 (Figure 2). We focus subsequent analyses of transmission on lineage 4. Overall, we found a low prevalence of antibiotic resistance across isolates, with 93.7% of isolates susceptible to all antibiotics (Figure 2).

### Phylogenetic structure of *M. tuberculosis* from prisons and the community

In a maximum likelihood phylogeny (Figure 2), *M. tuberculosis* isolates sampled from incarcerated and non-incarcerated people are distributed across the tree and do not form distinct clades, indicating recent shared evolutionary history of isolates sampled from prisons and the community.

We tested for phylogenetic signal in incarceration status of sampled tuberculosis patients—whether incarceration status was distributed randomly across the tree—with Fritz’s D,<sup>22</sup> a test of “clumping” in observed traits. The D statistic is equal to 1 if the trait is distributed randomly across tree tips with respect to the phylogeny and 0 if the trait is consistent with a pattern of clumping consistent with Brownian motion of the trait along the phylogenetic tree (S1 Text).

We predicted that if prison and community epidemics were distinct, incarceration status would exhibit an extremely clumped pattern on the *M. tuberculosis* phylogeny, with separate clades including samples from incarcerated people and community members. We found that incarceration status had a moderate phylogenetic signal (Fritz’s D=0.60), indicating a pattern of shared evolutionary history of isolates collected from individuals with and without an incarceration history, with some degree of “clumping” with respect to incarceration status.

We similarly found that city (including the two major cities, Campo Grande and Dourados) had a moderate phylogenetic signal (Fritz’s D=0.49), consistent with *M. tuberculosis* migration across the state, with some geographic structure, indicating more frequent local transmission (S1 Text).

### Genomic clustering of *M. tuberculosis* from prisons and the community

Patterns of genomic relatedness are often used to infer potential recent transmission; *M. tuberculosis* isolates that are more closely genetically related are hypothesized to be more likely linked through recent transmission rather than travel-associated importation or re-activation of genetically distinct latent infections.<sup>23,24</sup> To identify potential transmission clusters, we applied a commonly used 12-SNP threshold,<sup>23</sup> including all isolates within the threshold distance of at least one other clustered isolate. Eighty-three percent (777 of 935) of the isolates fell into 84 genomic clusters (each including 2 to 170 isolates), evidence that notifications were largely attributable to recent transmission (Figure 2).<sup>24</sup> Isolates from incarcerated people were more frequently clustered (93.3%, 434 of 465), than those from

formerly incarcerated (86.0%, 129 of 150;  $p < 0.0001$ ), or never incarcerated people (66.9%, 214 of 320;  $p < 0.0001$ ), again suggesting more recent transmission within prisons.

We predicted that if prison and community-associated epidemics were distinct, isolates from the community would be most closely related to and cluster with other isolates from the community and vice versa. The majority (70.8%; 46/65) of genomic clusters including people with no incarceration history also included individuals with a recent history of incarceration. Among cases in individuals with no incarceration history, 50.6% (162 of 320) were in clusters that included individuals with recent incarceration history, suggesting that transmission chains often span prisons and communities.

We found a similar pattern of clustering of isolates from community members and individuals with an incarceration history when using an alternative 5-SNP threshold for genomic clustering<sup>23</sup> (S2 Text).

Incarcerated people were sampled at a higher rate due to active case finding in three prisons in the state. From 2014 to 2019, in the cities Campo Grande and Dourados, our genomic collection includes *M. tuberculosis* genomes from 35% of notified cases of TB among incarcerated individuals, 17% of formerly incarcerated individuals, and 11% of never incarcerated individuals (Fig. S2). To test if our findings were robust to our sampling framework, we conducted a permutation test in which we resampled isolates to create sample sets with equivalent sampling coverage for each population. In 1000 resampled datasets, we found a pattern of clustering very similar to what we found in our original data-set: isolates from incarcerated people were more frequently clustered (mean 90.4% clustered), than those from formerly incarcerated (mean 83.2% clustered), or never incarcerated people (mean 60.0% clustered). This confirmed that our findings were robust to our genomic sampling approach. Our findings were consistent when including only isolates from Campo Grande and Dourados and when including only isolates collected from 2014 to 2019 (S2 Text).

### **A *M. tuberculosis* clone spans prisons and the community across the state**

The largest potential transmission cluster, including 170 isolates sampled from September 2010 to April 2019, was distributed across the state, including cases found across state's two largest cities, Campo Grande and Dourados, as well as the smaller cities Corumbá (on the Bolivian border) and Ponta Porã (on the Paraguayan border) (Figures 1b, 3a). The cluster had a most recent common ancestor in 1996 (95% HPD: 1989–2003; Figure 3a), indicating that it had circulated for approximately 23 years before the most recent samples were collected. 103 isolates were from people incarcerated at the time of diagnosis, 35 from people formerly incarcerated, and 32 from people with no incarceration history.

As observed in other clusters, isolates from people who are currently and formerly incarcerated were closely genetically related—and often, identical—to isolates from people who were not incarcerated at the time of notification, suggesting they are linked through recent transmission (Figure 3b). The three largest haplotypes (identical sequences), included both incarcerated and never incarcerated individuals, indicating very recent transmission between prisons and the community.

