Genomic Surveillance of Yellow Fever Virus Epidemic Waves in São Paulo, Brazil, 2017 – 2018

de Souza, R. P.*[1], Hill, S. C.*[2], Thézé, J.[2], Claro, I.[3], Aguiar, R. S. [4], Dellicour, S.[5], Abade, L.[6], Santos, F. C. P.[1], Cunha, M. S.[1], Nogueira, J. S.[1], Salles, F. C. S.[3], Rocco, I. M.[1], Maeda, A. Y.[1], Vasami, F. G. S.[1], du Plessis, L.[2], Silveira, P. P. [4], Giovanetti, M.[7], de Goes, J.[7], Quick, J.[8], Fernandes, N. C. C. A.[1], Guerra, J. M. [1], Réssio, R. A.[1], Cirqueira, C. S.[1], Iglezias, S. D.[1], Delgado, J.D. [1], Macedo, F. L. L.[1], Timenetsky, M. C. S. T.[1], de Paula, R. [9], Spinola, R. [9], Deus, J.T. [10], Mucci, L.F. [10], Tubaki, R.M. [10], Menezes, R.M.T.[10], Ramos, P.L.[11], Abreu A. L.[12], Cruz, L. N.[12], Loman, N.[8], Bispo, A.[7], Pybus, O. G.[2], Alcantara, L. C. J. [7], Sabino, E. C.[3], Faria, N. R.[2]

1. Instituto Adolfo Lutz. Av. Dr. Arnaldo, 355, São Paulo, Brazil
2. Department of Zoology, University of Oxford, South Parks Road, Oxford, UK
3. Departamento de Moléstias Infecciosas e Parasitarias da Faculdade de Medicina e Instituto de Medicina Tropical, Universidade de São Paulo, São Paulo, Brazil
4. Laboratório de Virologia Molecular, Departamento de Genética, Instituto de Biologia, Rio de Janeiro, Brazil
5. Department of Microbiology and Immunology, Rega Institute, KU Leuven – University of Leuven, Herestraat 49, 3000 Leuven, Belgium
6. The Global Health Network, Nuffield Department of Medicine, University of Oxford, Oxford, UK
7. Laboratório de Flavivírus, Instituto Oswaldo Cruz, FIOCRUZ, Rio de Janeiro, Brazil
8. Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK
9. Centro de Vigilância Epidemiológica "Prof. Alexandre Vranjac", São Paulo, Brazil
10. Superintedência do Controle de Endemias, São Paulo, Brazil
11. Fundação Parque Zoológico de São Paulo, São Paulo, Brazil

12. Secretaria de Vigilância em Saúde, Ministério da Saúde (SVS/MS), Brasília-DF, Brazil

*Denotes equal contribution.

Corresponding author mailing address: doencasporvetor@ial.sp.gov.br and nuno.faria@zoo.ox.ac.uk

Article Summary Line: Genomic surveillance of yellow fever in São Paulo during the yellow fever 2017-2018 epidemic reveals movement towards Atlantic coast.

Running title: Phylogeography of yellow fever virus in São Paulo

Keywords: Yellow fever, outbreak, Brazil

Abstract

São Paulo (SP), a densely populated state in southeast Brazil that contains one of the world's largest urban regions, has experienced its largest yellow fever virus (YFV) outbreak in decades. Surveillance in non-human primates (NHP) is important in order to detect YFV early during an epidemic or epizootic, to quantify the magnitude of the outbreak in NHP, and to evaluate the risk of YFV spillover infection in human populations. To better understand the genetic diversity and spatial distribution of YFV during the current outbreak in southeast Brazil, we generated 46 new virus genomes from YFV positive cases identified in 18 different municipalities in SP, mostly sampled from non-human primates between April 2017 and February 2018. Our data show that most NHP cases in São Paulo state were likely caused by the introduction of a single YFV lineage from Minas Gerais to São Paulo. Phylogenetic and phylogeographic analyses of these data indicate that YFV spread southwards from Minas Gerais into São Paulo state at a typical rate of <1km per day. These results shed light in the sylvatic transmission of yellow fever in highly fragmented forested regions in São Paulo state and highlight the importance of continued operational research and surveillance of zoonotic pathogens in sentinel populations.

Author´s Summary

Since July 2016, the southeast region of Brazil has experienced the largest yellow fever virus (YFV) outbreak in decades. São Paulo is the most densely populated state in southeast Brazil. YFV is not normally present in São Paulo state and therefore a large proportion of the 18 million inhabitants of the state have not been vaccinated against YFV. The presence of YFV in São Paulo state therefore represents a serious threat to public health. In Brazil, YFV typically circulates among non-human primates, with cases in humans representing isolated 'spillover' events from this predominantly sylvatic cycle. Understanding the epidemiological dynamics and spread of YFV in primates is therefore critical for contextualising human cases, and guiding vaccination strategies that can better protect local human populations. Here, we analyse the geographic and temporal distribution of observed cases of YFV in non-human primates in São Paulo state. We generate sequence data from 46 YFV positive cases, and perform phylogenetic and phylogeographic analyses aimed at understanding the spatial spread of YFV in São Paulo state. We show that most cases in non-human primates in the São Paulo state were likely caused by a single introduction of YFV from Minas Gerais to São Paulo. Analyses of these data indicate that YFV spread southwards from Minas Gerais into São Paulo state at a typical rate of <1 km per day, consistent with a scenario of continued spread in non-human primates and sylvatic vectors across forested patches, with occasional spillover to unvaccinated human populations.

