- 1 **Full Title:** The role of prisons in disseminating tuberculosis in Brazil: a genomic epidemiology study
- 2 Short Title: Prison dissemination of tuberculosis
- 3 Authors:
- 4 Katharine S. Walter PhD<sup>1</sup>, Paulo César Pereira dos Santos BS<sup>2</sup>, Thais Oliveira Gonçalves BS<sup>3</sup>, Brunna
- 5 Oliveira da Silva BS<sup>2</sup>, Andrea da Silva Santos PhD<sup>2</sup>, Alessandra de Cássia Leite MSc<sup>2</sup>, Alessandra Moura
- 6 da Silva BA<sup>4</sup>, Flora Martinez Figueira Moreira MSc<sup>2</sup>, Roberto Dias de Oliveira PhD<sup>5</sup>, Everton Ferreira
- 7 Lemos PhD<sup>4</sup>, Eunice Cunha PhD<sup>3</sup>, Yiran E. Liu BS<sup>1,6</sup>, Albert Ko MD<sup>7</sup>, Caroline Colijn PhD<sup>8</sup>, Ted Cohen
- 8 MD<sup>7</sup>, Barun Mathema PhD<sup>9</sup>, Julio Croda MD<sup>4,7,10</sup><sup>†</sup>, Jason R. Andrews MD<sup>1</sup><sup>†</sup>
- 9 †These authors contributed equally to this work.
- <sup>1</sup>Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine;
- 11 Stanford, United States.
- <sup>2</sup> Health Sciences Research Laboratory, Federal University of Grande Dourados; Dourados, Brazil.
- <sup>13</sup> Laboratory of Bacteriology, Central Laboratory of Mato Grosso do Sul; Campo Grande, Brazil.
- <sup>4</sup> School of Medicine, Federal University of Mato Grosso do Sul, School of Medicine; Campo Grande,
- 15 Brazil.
- <sup>5</sup> Nursing Course, State University of Mato Grosso do Sul; Dourados, Brazil.
- 17 <sup>6</sup>Cancer Biology Graduate Program, Stanford University School of Medicine; Stanford, United States.
- 18 <sup>7</sup>Department of Epidemiology of Microbial Diseases, Yale School of Public Health; New Haven, United
- 19 States.
- <sup>8</sup>Department of Mathematics, Simon Fraser University; Burnaby, Canada.
- <sup>9</sup>Department of Epidemiology, Columbia University Mailman School of Public Health; New York,
- 22 United States.
- <sup>10</sup> Mato Grosso do Sul Office, Oswaldo Cruz Foundation; Campo Grande, Brazil.
- 24 Corresponding Author: <a href="https://www.kwalter@stanford.edu">kwalter@stanford.edu</a> (KSW)
- 25 Word count: 4585

It is made available under a CC-BY-NC-ND 4.0 International license .

#### 26 Abstract

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

30 Methods and Findings: To quantify the role of prisons in driving wider community *M. tuberculosis* 

31 transmission, we conducted prospective genomic surveillance in Central West Brazil from 2014 to 2019.

32 We whole genome sequenced 1,152 *M. tuberculosis* isolates collected during active and passive

33 surveillance inside and outside prisons and linked genomes to detailed incarceration histories. We applied

34 multiple phylogenetic and genomic clustering approaches and inferred timed transmission trees. We

35 found that *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

37 with no incarceration history also included individuals with a recent history of incarceration. Among

38 cases in individuals with no incarceration history, 50.6% (162/320) were in clusters that included

39 individuals with recent incarceration history, suggesting that transmission chains often span prisons and

40 communities. We identified a minimum of 18 highly probable spillover events, *M. tuberculosis* 

41 transmission from people with a recent incarceration history to people with no prior history of

42 incarceration, occurring in the state's four largest cities and across sampling years. We additionally found

43 that frequent transfers of people between the state's prisons creates a highly connected prison network

44 that likely disseminates *M. tuberculosis* across the state.

45 Conclusions: We developed a framework for measuring spillover from high-incidence environments to
46 surrounding communities by integrating genomic and spatial information to reconstruct tuberculosis
47 transmission chains. Our findings indicate that, in this setting, prisons serve not only as disease reservoirs,
48 but also disseminate *M. tuberculosis* across highly connected prison networks, both amplifying and

but also disseminate *m. tabereatosis* deross inginy connected prison networks, both ampinying

49 propagating *M. tuberculosis* risk in surrounding communities.

50 Abstract word count: 294

51 Key words: tuberculosis, epidemiology, genomics, disease reservoir, spillover, prisons

It is made available under a CC-BY-NC-ND 4.0 International license .

#### 52 Introduction

53	Tuberculosis remains one of the leading causes of death by an infectious disease and led to 1.4
54	million deaths in 2019.[1] Locating where M. tuberculosis transmission occurs is critical for developing
55	public health interventions to block transmission and prevent future infections. Yet in high-incidence
56	settings, where most tuberculosis is transmitted remains unknown.[2–7] Built environments often shape
57	M. tuberculosis transmission risk, and high-transmission settings like prisons, hospitals, and mines may
58	have epidemiological impacts that transcend their walls[8-10], serving as institutional amplifiers of
59	infection.[11,12] In particular, the conditions of detention, including severe overcrowding, poor
60	ventilation, poor nutrition, and limited access to healthcare, put incarcerated people at elevated risk of
61	many infectious diseases, including tuberculosis.[8–10]
62	Globally, incidence of tuberculosis in incarcerated populations is more than ten times that in the
63	general population.[9,13] The role of prisons in driving community tuberculosis epidemics is of particular
64	concern in Central and South America, where the incarcerated population has increased by over 200%
65	since 2000 and tuberculosis cases among the incarcerated population have increased by over 269% [14].
66	Alarmingly, increasing tuberculosis notifications within prisons more than offset declines in tuberculosis
67	outside prisons across the region[14].
68	The full impact of increasing incarceration rates on the tuberculosis epidemic in the Americas and
69	elsewhere is underestimated by total incident cases within prisons. Many people are infected in prison,
70	but only diagnosed later.[12,15,16] Further, spillover may occur when incarcerated people, prison staff, or
71	visitors are infected within high-incidence prison environments and transmit M. tuberculosis onwards in
72	communities outside prison. Previous studies have found that the excess risk of tuberculosis within
73	prisons extends to surrounding communities[17,18] and that M. tuberculosis genotypes responsible for
74	jail and prison outbreaks are also found in the surrounding communities[15,17,19,20]. Because
75	transmission is rarely observed, however, it is often difficult to link community cases to transmission
76	from prisons and previous studies have largely been associative.