### Prison networks disseminate *M. tuberculosis* across space

We hypothesized that the criminal justice system's frequent transfer of people between prisons and jails could disseminate TB across the state. We analyzed movements in the state's incarceration database and found that the average duration of incarceration was 230 days in a prison and 25 days in a jail. The incarceration database documents 7982 mean yearly transfers between prisons (including closed and semi-open prisons) from 2015 to 2018 in a prison population of 17,221 in August 2018, including frequent transfers between cities across the state (Figure 4a).

To investigate the role of the criminal justice system in disseminating infection, we tracked the spread of the largest sampled *M. tuberculosis* cluster (Figure 3) across prisons in Mato Grosso do Sul (Figure 4b). From early 2011 to 2014, sampled isolates from the cluster were concentrated in the maximum security prison in Dourados, after which it was identified within two prisons in Campo Grande and then exported to prisons and jails in other cities. This genomic cluster remained the dominant sampled genotype across the city of Dourados for the duration of the study period; however beginning in 2015, other genotypes replaced this cluster as the most common circulating genotype in prisons. The cluster was not contained within the state's prison network; notified tuberculosis cases among community members outside of prison occurred over the duration of the clusters' state-wide spread.

### Transmission trees reveal frequent spillover of infection

We further investigated the directionality of recent transmission in the state by inferring probabilistic transmission trees representing who-infected-whom for the 84 identified *M. tuberculosis* genomic clusters with TransPhylo.<sup>16</sup> Because our sampling of the TB epidemic in Mato Grosso do Sul was incomplete and we did not observe the majority of transmission events resulting in sampled cases, we quantified the minimum number of recent infections attributable to prison spillover rather than the total role of spillover on community infections. We identified a minimum of 16.8% (18 of 107) infections among never incarcerated individuals were directly attributable to observed spillover from prisons, not including spillover to individuals with an incarceration history. These probable spillover events included transmission from ten genomic clusters; in the three most highly sampled cities of Campo Grande, Corumbá, and Dourados; and every year, from 2014 to 2018. Two of the community members infected via spillover had reported contact with people currently or formerly incarcerated (these contact data were incomplete). Notification dates, which also represent the date at which sputa were collected, differed between primary and secondary infections by a mean of 245 days. Sixteen of the eighteen spillover events were each directly attributable to a single person transmission event (i.e. transmission probability of >50% from a single infector). Interestingly, two of the individuals infected through spillover each had two potential infectors, each currently or formerly incarcerated, such that the total probability of transmission from prison was >50%, despite uncertainty in the identity of the primary case.

For the majority of notified TB cases in our dataset, our approach did not infer a probable donor (source of transmission); therefore, the number of inferred spillover events represents a conservative or minimum estimate of likely spillovers contributing to infection



among never incarcerated individuals. Our transmission inferences, including the number of probable spillover events, were sensitive to epidemiological priors (below), reflecting the uncertainty inherent in transmission reconstruction for an incompletely sampled epidemic.

Together, these findings indicate that spillover of *M. tuberculosis* infection from prisons to surrounding communities is frequent. The result is that the 0.84% of the state's population incarcerated at a given time (22,706 of 2,713,147 in 2017) or 2.6% currently or formerly incarcerated at a given time (71,703 of 2,713,147), have a disproportionate effect on TB transmission outside of prisons.

### Transmission inferences are sensitive to epidemiological distributions

We conducted a sensitivity analysis to determine the effect of epidemiological distributions (Table S2) on transmission inferences for the largest genomic cluster (170 isolates). Transmission probabilities inferred for pairs of individuals in this cluster were well correlated across generation time distributions (Spearman's correlation: 0.85–0.94, Fig. S3) and sampling time distributions (Spearman's correlation: 0.83–0.96, Fig. S4), although the total number of observed transmission events was strongly influenced by both distributions (Fig. S5). This suggests that, as expected, epidemiological information impacts the magnitude of inferred transmission probabilities rather than the identity of predicted transmission pairs. As expected, the minimum number of observed spillover events in the largest cluster were sensitive to epidemiological distributions, when transmission was defined as a transmission probability greater than or equal to 50%, and ranged from 0 (short sampling and generation times) to 18 (medium generation time and very long sampling time) (Fig. S6).

As expected, the minimum number of identified spillover events to individuals with no previous incarceration was dependent on the probability threshold used to infer a likely transmission event (50%, above). When increasing the transmission probability threshold from 25% to 75%, the inferred number of spillovers varied from 33 to 6 (base case, 18).

## Discussion

Here, we linked intensive *M. tuberculosis* genomic surveillance to detailed individual-level incarceration histories, allowing us to reconstruct high-resolution population-wide tuberculosis transmission in Central West Brazil, a setting with an extreme disparity in tuberculosis incidence between prisons and surrounding communities. We found that *M. tuberculosis* sequences from incarcerated and non-incarcerated individuals were closely evolutionarily related and sometimes identical. Locally circulating *M. tuberculosis* clones spanned prisons and the community, and the majority of genomic clusters included individuals with and without a recent history of incarceration. In reconstructed transmission trees, we found evidence of frequent spillover of infection in all the major cities in the state and across sampling years, indicating that spillover is not the effect of a single prison but a widespread phenomenon. We found that frequent transfers of people between prisons in Mato Grosso do Sul, Brazil creates a highly connected prison contact network that likely facilitates the spread of *M. tuberculosis* between prisons and may disseminate infection risk across the state.

While alarming disparities in tuberculosis incidence between prisons and surrounding communities are well-documented,<sup>5,9</sup> it has been more challenging to directly link tuberculosis cases in the community to spillover from prisons. Previous studies of the broader effects of prisons on general TB epidemics have largely been associative. Here, we describe a probabilistic framework for identifying spillover events that can be applied more generally to measure the impact of high-transmission environments such as mines or hospitals on tuberculosis transmission in the general population. Our finding of the likely role of prison transfers in disseminating *M. tuberculosis* highlights an additional role that carceral systems may play in propagating infectious diseases both within highly-connected prison systems and across space.

The rapid rise in incarceration rates in many countries<sup>25</sup> and stark global disparities in tuberculosis incidence inside compared to outside prisons<sup>4,5,9,10</sup> suggests that spillover of infection may contribute to community TB transmission outside of Brazil. A study of multi-drug resistant (MDR) TB transmission found that up to 31% of MDR TB in the country of Georgia was directly or indirectly linked to prisons.<sup>17</sup> Our transmission inference framework is conservative in that it identifies a minimum contribution of spillover to total community transmission, and therefore, is consistent with these findings. Prisons and other detention centers have served as reservoirs for other pathogens including meningococcus<sup>26</sup> and SARS-CoV-2.<sup>27–29</sup> Prison transfers have also spread of SARS-CoV-2 across prisons in California, resulting in widespread outbreaks.<sup>27</sup>

The frequent spillover of infection from prisons to surrounding communities indicates that for TB control to be achieved in Central West Brazil, TB control efforts must focus on reducing the extraordinarily high transmission risk within prisons. Tuberculosis control programs should expand routine active screenings during<sup>8,30</sup> and following incarceration so that cases can be diagnosed early and linked to treatment. Preventive therapy is not provided in prisons in Brazil and many other low- and middle-income countries and could represent an effective means for tuberculosis prevention, particularly among individuals exiting prisons who are at high risk of infection. More broadly, improving the unsanitary, inhumane conditions of incarceration and expanding access to primary healthcare and nutrition are critical.

Reducing this excess burden of disease will require work that extends beyond biomedical interventions. The most direct way to mitigate the excess tuberculosis risk created by prisons is for governments to rapidly reduce incarcerated populations, as advocated by the American Public Health Association, to minimize the broader harms associated with incarceration.<sup>29</sup> Prison releases have already begun as a means of reducing infectious disease transmission risk.<sup>28</sup> In an attempt to reduce the risk of COVID-19 in prisons, for example, Brazil released more than thirty thousand people by mid-2020.<sup>31</sup>

Our study has several limitations. First, our study highlights the uncertainty inherent in transmission inferences drawn from genomic surveillance data, which almost always constitutes an incomplete sample of population-wide infections. We addressed the incomplete sampling of *M. tuberculosis* genomes in two ways. First, we took a probabilistic transmission inference approach that accounts for uncertainty in transmission linkages

and also includes unsampled individuals contributing to transmission.<sup>16</sup> Second, in order to collect the most comprehensive set of *M. tuberculosis* genomes possible, we combined isolates collected during active mass screening within prisons and passive routine surveillance in the community. The result was that isolates from prisons are overrepresented in our genomic collection. Because of the over-sampling within prisons, we could not directly quantify the contribution of transmission within prisons and spillover from prisons on overall recent infections in the study area. However, by taking a probabilistic transmission inference approach that accounted for uncertainty in transmission linkages as well as unsampled individuals contributing to transmission, we were able to quantify the minimum number of reconstructed spillover events consistent with inferred transmission trees. Because we cannot infer the identity or incarceration history of unsampled people who contribute to transmission, our estimates of the contribution of spillover are conservative and likely underestimate the full impact of spillover on infections in the community.

Additionally, it remains difficult to distinguish genomic sampling proportion from generation time and sampling time distributions. For example, two sampled individuals in our dataset could be linked through transmission via an unsampled intermediate contact or could be directly linked in transmission, but have different notification dates due to a relatively long generation time or due to delayed diagnosis and sampling.

Because we conducted transmission inference in a tuberculosis-endemic setting, rather than in a discrete outbreak, our rate of genomic sampling was lower than in some previous tuberculosis transmission studies. To reduce uncertainty in sampling proportion, we conducted transmission inference on recent subtrees for which we had sampled a greater proportion of isolates. To investigate the influence of epidemiological distributions on transmission inferences, we conducted a sensitivity analysis (Table S2) and found that our conclusions about the identity of specific transmission linkages were largely unchanged (S2 Text, Figs. S3, S4). However, as expected, the number of transmission events and thus spillover events was dependent on epidemiological distributions (Figs. S5, S6).

As expected, longer sampling time distributions resulted in a greater number of observed transmission events, as defined either by the sum of transmission probabilities (Fig. S5) or the minimum number of transmission events (transmission probabilities with >50% probability; Fig. S6). The generation time distribution also impacted the number of observed transmissions. Interestingly, the “medium” generation time distribution maximized the number of observed transmission events in our data (Figs. S5 and S6).