77	Pathogen genomes can be powerfully harnessed to infer high-resolution evolutionary and
78	transmission histories, including who-infected-whom.[21] One previous study in Georgia used M.
79	tuberculosis genomes to identify frequent transmission of multi-drug resistant tuberculosis from prisons
80	to surrounding communities.[22] However, studies have not yet combined M. tuberculosis genomic,
81	epidemiologic, and location data to reconstruct transmission linkages between potential institutional
82	amplifiers and the community. To quantify the contribution of prison spillover to tuberculosis
83	transmission in the community, we conducted prospective genomic surveillance for tuberculosis in the
84	Central West Brazilian state of Mato Grosso do Sul. We integrated M. tuberculosis genomes with detailed
85	incarceration histories and took multiple phylogenetic and genomic clustering approaches to reconstruct
86	spillover events, transmission of tuberculosis from prisons to the community.
87	
88	Methods
89	Study design
90	We conducted population-based tuberculosis surveillance in the state of Mato Grosso do Sul in
91	Central West Brazil from January 2014 through May 2019. Surveillance included active screening in
92	three of the largest prisons in the state as well as ongoing passive surveillance focused on the two largest
93	cities in the state, and three cities at the state's border with Paraguay and Bolivia (Fig. 1a,b; S1 Text). All
94	participants provided written consent, and this study was conducted with the approval of the Research
95	Ethics Committee from the Federal University of Grande Dourados, Federal University of Mato Grosso
96	do Sul and National Research Ethics Committee (CONEP) (CAAE 37237814.4.0000.5160,
97	2676613.3.1001.5160, and 26620619.6.0000.0021) and Stanford University Institutional Review Board
98	(IRB-40285) ( <b>S1 Text</b> ).
99	Incarceration history
100	We obtained permission from the Mato Grosso do Sul state prison administration agency,
101	Agência Estadual de Administração do Sistema Penitenciário, to access a database of all prison entries,

102 exits, and transfers within the criminal justice record system from January 1, 2005 through December 31,

103 2018 (**S1 Text**).

104

105 Laboratory diagnosis and culture

106 All sputum specimens collected in the participating laboratories were examined by microscopy or

107 GeneXpert® MTB/RIF, decontaminated with sodium hydroxide (NaOH), and inoculated in Ogawa-

108 Kudoh culture medium. Cultures were incubated at 37°C for up to eight weeks and checked weekly for

109 visible colonies at the participating laboratories. Microbial species determination was made with the

110 MPT64 protein detection-based immunochromatographic rapid test (SD Bioline Kit, Standard

111 Diagnostics, Inc., Korea). DNA was extracted from cultured isolates with the manual CTAB (Cetyl

trimethylammonium bromide) method and whole genome sequenced on an Illumina NextSeq (2 x 151-

113 bp).

114

## 115 Whole genome sequencing

116 We sequenced whole *M. tuberculosis* genomes from cultured isolates on an Illumina NextSeq (2

117 x 151-bp). Sequence data is available on the Sequence Read Archive, in BioProject PRJNA671770. We

118 applied variant calling methods closely following those described in Menardo et al.[23] to be consistent

119 with the methods used for molecular clock estimation (S1 Text).

120

#### 121 Phylogenetic and Bayesian evolutionary analysis

We fit a maximum likelihood tree with RAxML-ng 1.0.1 and rooted the tree on two Lineage 1 isolates from this study.[24] 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[25] with a strict clock, constant coalescent

125 population size model using tuberculosis notification dates to calibrate tips (S1 Text).

126

# 127 Transmission inference

It is made available under a CC-BY-NC-ND 4.0 International license .

128	M. tuberculosis phylogenetic trees represent patterns of evolutionary relatedness between the
129	consensus bacterial genomes sampled from different individuals. Because most outbreaks are
130	incompletely sampled and individuals may be infected with diverse populations of <i>M. tuberculosis</i> ,
131	phylogenies do not represent the underlying history of transmission.[21] We used TransPhylo[21] to infer
132	transmission linkages-including unsampled hosts-that were consistent with the underlying timed
133	phylogenies and epidemiology of tuberculosis (i.e. generation time and sampling intervals). TransPhylo
134	samples transmission trees that are consistent with timed phylogenetic trees and epidemiological priors
135	governing the distribution of likely <i>M. tuberculosis</i> generation times (time between subsequent infections)
136	and sampling time (time from an individual's infection to diagnosis) (S1 Text).
137	We then combined transmission trees with incarceration histories to identify spillover events from
138	prisons to surrounding communities. We defined spillover events as transmission that occurred from an
139	individual with a history of current or former incarceration to the community, including individuals with
140	no incarceration history and prior incarceration. We considered transmission probabilities >50% as
141	evidence of spillover and included spillover events no more than two degrees removed from prisons.
142	TransPhylo infers transmission linkages including transmission to and from unsampled hosts and thus
143	does not require that all transmission events be observed. Therefore, we are able to quantify the minimum
144	number of transmission events, but not estimate the total role of spillover on community transmission.
145	
146	Results
147	Tuberculosis and incarceration trends in Mato Grosso do Sul
148	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% (7,891 to 16,976) over the same period (Fig. 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

It is made available under a CC-BY-NC-ND 4.0 International license .