Together, the uncertainty in the number of spillover events reflects the difficulty in transmission reconstruction in an endemic setting where *M. tuberculosis* sampling is incomplete. Our findings also highlight the potential for harnessing additional information to reduce uncertainty in transmission inferences. While we did not use location information in transmission inferences here (because quantifying spillover was a major outcome of the study), such information has been used previously. Additionally, approaches are needed to recover more variation from individual *M. tuberculosis* infections, both by incorporating within-host *M. tuberculosis* variation or by recovering variation along the full *M. tuberculosis* genome, including the diverse PE/PPE genes, commonly excluded from

analysis.<sup>18,32,33</sup> Finally, to speed computation, we fit timed phylogenies to genomic clusters in parallel. Approaches for incorporating rapid tree fitting methods into phylogenetic-based transmission frameworks could overcome this and allow for transmission inference for the full sample collection.

Here, we present genomic evidence that spillover of tuberculosis infection risk from prisons to surrounding communities is frequent in Central West Brazil. The dramatic expansion of incarceration in recent decades has put an increasing population at extremely high risk of tuberculosis; this risk extends to surrounding communities. Reducing the excess tuberculosis transmission risk within prisons and other detention centers is therefore an urgent public health priority.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We would like to acknowledge the contributions of AGEPEN, the Mato Grosso do Sul state criminal justice agency; LACEN, the state reference laboratory; and the Mato Grosso do Sul State Health Department. All authors had full access to all the data in the study and had final responsibility for the decision to submit.

## Funding

National Institutes of Health grant R01AI130058 (JRA) and R01 AI149620 (JRA and JC). Brazil's National Council for Scientific and Technological Development grant 404237/2012-6 (JC).

## Data sharing

Sequence data is available on the Sequence Read Archive, in BioProject PRJNA671770.

## References

1. The World Health Organization. Global Tuberculosis Report 2021. Geneva, Switzerland; 2021.
2. Churchyard G, Kim P, Shah NS, et al. What we know about tuberculosis transmission: an overview. *J Infect Dis.* 2017;216:S629–S635. [PubMed: 29112747]
3. Auld SC, Sarita Shah N, Mathema B, et al. Extensively drug-resistant tuberculosis in South Africa: genomic evidence supporting transmission in communities. *Eur Respir J.* 2018;52(4).
4. Dolan K, Wirtz AL, Moazen B, et al. Global burden of HIV, viral hepatitis, and tuberculosis in prisoners and detainees. *Lancet.* 2016;388(10049):1089–1102. [PubMed: 27427453]
5. Cords O, Martinez L, Warren JL, et al. Incidence and prevalence of tuberculosis in incarcerated populations: a systematic review and meta-analysis. *Lancet Public Health.* 2021;6(5):300–308.
6. Urrego J, Ko AI, Da Silva Santos Carbone A, et al. The impact of ventilation and early diagnosis on tuberculosis transmission in Brazilian prisons. *Am J Trop Med Hyg.* 2015;93(4):739–746. [PubMed: 26195459]
7. Basu S, Stuckler D, McKee M. Addressing institutional amplifiers in the dynamics and control of tuberculosis epidemics. *Am J Trop Med Hyg.* 2011;84(1):30–37. [PubMed: 21212197]
8. Mabud TS, de Lourdes Delgado Alves M, Ko AI, et al. Evaluating strategies for control of tuberculosis in prisons and prevention of spillover into communities: an observational and modeling study from Brazil. *PLoS Med.* 2019;16(1) e1002737. [PubMed: 30677013]
9. Baussano I, Williams BG, Nunn P, Beggiano M, Fedeli U, Scano F. Tuberculosis incidence in prisons: a systematic review. *PLoS Med.* 2010;7(12):1–10.