153 100,000) in 2019,[26] the notification rate within prisons was 43 times as high (1,666 per 100,000), again
154 similar to Brazil's national notification rate within prisons (1,596 per 100,000) in 2019 (Fig. 1d,e).

155

## 156 Fig 1. Tuberculosis is increasingly concentrated within prisons in Mato Grosso do Sul, Brazil. (a) 157 Map of Brazil with states colored by the 2019 incarceration rate with Mato Grosso do Sul outlined in 158 black. (b) Map of Mato Grosso do Sul state with the cities with passive surveillance only and where 159 active surveillance was additionally conducted within prisons. (c) The incarcerated population size in 160 Mato Grosso do Sul grew by 142% from 2005 to 2020, increasing from 7,891 to 19,065 people. (d) Mato 161 Grosso do Sul state's annual new and retreatment tuberculosis notifications, colored by incarceration 162 status and (e) the notification rate per 100,000 people for the incarcerated population and non-163 incarcerated populations from 2009 to 2019. The y-axis is log-scaled. The mean tuberculosis notification 164 rate was more than 46 times greater among the incarcerated population compared to the non-incarcerated 165 population. The green and yellow bars (d) and (e) indicate the period of active surveillance in prisons in 166 Campo Grande and Dourados respectively. 167 168 Population-based genomic surveillance across Mato Grosso do Sul state 169 From 2014 to 2019, 3,491 new and retreatment cases of tuberculosis were reported in the Campo 170 Grande and Dourados, the two largest cities in Mato Grosso do Sul, Brazil. Of these, 1,249 had positive 171 cultures and we sequenced 787 M. tuberculosis isolates. We sequenced additional isolates from three 172 other cities in the state and from earlier years in Dourados for a total of 1,090 isolates from unique 173 tuberculosis notifications (Fig. 1b). Whole genome sequences (WGS) for a total of 1,043 isolates met our 174 coverage and quality criteria (Methods). We excluded 10% (108 of 1043) of isolates with mixed 175 infection, resulting in 935 high-quality genomes from 935 unique tuberculosis episodes from 918 176 individuals.

177 Of the 935 isolates in our analyses, 50% (465 of 935) were from patients who were incarcerated 178 at the time of tuberculosis notification; 16% (150 of 935) were formerly incarcerated; and 34% (320 of

It is made available under a CC-BY-NC-ND 4.0 International license .

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 E1. 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 (Fig. 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 (Fig. 2).

185

#### 186 Fig 2. *M. tuberculosis* isolates from incarcerated and non-incarcerated people are closely related in

187 Mato Grosso do Sul, Brazil. A maximum likelihood phylogeny of 932 tuberculosis isolates from

188 Lineage 4 inferred from a multiple sequence alignment of 19,753 SNPs and rooted on two Lineage 1

189 isolates from this study. Branch lengths are in units of substitutions per site. Branches are colored by sub-

190 lineage. From the inside, rings are colored by incarceration status, antimicrobial resistance prediction, and

191 genomic cluster (Methods). "Other" resistance indicates resistance to at least one antibiotic other than

192 isoniazid or rifampicin. The clade containing the largest predicted genomic cluster (Cluster 5), including

- 193 170 isolates, is highlighted in purple.
- 194

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

196 In a maximum likelihood phylogeny (Fig. 2), M. tuberculosis isolates sampled from incarcerated 197 and non-incarcerated people are distributed across the tree and do not form distinct clades, indicating 198 recent shared evolutionary history of isolates sampled from prisons and the community. Terminal branch 199 lengths, a proxy for the degree of recent transmission, [27] were significantly longer for isolates from 200 people with no incarceration history at the time of tuberculosis notification (9.4 x 10<sup>-4</sup> substitutions/site) 201 compared to isolates from incarcerated individuals (2.2 x  $10^{-4}$ ; p < 0.001) or formerly incarcerated 202 individuals ( $3.8 \times 10^{-4}$ ; p=0.019), suggesting that individuals with an incarceration history were more 203 likely to be recently infected, although this could alternatively reflect a higher rate of genomic sampling 204 (Fig. S1; permutation test in S1 Text).

It is made available under a CC-BY-NC-ND 4.0 International license .

205 We tested for phylogenetic signal in incarceration status of sampled tuberculosis patients— 206 whether incarceration status was distributed randomly across the tree—with Fritz's D[28], a test 207 "clumping" in observed traits. The D statistic is equal to 1 if the trait is distributed randomly across tree 208 tips with respect to the phylogeny and 0 if the trait is consistent with a pattern of clumping consistent with 209 Brownian motion of the trait along the phylogenetic tree. 210 We predicted that if prison and community epidemics were distinct, incarceration status would 211 exhibit an extremely clumped pattern on the *M. tuberculosis* phylogeny, with separate clades including 212 samples from incarcerated people and community members. Conversely, if transmission between prisons 213 and the community occurred just as frequently as transmission within prisons, there would be a random 214 pattern of incarceration status across the phylogeny. We found that incarceration status had a moderate 215 phylogenetic signal (Fritz's D=0.59), indicating a pattern of shared evolutionary history of isolates 216 collected from individuals with and without an incarceration history, with some degree of "clumping" 217 with respect to incarceration status.

218

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

220 Patterns of genomic relatedness are often used to infer potential recent transmission; M. 221 tuberculosis isolates that are more closely genetically related are hypothesized to be more likely linked 222 through recent transmission[29]. To identify potential transmission clusters, we applied a commonly used 223 12-SNP threshold [29], including all isolates within the threshold distance of at least one other clustered 224 isolate. Eighty-three percent (777 of 935) of the isolates fell into 84 genomic clusters (each including 2 to 225 170 isolates), providing evidence that notifications were largely attributable to recent, local transmission 226 rather than travel-associated importation or re-activation of genetically distinct latent infections (Fig. 2). 227 Isolates from incarcerated people were more frequently clustered (93.3%, 434/465), than those from 228 formerly incarcerated (86.0%, 129 of 150; p <0.0001), or never incarcerated people (66.9%, 214 of 320; p 229 <0.0001), again suggesting more recent transmission within prisons.

It is made available under a CC-BY-NC-ND 4.0 International license .

230	We predicted that if prison and community-associated epidemics were distinct, isolates from the
231	community would be most closely related to and cluster with other isolates from the community and vice
232	versa. The majority (70.8%; 46/65) of genomic clusters including people with no incarceration history
233	also included individuals with a recent history of incarceration. Among cases in individuals with no
234	incarceration history, 50.6% (162/320) were in clusters that included individuals with recent incarceration
235	history, suggesting that transmission chains often span prisons and communities.
236	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[30]. With

this threshold, of the 80 clusters, 56 included isolates from people with no incarceration history and

239 78.6% (44 of 56) also included isolates from people with an incarceration history. Among all tuberculosis

cases occurring among individuals with no incarceration history, 46.9% (150 of 320) were in 5-SNP

241 clusters that included individuals with incarceration history.

242

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

244 The largest potential transmission cluster, including 170 isolates sampled from September 2010 245 to April 2019, was distributed across the state, including cases found across state's two largest cities, 246 Campo Grande and Dourados, as well as the smaller cities Corumbá (on the Bolivian border) and Ponta 247 Porã (on the Paraguayan border) (Fig. 1b, Fig. 3a). The cluster had a most recent common ancestor in 248 1996 (95% HPD: 1989-2003; Fig. 3a), indicating that it had circulated for approximately 23 years before 249 the most recent samples were collected. 103 isolates were from people incarcerated at the time of 250 diagnosis, 35 from people formerly incarcerated, and 32 from people with no incarceration history at the 251 time of tuberculosis notification.

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 (Fig. 3b). The three largest haplotypes (identical sequences), including 14, 9, and 8 samples, were dominated by isolates from

It is made available under a CC-BY-NC-ND 4.0 International license .

256 people who were incarcerated. Each haplotype also included isolates from people who were formerly

257 incarcerated and from people with no incarceration history, indicating very recent transmission between

258 incarcerated individuals and individuals in the community.

259

260 Fig 3. A single, recently emerged *M. tuberculosis* clone spans prisons and the community across the 261 major cities Mato Grosso do Sul state. (a) A Bayesian time-calibrated phylogeny of the largest sampled 262 genomic cluster of 170 isolates. The cluster emerged approximately 23 years before the most recently 263 sampled isolate, with a most recent common ancestor in 1996 (95% HPD: 1989-2003). Tip point color 264 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. 265 266 Clade posterior support values are shown on the middle of branches for clades with posterior support > 267 0.5. (b) A haplotype network of the single largest genomic cluster, including 170 isolates. Nodes 268 represent unique haplotypes and are scaled to size. Points along branches indicate SNP distances between 269 haplotypes. Node color indicates incarceration status at the time of diagnosis. 270 271 Prison networks disseminate *M. tuberculosis* across space 272 We hypothesized that the criminal justice system's frequent transfer of people between prisons 273 and jails could disseminate TB across the state. We analyzed movements in the state's incarceration 274 database and found that the average duration of incarceration was 230 days in a prison and 25 days in a

275 jail. The incarceration database documents 7,982 mean yearly transfers between prisons (including closed

and semi-open prisons) from 2015-2018 in a prison population of 17,221 in August 2018, including

277 frequent transfers between cities across the state (Fig. 4a).

To investigate the role of the criminal justice system in disseminating infection, we tracked the spread of the largest sampled *M. tuberculosis* cluster (Fig. 3) across prisons in Mato Grosso do Sul (Fig. 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

It is made available under a CC-BY-NC-ND 4.0 International license .

exported to prisons and jails in other cities. 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.

285

#### Fig 4. Prisons disseminate tuberculosis through the frequent movement of people between prisons.

287 (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 2015-2018,

290 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

292 (depicted in Figure 3) spreads throughout multiple prisons, jails, and the community. Each line represents

293 an individual patient infected with an isolate in the largest genomic cluster, ordered by notification date.

294 Line color indicates patient location, determined by matching patient names with the state incarceration

295 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.

297

#### 298 Transmission trees reveal frequent spillover of infection

299 We further investigated recent transmission in the state by inferring probabilistic transmission 300 trees representing who-infected-whom for the 84 identified M. tuberculosis genomic clusters with 301 TransPhylo[21]. Because our sampling of the TB epidemic in Mato Grosso do Sul was incomplete and we 302 did not observe the majority of transmission events resulting in sampled cases, we quantified the 303 minimum number of recent infections attributable to prison spillover rather than the total role of spillover 304 on the community epidemic. We identified a minimum of 16.8% (18 of 107) infections among never 305 incarcerated individuals were directly attributable to observed spillover from prisons. Observed spillover 306 of infection to never or formerly incarcerated individuals occurred across the state in the four largest cities 307 sampled (Campo Grande, Dourados, Ponta Pora, and Corumba) and every year, from 2014-2019.