10. Walter KS, Martinez L, Arakaki-Sanchez D, et al. The escalating tuberculosis crisis in central and South American prisons. *Lancet*. 2021;397(10284):1591–1596. [PubMed: 33838724]
11. Sacchi FPC, Praça RM, Tatará MB, et al. Prisons as reservoir for community transmission of tuberculosis, Brazil. *Emerg Infect Dis*. 2015;21(3):452–455. [PubMed: 25642998]
12. Coker R, McKee M, Atun R, et al. Risk factors for pulmonary tuberculosis in Russia: case-control study. *Br Med J*. 2006;332(7533):85–87. [PubMed: 16339219]
13. Warren JL, Grandjean L, Moore DAJ, et al. Investigating spillover of multidrug-resistant tuberculosis from a prison: a spatial and molecular epidemiological analysis. *BMC Med*. 2018;16(1):122. [PubMed: 30071850]
14. Jenkins HE, Plesca V, Ciobanu A, et al. Assessing spatial heterogeneity of multidrug-resistant tuberculosis in a high-burden country. *Eur Respir J*. 2013;42(5):1291–1301. [PubMed: 23100496]
15. S eraphin MN, Didelot X, Nolan DJ, et al. Genomic investigation of a *Mycobacterium tuberculosis* outbreak involving prison and community cases in Florida, United States. *Am J Trop Med Hyg*. 2018;99 (4):867–874. [PubMed: 29987998]
16. Didelot X, Fraser C, Gardy J, Colijn C, Malik H. Genomic infectious disease epidemiology in partially sampled and ongoing outbreaks. *Mol Biol Evol*. 2017;34(4):997–1007. [PubMed: 28100788]
17. Gygli SM, Loiseau C, Jugheli L, et al. Prisons as ecological drivers of fitness-compensated multidrug-resistant *Mycobacterium tuberculosis*. *Nat Med*. 2021;27(7):1171–1177. [PubMed: 34031604]
18. Menardo F, Duchêne S, Brites D, Gagneux S. The molecular clock of *Mycobacterium tuberculosis*. *PLoS Pathog*. 2019;15(9).
19. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable, and user-friendly tool for maximum likelihood phylogenetic inference. Wren J, editor *Bioinformatics*. 2019;35(21):4453–4455. [PubMed: 31070718]
20. Bouckaert R, Vaughan TG, Barido-Sottani J, et al. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol*. 2019;15(4) e1006650. [PubMed: 30958812]
21. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst Biol*. 2018;67(5):901–904. [PubMed: 29718447]
22. Fritz SA, Purvis A. Selectivity in mammalian extinction risk and threat types: a new measure of phylogenetic signal strength in binary traits. *Conserv Biol*. 2010;24(4):1042–1051. [PubMed: 20184650]
23. 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(2):137–146. [PubMed: 23158499]
24. Borgdorff MW, Van Soolingen D. The re-emergence of tuberculosis: what have we learnt from molecular epidemiology? *Clin Microbiol Infect*. 2013;19:889–901. [PubMed: 23731470]
25. Walmsley R *World Prison Population List* (twelfth edition). Institute for Criminal Policy Research; 2018. Vol. 12.
26. Tappero JW, Reporter R, Wenger JD, et al. Meningococcal disease in Los Angeles County, California, and among Men in the County Jails. *N Engl J Med*. 1996;335(12):833–841. [PubMed: 8778600]
27. McCoy S, Bertozze SM, Sears D, Duarte C, Cameron D, Williams B. Urgent memo: COVID-19 outbreak: San Quentin Prison. San Francisco, CA; 2020.
28. Wang EA, Western B, Berwick DM. COVID-19, decarceration, and the role of clinicians, health systems, and payers: a report from the National Academy of Sciences, Engineering, and Medicine. *JAMA*. 2020;324(22):2257–2258. [PubMed: 33196762]
29. American Public Health Association. *Advancing Public Health Interventions to Address the Harms of the Carceral System*. American Public Health Association; 2021. Vol. 2019.
30. Santos A da S, Oliveira RD, Lemos EF, et al. Yield, efficiency and costs of mass screening algorithms for tuberculosis in Brazilian prisons. *Clin Infect Dis*. 2020;72(5):771–777.
31. Oliveira De Andrade R. Covid-19: prisons exposed in Brazil's crisis. *BMJ*. 2020;370.

32. Meehan CJ, Goig GA, Kohl TA, et al. Whole genome sequencing of *Mycobacterium tuberculosis*: current standards and open issues. *Nat Rev Microbiol.* 2019;17(9):533–545. [PubMed: 31209399]
33. Walter KS, Colijn C, Cohen T, et al. Genomic variant-identification methods may alter *Mycobacterium tuberculosis* transmission inferences. *Microb Genom.* 2020;6(8):1–16.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