It is made available under a CC-BY-NC-ND 4.0 International license .

308 Together, these findings indicate that spillover of *M. tuberculosis* infection from prisons to surrounding 309 communities is frequent. The result is that the 0.84% of the state's population incarcerated at a given time 310 (22,706 of 2,713,147 in 2017) or 2.6% currently or formerly incarcerated at a given time (71,703 of 311 2,713,147), have a disproportionate effect on TB transmission outside of prisons. 312 313 Transmission inferences are sensitive to epidemiological priors 314 We conducted a sensitivity analysis to determine the effect of epidemiological priors (Table S2) 315 on transmission inferences for the largest genomic cluster (170 isolates). Transmission probabilities 316 inferred for pairs of individuals were closely correlated across generation time priors (Pearson's r: 0.94 – 317 0.98, Fig. S3) and sampling time priors (Pearson's r: 0.84 - 0.98, Fig. S4), although the total number of 318 observed transmission events was strongly influenced by both priors (Figure S5). This suggests that, as 319 expected, epidemiological priors impact the magnitude of inferred transmission probabilities rather than 320 the identity of predicted transmission pairs. As expected, the minimum number of observed spillover 321 events in the largest cluster were sensitive to epidemiological priors, when transmission was defined as a

transmission probability greater than or equal to 50%, and ranged from 0 (short sampling and generation

times) to 28 (medium generation time and very long sampling time) (Fig. S6).

324

#### 325 Discussion

326 Here, we linked intensive *M. tuberculosis* genomic surveillance to detailed individual-level 327 incarceration histories, allowing us to reconstruct high-resolution population-wide tuberculosis 328 transmission in Central West Brazil, a setting with an extreme disparity in tuberculosis incidence between 329 prisons and surrounding communities. We found that *M. tuberculosis* sequences from incarcerated and 330 non-incarcerated individuals were closely evolutionarily related and sometimes identical. Locally 331 circulating *M. tuberculosis* clones spanned prisons and the community, and the majority of genomic 332 clusters included individuals with and without a recent history of incarceration. In reconstructed 333 transmission trees, we found evidence of frequent spillover of infection in all the major cities in the state

It is made available under a CC-BY-NC-ND 4.0 International license .

and across sampling years, indicating that spillover is not the effect of a single prison but a widespread

335 phenomenon. We found that frequent transfers of people between prisons in Mato Grosso do Sul, Brazil

336 creates a closely connected prison contact network that likely facilitates the spread of *M. tuberculosis* 

337 between prisons and may disseminate infection risk across the state.

338 While alarming disparities in tuberculosis incidence between prisons and surrounding 339 communities are well-documented, [9,13] it has been more challenging to directly link tuberculosis cases 340 in the community to spillover from prisons. Previous studies of the broader effects of prisons on general 341 TB epidemics have largely been associative. Here, we describe a probabilistic framework for identifying 342 spillover events that can be applied more generally to measure the impact of high-transmission 343 environments such as mines or hospitals on tuberculosis transmission in the general population. Our 344 finding of frequent and widespread spillover events demonstrates the outsized effect of prison 345 environments on community-wide *M. tuberculosis* transmission. This also suggests that the introduction 346 of *M. tuberculosis* lineages or resistance mutations into prisons is likely to be a source of community 347 epidemics. Our finding of the likely role of prison transfers in disseminating *M. tuberculosis* highlights an 348 additional role that carceral systems may play in propagating infectious diseases both within highly-349 connected prison systems and across space.

350 The rapid rise in incarceration rates in many countries[31] and stark global disparities in 351 tuberculosis incidence inside compared to outside prisons[8,9,13,32] suggests that spillover of infection 352 may contribute to community TB transmission outside of Brazil. A study of multi-drug resistant (MDR) 353 TB transmission found that up to 31% of MDR TB in the country of Georgia was directly or indirectly 354 linked to prisons.[22] Our transmission inference framework is conservative in that it identifies a 355 minimum contribution of spillover to total community transmission, and therefore, is consistent with 356 these findings. The role of prisons and other detention centers as disease reservoirs is not unique to 357 tuberculosis. Previous studies have identified spillover of meningococcal disease[33] and SARS-CoV-358 2[34] from prisons to the community. Prison transfers have also spread of SARS-CoV-2 across prisons in 359 California, resulting in widespread outbreaks.[35]

360	The frequent spillover of infection from prisons to surrounding communities indicates that for TB
361	control to be achieved in Central West Brazil, TB control efforts must focus on reducing the
362	extraordinarily high transmission risk within prisons. Tuberculosis control programs should expand
363	routine active screenings during[12,36] and following incarceration so that cases can be diagnosed early
364	and linked to treatment. Individuals leaving prison are at heightened risk of tuberculosis and are a
365	particularly vulnerable population. Preventive therapy is not provided in prisons in Brazil and many other
366	low- and middle-income countries and could represent an effective means for tuberculosis prevention,
367	particularly among individuals exiting prisons for community environments where the risk of reinfection
368	is lower. More broadly, improving the unsanitary, inhumane conditions of incarceration and expanding
369	access to primary healthcare and nutrition are critical.
370	Reducing this excess burden of disease will require work that extends beyond biomedical
371	interventions. The most direct way to mitigate the excess tuberculosis risk created by prisons is for
372	governments to rapidly reduce incarcerated populations, as advocated by the American Public Health
373	Association, to reduce the broader harms associated with incarceration.[37] Prison releases have already
374	begun as a means of reducing infectious disease transmission risk. [38] In an attempt to reduce the risk of
375	COVID-19 in prisons, for example, Brazil released more than thirty thousand people by mid-2020.[39]
376	Our study has several limitations. First, our study highlights the uncertainty inherent in
377	transmission inferences drawn from genomic surveillance data, which almost always constitutes an
378	incomplete sample of population-wide infections. We addressed the incomplete sampling of $M$ .
379	tuberculosis genomes in two ways. First, we took a probabilistic transmission inference approach that
380	accounts for uncertainty in transmission linkages and also includes unsampled individuals contributing to
381	transmission.[21] Second, in order to collect the most comprehensive set of <i>M. tuberculosis</i> genomes
382	possible, we combined isolates collected during active mass screening within prisons and passive routine
383	surveillance in the community. The result was that isolates from prisons are overrepresented in our
384	genomic collection. Because of the over-sampling within prisons, we could not directly quantify the
385	contribution of transmission within prisons and spillover from prisons on overall recent infections in the

It is made available under a CC-BY-NC-ND 4.0 International license .

386 study area. However, by taking a probabilistic transmission inference approach that accounted for 387 uncertainty in transmission linkages as well as unsampled individuals contributing to transmission, we 388 were able to quantify the minimum number of reconstructed spillover events consistent with inferred 389 transmission trees. Because we cannot infer the identity or incarceration history of unsampled people who 390 contribute to transmission, our estimates of the contribution of spillover are conservative and likely 391 underestimate the full impact of spillover on infections in the community. However, our finding of 392 geographically widespread spillover occurring over several years suggests that spillover is common and 393 an important contributor to population-wide *M. tuberculosis* transmission.

394 Additionally, transmission trees reflect information from timed *M. tuberculosis* phylogenies and 395 epidemiological priors and it remains difficult to distinguish genomic sampling proportion from 396 generation time and sampling time distributions. Because we conducted transmission inference in a 397 tuberculosis-endemic setting, rather than in a discrete outbreak, our rate of genomic sampling was lower 398 than in some previous tuberculosis transmission studies. To reduce uncertainty in sampling proportion, 399 we conducted transmission inference on recent subtrees for which we had sampled a greater proportion of 400 isolates. To investigate the influence of epidemiological priors on transmission inferences, we conducted 401 a sensitivity analysis (Table E2) and found that our conclusions about transmission linkages were largely 402 unchanged (Supplementary Text, Figs. S3-S4), although, as expected, the number of transmission events 403 and thus spillover events was dependent on epidemiological priors (Figs. S5-S6).

Finally, we conducted whole genome sequencing of *M. tuberculosis* isolates cultured from
sputum and generated single consensus genomes from each participant, excluding within-host diversity.
We also employed an existing bioinformatic approach to identify *M. tuberculosis* variants and excluded
repetitive genomic regions as is common practice.[23,40] Future studies that leverage within-host *M. tuberculosis* diversity and/or include the diverse PE/PPE genes could provide greater resolution and
reduce the uncertainty of transmission inferences.[41]

In this study, we present genomic evidence that spillover of tuberculosis infection risk fromprisons to surrounding communities is frequent in Central West Brazil. The dramatic expansion of

412	incarceration in recent decades has put an increasing population at extremely high risk of tuberculosis;
413	this risk extends to surrounding communities. Reducing the excess tuberculosis transmission risk within
414	prisons and other detention centers is therefore an urgent public health priority. Our findings additionally
415	highlight the potential for pathogen genomic surveillance to be used to identify other key environments
416	with a disproportionate role in pathogen transmission in order to prioritize public health interventions.
417	
418	Acknowledgments
419	We would like to acknowledge the contributions of AGEPEN, the Mato Grosso do Sul state criminal
420	justice agency; LACEN, the state reference laboratory; and the Mato Grosso do Sul State Health
421	Department. All authors had full access to all the data in the study and had final responsibility for the
422	decision to submit.
423	
424	

It is made available under a CC-BY-NC-ND 4.0 International license .