## Research in context

### Evidence before this study

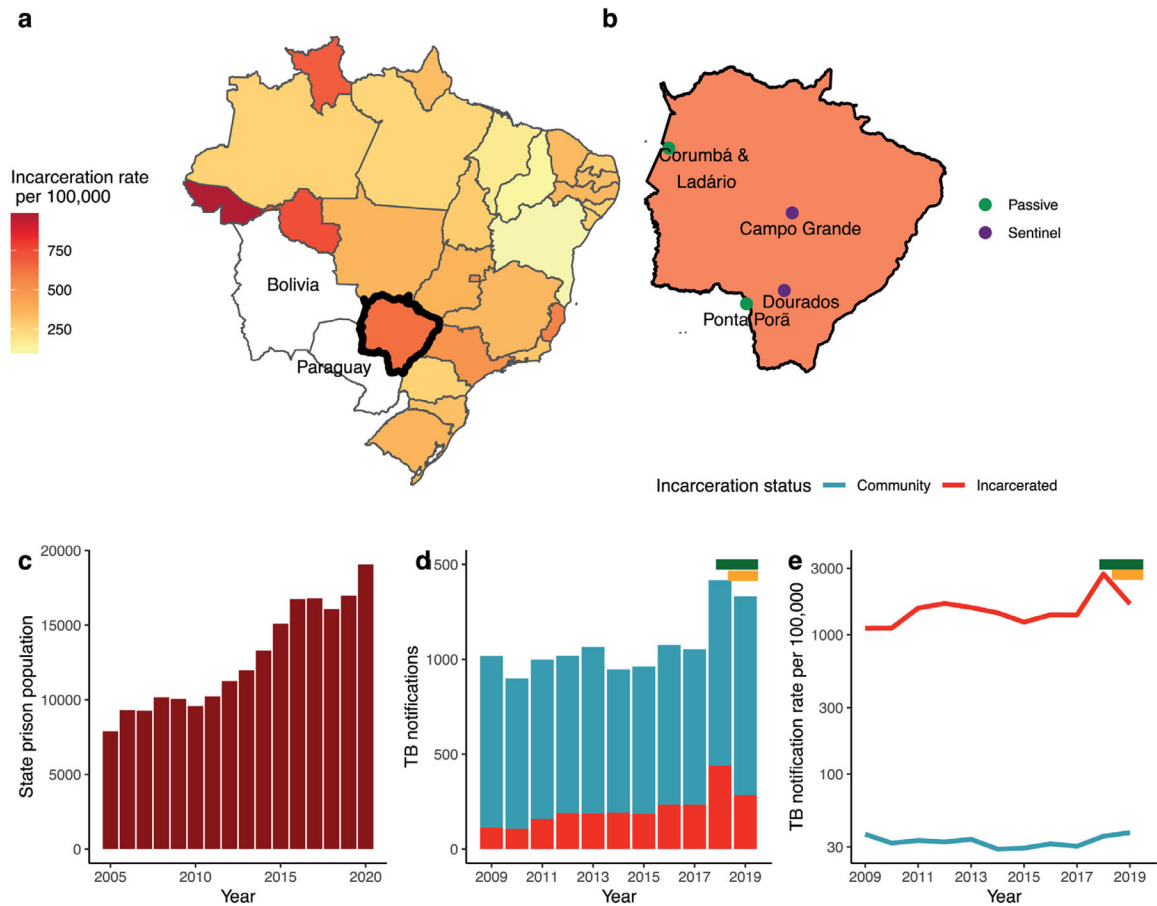
We searched PubMed with the key terms “tuberculosis”, “transmission”, “prison or jail”, and “community” for articles published in English before April 9, 2021. This search identified 96 articles, 10 of which found *M. tuberculosis* genotypes were shared between prisons or jails and adjacent communities, suggesting recent transmission. Several studies from different settings additionally documented the extremely high incidence and/or prevalence of tuberculosis in prisons compared to surrounding communities. We additionally searched SciELO with the key terms “tuberculose”, “prisão or presídio or cadeia”, and “transmissão” for articles published in Portuguese or in Portuguese or Spanish language journals before April 9, 2021. This search identified 12 articles, one of which took a molecular epidemiology approach to identify recent transmission of multi-drug resistant tuberculosis within a Brazilian prison. We did not find studies that quantify the role of prison spillover on community tuberculosis burden.

### Added value of this study

We integrated genomic, epidemiological, and incarceration data to reconstruct tuberculosis transmission patterns in Central West, Brazil, a setting with among the highest incarceration rates in the world. We found that *M. tuberculosis* isolates from inside and outside prisons were closely evolutionarily related and often identical. We identified frequent spillover events, transmission from prisons to the surrounding community. In addition to serving as reservoirs of infection, we found that through the frequent transfers of incarcerated people between prisons, prisons disseminate *M. tuberculosis* infection risk across prison networks and across the state. Our results strongly suggest that prisons play an outsized role in driving the tuberculosis epidemic in Mato Grosso do Sul; the extremely high risk of tuberculosis created by prison environments is not confined within prisons, but extends into the neighboring communities.