# 425 References

#### 426

- The World Health Organization, World Health Organization. Global Tuberculosis Report
   2020. Geneva; 2020.
- Churchyard G, Kim P, Shah NS, Rustomjee R, Gandhi N, Mathema B, et al. What We
  Know about Tuberculosis Transmission: An Overview. J Infect Dis. 2017;216: S629–
  S635. doi:10.1093/infdis/jix362
- Auld SC, Sarita Shah N, Mathema B, Brown TS, Ismail N, Omar SV, et al. Extensively
  drug-resistant tuberculosis in South Africa: Genomic evidence supporting transmission in
  communities. Eur Respir J. 2018;52. doi:10.1183/13993003.00246-2018
- 435
  4. Middelkoop K, Mathema B, Myer L, Shashkina E, Whitelaw A, Kaplan G, et al.
  436
  437
  437
  437
  438
  439
  439
  439
  439
  430
  430
  430
  430
  431
  431
  431
  432
  433
  434
  434
  435
  435
  435
  436
  437
  437
  437
  437
  437
  437
  437
  437
  437
  437
  437
  437
  438
  438
  439
  439
  439
  430
  430
  430
  431
  431
  431
  432
  432
  433
  434
  434
  434
  435
  435
  436
  437
  437
  437
  437
  437
  437
  437
  437
  437
  437
  437
  437
  437
  437
  438
  448
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
  449
- Andrews JR, Morrow C, Walensky RP, Wood R. Integrating social contact and
  environmental data in evaluating tuberculosis transmission in a South African township. J
  Infect Dis. 2014;210: 597–603. doi:10.1093/infdis/jiu138
- 441 6. Yates TA, Khan PY, Knight GM, Taylor JG, McHugh TD, Lipman M, et al. The
  442 transmission of Mycobacterium tuberculosis in high burden settings. Lancet Infect Dis.
  443 2016;16: 227–238. doi:10.1016/S1473-3099(15)00499-5
- Andrews JR, Morrow C, Wood R. Modeling the role of public transportation in sustaining
  tuberculosis transmission in South Africa. Am J Epidemiol. 2013;177: 556–561.
  doi:10.1093/aje/kws331
- B. Bolan K, Wirtz AL, Moazen B, Ndeffo-mbah M, Galvani A, Kinner SA, et al. Global
  burden of HIV, viral hepatitis, and tuberculosis in prisoners and detainees. Lancet.
  2016;388: 1089–1102. doi:10.1016/S0140-6736(16)30466-4
- 450 9. Cords O, Martinez L, Warren JL, O'Marr JM, Walter KS, Cohen T, et al. Incidence and
  451 prevalence of tuberculosis in incarcerated populations: a systematic review and meta452 analysis. Lancet Public Heal. 2021. doi:10.1016/S2468-2667(21)00025-6
- Urrego J, Ko AI, Da Silva Santos Carbone A, Paião DSG, Sgarbi RVE, Yeckel CW, et al.
  The impact of ventilation and early diagnosis on tuberculosis transmission in Brazilian
  prisons. Am J Trop Med Hyg. 2015;93: 739–746. doi:10.4269/ajtmh.15-0166
- 456 11. Basu S, Stuckler D, McKee M. Addressing institutional amplifiers in the dynamics and
  457 control of tuberculosis epidemics. Am J Trop Med Hyg. 2011;84: 30–37.
  458 doi:10.4269/ajtmh.2011.10-0472
- Mabud TS, de Lourdes Delgado Alves M, Ko AI, Basu S, Walter KS, Cohen T, et al.
  Evaluating strategies for control of tuberculosis in prisons and prevention of spillover into communities: An observational and modeling study from Brazil. Metcalfe JZ, editor.
  PLoS Med. 2019;16: e1002737. doi:10.1371/journal.pmed.1002737
- Baussano I, Williams BG, Nunn P, Beggiato M, Fedeli U, Scano F. Tuberculosis
  incidence in prisons: A systematic review. PLoS Med. 2010;7: 1–10.
  doi:10.1371/journal.pmed.1000381
- 466 14. Walter KS, Martinez L, Arakaki-Sanchez D, Sequera VG, Estigarribia Sanabria G, Cohen
  467 T, et al. The escalating tuberculosis crisis in central and South American prisons. Lancet.
  468 2021 [cited 10 Apr 2021]. doi:10.1016/s0140-6736(20)32578-2
- 469 15. Sacchi FPC, Praça RM, Tatara MB, Simonsen V, Ferrazoli L, Croda MG, et al. Prisons as
  470 reservoir for community transmission of tuberculosis, brazil. Emerg Infect Dis. 2015;21:

471		452–455. doi:10.3201/eid2103.140896
472	16.	Coker R, McKee M, Atun R, Dimitrova B, Dodonova E, Kuznetsov S, et al. Risk factors
473		for pulmonary tuberculosis in Russia: Case-control study. Br Med J. 2006;332: 85–87.
474		doi:10.1136/bmj.38684.687940.80
475	17.	Warren JL, Grandjean L, Moore DAJ, Lithgow A, Coronel J, Sheen P, et al. Investigating
476		spillover of multidrug-resistant tuberculosis from a prison: a spatial and molecular
477		epidemiological analysis. BMC Med. 2018;16: 122. doi:10.1186/s12916-018-1111-x
478	18.	Jenkins HE, Plesca V, Ciobanu A, Crudu V, Galusca I, Soltan V, et al. Assessing spatial
479		heterogeneity of multidrug-resistant tuberculosis in a high-burden country. Eur Respir J.
480		2013;42: 1291–1301. doi:10.1183/09031936.00111812
481	19.	Séraphin MN, Didelot X, Nolan DJ, May JR, Khan MSR, Murray ER, et al. Genomic
482		investigation of a Mycobacterium tuberculosis outbreak involving prison and community
483		cases in Florida, United States. Am J Trop Med Hyg. 2018;99: 867–874.
484		doi:10.4269/ajtmh.17-0700
485	20.	Jones TF, Woodley CL, Fountain FF, Schaffner W. Increased incidence of the outbreak
486		strain of Mycobacterium tuberculosis in the surrounding community after an outbreak in a
487		jail. (Original Article). South Med J. 2003;96: 155–158. Available:
488		https://go.gale.com/ps/i.do?p=AONE&sw=w&issn=00384348&v=2.1⁢=r&id=GALE%
489		7CA98828112&sid=googleScholar&linkaccess=fulltext
490	21.	Didelot X, Fraser C, Gardy J, Colijn C, Malik H. Genomic infectious disease
491		epidemiology in partially sampled and ongoing outbreaks. Mol Biol Evol. 2017;34: 997–
492		1007. doi:10.1093/molbev/msw275
493	22.	Gygli SM, Loiseau C, Jugheli L, Adamia N, Trauner A, Reinhard M, et al. Prisons as
494		ecological drivers of fitness-compensated multidrug-resistant Mycobacterium
495		tuberculosis. Nat Med 2021 277. 2021:27: 1171–1177. doi:10.1038/s41591-021-01358-x
496	23.	Menardo F, Duchêne S, Brites D, Gagneux S. The molecular clock of mycobacterium
497		tuberculosis. PLoS Pathog. 2019;15. doi:10.1371/journal.ppat.1008067
498	24.	Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A, RAxML-NG: A fast, scalable.
499		and user-friendly tool for maximum likelihood phylogenetic inference. Wren J. editor.
500		Bioinformatics, 2019 [cited 13 May 2019], doi:10.1093/bioinformatics/btz305
501	25.	Bouckaert R. Vaughan TG. Barido-Sottani J. Duchêne S. Fourment M. Gavryushkina A.
502		et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis.
503		PLoS Comput Biol. 2019:15: e1006650. doi:10.1371/journal.pcbi.1006650
504	26.	World Health Organization. Global tuberculosis report 2020. Geneva: 2020.
505	2.7.	Holt KE, Mcadam P, Vuong P, Thai K, Thuy N, Thuong T, et al. Frequent transmission of
506	27.	the Mycobacterium tuberculosis Beijing lineage and positive selection for the EsxW
507		Beijing variant in Vietnam Nat Genet 2018 doi:10.1038/s41588-018-0117-9
508	28	Fritz SA Purvis A Selectivity in mammalian extinction risk and threat types: A new
509	20.	measure of phylogenetic signal strength in binary traits. Conserv Riol 2010:24: 1042–
510		$1051 \text{ doi:}10.1111/i.1523_1739.2010.01455 \text{ x}$
511	29	Walker TM Kohl TA Omar S.V. Hedge I. Del Oio Elias C. Bradley P. et al. Whole-
512	2).	genome sequencing for prediction of Mycobacterium tuberculosis drug suscentibility and
512		resistance: A retrospective cohort study I ancet Infect Dis 2015.15. 1103 1202
51/		doi:10.1016/\$1473_3099(15)00062_6
515	30	Walker TM In CLC Harrell RH Evans IT Kanatai G Dedicaat ML et al Whale
516	50.	genome sequencing to delineate Mycobacterium tuberculosis outbroaks: A retrospective
510		genome sequencing to demicate mycobacterium tuberculosis outoreaks. A retrospective

517		observational study. Lancet Infect Dis. 2013;13: 137-146. doi:10.1016/S1473-
518		3099(12)70277-3
519	31.	Walmsley R. World Prison Population List (twelfth edition). Inst Crim Policy Res. 2018.
520		Available: www.prisonstudies.org
521	32.	Walter K, Martinez L, Arakaki-Sanchez D, Sequera V, G ES, T C, et al. The escalating
522		tuberculosis crisis in Central and South American prisons. Lancet.
523	33.	Tappero JW, Reporter R, Wenger JD, Ward BA, Reeves MW, Missbach TS, et al.
524		Meningococcal Disease in Los Angeles County, California, and among Men in the County
525		Jails. N Engl J Med. 1996;335: 833–841. doi:10.1056/nejm199609193351201
526	34.	Reinhart E, Chen DL. Incarceration and its disseminations: COVID-19 pandemic lessons
527		from Chicago's cook county jail. Health Aff. 2020;39: 1412–1418.
528		doi:10.1377/hlthaff.2020.00652
529	35.	Sandra McCoy, Stefano M. Bertozze, David Sears, Catherine Duarte, Drew Cameron,
530		Brie Williams. Urgent Memo: COVID-19 Outbreak: San Quentin Prison. San Francisco,
531		CA; 2020. Available: https://www.documentcloud.org/documents/6956448-06-15-20-San-
532		Quentin-Urgent-Memo.html
533	36.	Santos A da S, Oliveira RD de, Lemos EF, Lima F, Cohen T, Cords O, et al. Yield,
534		Efficiency and Costs of Mass Screening Algorithms for Tuberculosis in Brazilian Prisons.
535		Clin Infect Dis. 2020 [cited 3 Apr 2020]. doi:10.1093/cid/ciaa135
536	37.	American Public Health Association. Advancing Public Health Interventions to Address
537		the Harms of the Carceral System. 2021. Available: https://www.apha.org/policies-and-
538		advocacy/public-health-policy-statements/policy-database/2021/01/14/advancing-public-
539		health-interventions-to-address-the-harms-of-the-carceral-system
540	38.	Wang EA, Western B, Berwick DM. COVID-19, Decarceration, and the Role of
541		Clinicians, Health Systems, and Payers: A Report From the National Academy of
542		Sciences, Engineering, and Medicine. Jama. 2020;06511. doi:10.1001/jama.2020.22109
543	39.	De Oliveira Andrade R. Covid-19: Prisons exposed in Brazil's crisis. The BMJ. 2020.
544		doi:10.1136/bmj.m2884
545	40.	Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, et al. Whole
546		genome sequencing of Mycobacterium tuberculosis: current standards and open issues.
547		Nat Rev Microbiol. 2019;17: 533–545. doi:10.1038/s41579-019-0214-5
548	41.	Walter KS, Colijn C, Cohen T, Mathema B, Liu Q, Bowers J, et al. Genomic variant-
549		identification methods may alter mycobacterium tuberculosis transmission inferences.
550		Microb Genomics. 2020;6: 1–16. doi:10.1099/mgen.0.000418
551		



Year

Year

Year



Lineage - 4.1 - 4.3 - 4.7 - 4.9 - 4.2 - 4.4 - 4.8



а





Incarceration status

- Formerly incarcerated
- Incarcerated
- No incarceration history

#### City

Campo Grande Corumba/ Ladario Dourados Ponta Pora





#### Location at TB notification

- --- Community
- 🗕 Jail
- Prison: Campo Grande
- Prison: Dourados
- Prison: Other