### Implications of all the available evidence

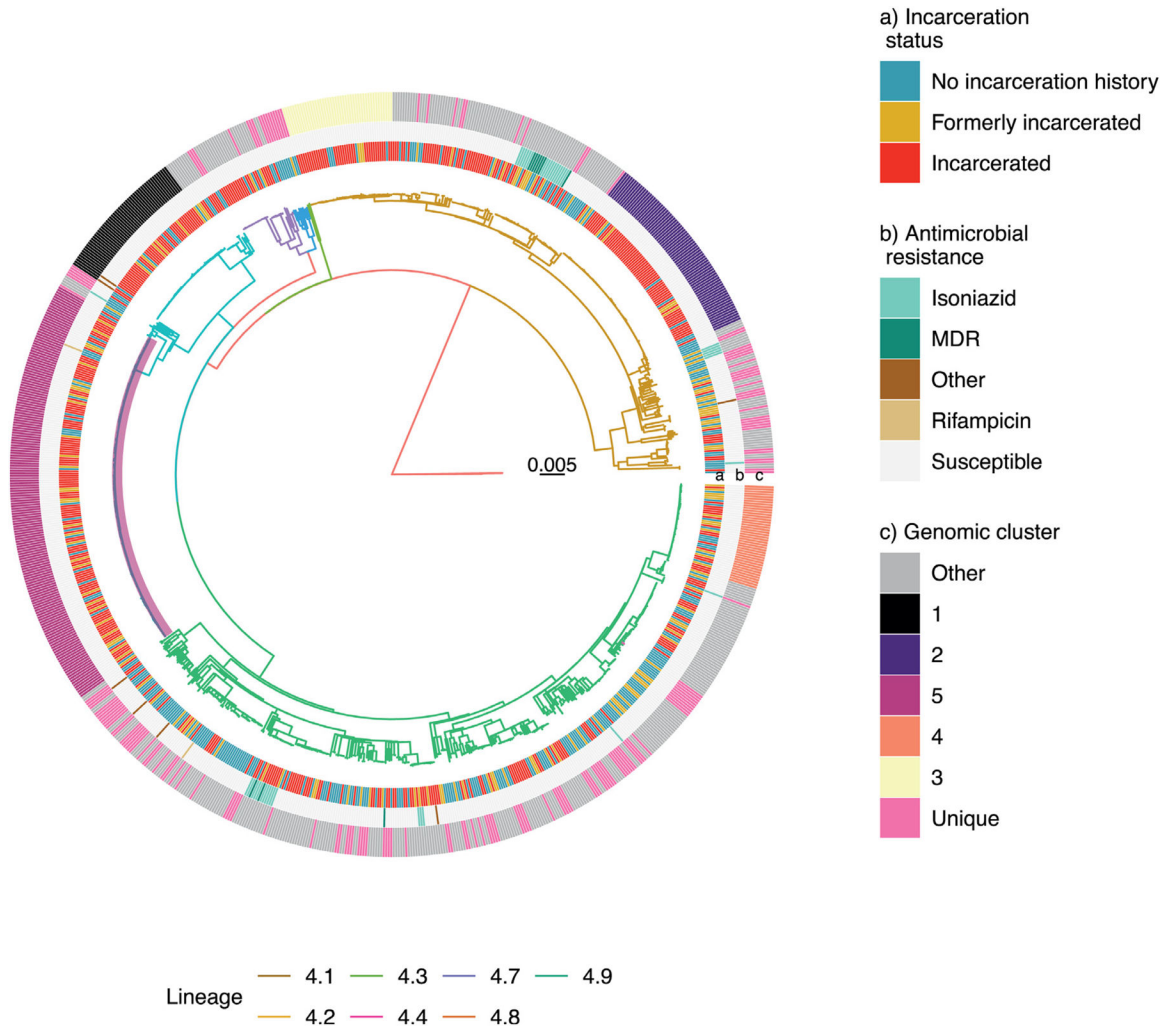
Our findings of the outsized role of prisons in driving community tuberculosis transmission illustrates how high transmission rates in small core populations or institutions can undermine community-level tuberculosis control. Addressing tuberculosis in carceral settings is critical for protecting the health and rights of incarcerated individuals and for achieving successful tuberculosis control in the broader population.



**Figure 1. Tuberculosis is increasingly concentrated within prisons in Mato Grosso do Sul, Brazil.**

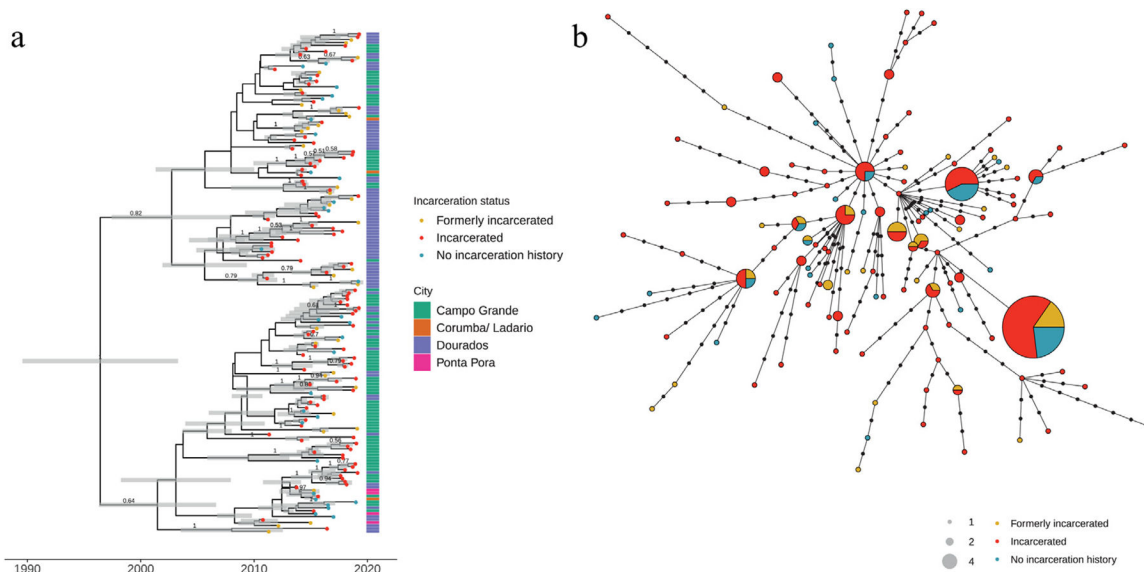
(a) Map of Brazil with states colored by the 2019 incarceration rate with Mato Grosso do Sul outlined in black. (b) Map of Mato Grosso do Sul state with the cities with passive surveillance only and where active surveillance was additionally conducted within prisons. (c) The incarcerated population size in Mato Grosso do Sul grew by 142% from 2005 to 2020, increasing from 7891 to 19,065 people. (d) Mato Grosso do Sul state’s annual new and retreatment tuberculosis notifications, colored by incarceration status and (e) the notification rate per 100,000 people for the incarcerated population and non-incarcerated populations from 2009 to 2019. The y-axis is log-scaled. The mean tuberculosis notification rate was more than 46 times greater among the incarcerated population compared to the non-incarcerated population. The green and yellow bars (d) and (e) indicate the period of active surveillance in prisons in Campo Grande and Dourados, respectively.





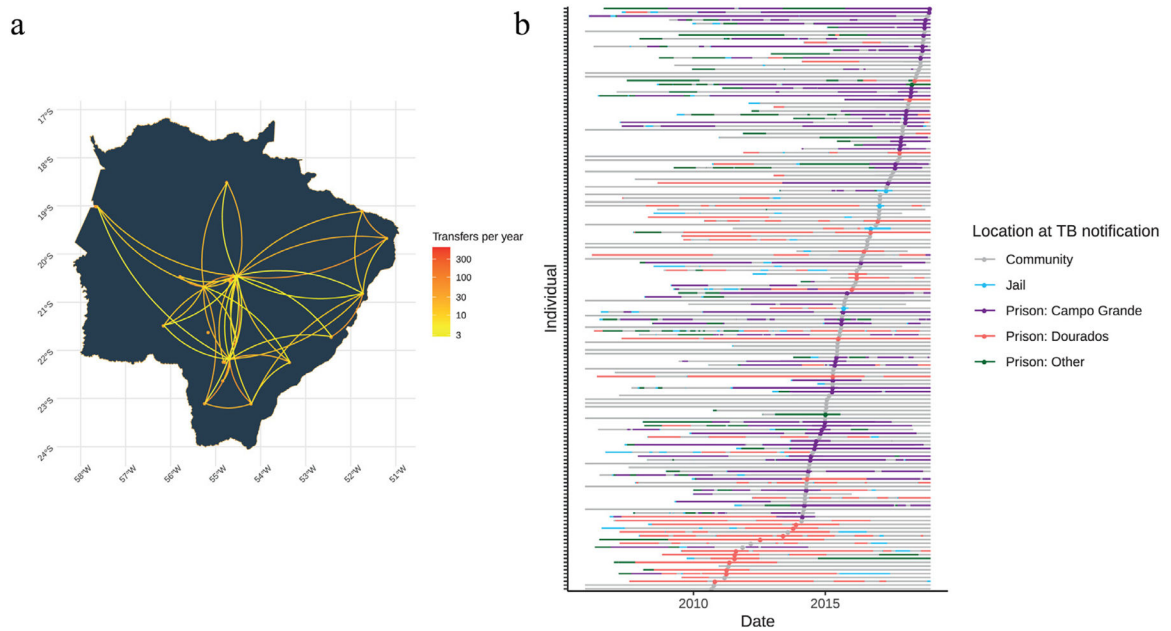
**Figure 2. *M. tuberculosis* isolates from incarcerated and non-incarcerated people are closely related in Mato Grosso do Sul, Brazil.**

A maximum likelihood phylogeny of 932 tuberculosis isolates from Lineage 4 inferred from a multiple sequence alignment of 19,753 SNPs and rooted on two Lineage 1 isolates from this study. Branch lengths are in units of substitutions per site. Branches are colored by sub-lineage. From the inside, rings are colored by incarceration status at time of TB notification, antimicrobial resistance prediction, and genomic cluster (Methods). “Other” resistance indicates resistance to at least one antibiotic other than isoniazid or rifampicin. The clade containing the largest predicted genomic cluster (Cluster 5), including 170 isolates, is highlighted in purple.



**Figure 3. A single, recently emerged *M. tuberculosis* clone spans prisons and the community across the major cities Mato Grosso do Sul state.**

(a) A Bayesian time-calibrated phylogeny of the largest sampled genomic cluster of 170 isolates. The cluster emerged approximately 23 years before the most recently sampled isolate, with a most recent common ancestor in 1996 (95% HPD: 1989–2003). Tip point color indicates patient’s incarceration status at the time of tuberculosis notification. Annotation bar colors indicate city. Grey error bars indicate the 95% Bayesian highest posterior density intervals for node date. Clade posterior support values are shown on the middle of branches for clades with posterior support > 0.5. (b) A haplotype network of the single largest genomic cluster, including 170 isolates. Nodes represent unique haplotypes and are scaled to number of isolates. Points along branches indicate SNP distances between haplotypes. Node color indicates incarceration status at the time of diagnosis.



**Figure 4. Prisons disseminate tuberculosis through the frequent movement of people between prisons.**

(a) Incarceration locates people within highly connected contact networks. Map of Mato Grosso do Sul state with orange points indicating the 21 prisons which most frequently transfer people between prisons. Edges are colored by the mean number of yearly transfers made between prisons from 2015 to 2018, including transfers in both directions. Edges were drawn between prison pairs with more than 10 transfers in the study period. Transfers between prisons within a city are not shown. (b) An *M. tuberculosis* clone (depicted in Figure 3) spreads throughout multiple prisons, jails, and the community. Each line represents an individual patient infected with an isolate in the largest genomic cluster, ordered by notification date. Line color indicates patient location, determined by matching patient names with the state incarceration database (Methods). Points indicate TB notification dates; point color indicates patient location at the time of TB notification. Individuals with notification dates after December 31, 2018 are not shown.