74 ## Introduction

75 Yellow fever (YF) is an acute hemorrhagic disease caused by the yellow fever virus (YFV),
76 an single-stranded positive-sense RNA virus from the Flavivirus genus. Clinical manifestations of
77 YF in humans range from inapparent or mild disease in up to 80% of infected cases, to severe
78 hepatitis and hemorrhagic disease. Among patients who develop visceral disease case fatality rate
79 can range from 20% to 60% [1]. YFV is commonly classified into four genotypes, denoted West
80 African, East African, South American genotype I (SA1) and South American genotype II (SA2) [2,
81 3]. To date, only the SA1 and the SA2 genotypes have been detected in the Americas [2, 4].YF is
82 preventable in humans by administration of a single dose of an extremely effective vaccine (17D)
83 that provides life-long protection against the disease.

84 Yellow fever is endemic to the American and African tropics, where nearly 400 million
85 people are estimated to be at risk of infection [5]. In the Americas, YFV transmission is thought to
86 occur through two main transmission cycles: (i) a sylvatic cycle that involves transmission between
87 tree-dwelling mosquitoes (e.g. *Haemagogus janthinomys*, *Haemagogus leucocelaenus* and in some
88 local areas, mosquitoes of the Sabethes genus) and non-human primates in forested areas ([6-8]);
89 and (ii) an urban cycle, undetected since the 1940s with transmission between humans and *Aedes*
90 *aegypti* mosquitoes in heavily populated areas [1, 9]. Alongside the urban and sylvatic cycles, in
91 Africa, YFV transmission also occurs through (iii) an intermediary cycle, in which semi-domestic
92 mosquitos that inhabit both urban and forested environment scan transmit YFV to both humans and
93 non-human primates [1]. In Brazil, epizootics of YF in nonhuman primates (NHP) typically ignite
94 in neotropical forests and subsequently cause human epidemics that mostly affect unvaccinated
95 male adults living in rural areas. These events occur approximately every 5-14 years [10].
96 Importantly, surveillance of primate epizootics can be used as an indicator of epidemic risk (i.e. as a
97 sentinel event), alerting us to the likelihood of future human cases [11]. Since 1999, most YFV
98 cases were registered in Southeast region, outside the regions in North and Center-West where
99 transmission of YFV predominated in the past, and indicating a recent and unprecedent expansion

100 of the virus towards highly populated urban centres in Southeast Brazil (e.g. São Paulo, Rio de
101 Janeiro and Belo Horizonte,) where vaccination had not been recommended before.

102 Since July 2016, the southeast region of Brazil has experienced the largest YFV outbreak in
103 decades. The epizootic/epidemic is caused by a rapidly-spreading lineage of the SA1 genotype that
104 is thought to have originated from the Amazon region [12]. By March 2019 there were 2,204
105 confirmed cases of YF and 757 fatalities, according to the Brazilian Ministry of Health [13]. São
106 Paulo state is a highly densely populated state of southeast Brazil that contains one of the world's
107 largest urban agglomerations. From January 2017 to 12 April 2019, São Paulo state reported 656
108 human cases and 827 NHP cases of YF [14]. Despite the magnitude and expansion of the outbreak
109 in São Paulo state, little is known about how fast the virus is spreading through NHP populations in
110 the region. Concerningly, as of late September 2018, 46% of people in São Paulo state (12 million
111 people) were still not vaccinated for YFV [13]. The risk of reestablishment of urban transmission of
112 YFV remains uncertain, but if established, would have very serious consequences for public health.

113 Analytical insights into the spread of YFV among NHP populations in São Paulo state (SP)
114 can help to guide public health efforts and targeted vaccination campaigns. Here, we report the
115 screening of dead NHP in São Paulo state for YFV using a previously-developed portable
116 sequencing protocol, resulting in the generation of 46 YFV genomes from positive samples. We
117 analyse epidemiological case data and virus phylogeography to characterize the spread of YFV in
118 São Paulo state during the first two waves of the current outbreak.

119

Methods

Sample collection and mapping of non-human primate cases

Notifications of dead NHP are made to the São Paulo's State Epidemiological Health Department by telephone, by fax, or electronically. Where multiple NHP were found dead at the same mortality event, the number of primates at each location was recorded but in most instances only one or a few animals from each event were tested. Despite the possibility of epidemiological association with positive, dead animals, these untested animals are not considered positive by current surveillance systems. Each YFV-confirmed NHP case was georeferenced according to the original case location using a handheld Garmin Etrex® GPS. Quantum GIS [15] was used to create a choropleth depicting the geographic distribution and number of confirmed cases across different municipalities in São Paulo state.

Sample collection of human cases

We tested 6 randomly selected samples from human cases collected in São Paulo state from four municipalities (**Table S1**). Samples were collected for molecular diagnostics in several municipalities of Sao Paulo state, and sent for testing at Instituto Adolfo Lutz.

Ethics statement

Research on human cases was supported by the Brazilian Ministry of Health (MoH) as part of the arboviral genomic surveillance efforts within the terms of Resolution 510/2016 of CONEP (Comissão Nacional de Ética em Pesquisa, Ministério da Saúde; National Ethical Committee for Research, Ministry of Health). Only already dead non-human primates were sampled for YFV surveillance. The surveillance protocol for dead non-human primates was approved by the Ethics Committee for the use of Animals In Research, Instituto Adolfo Lutz, under the numbers 0135D/2012 and 020G/2014.

146

147 ***Identification of YFV positive cases***

148 YFV positive cases were detected using at least one of the three methods described below;
149 immunofluorescence, immunohistochemistry, or reverse transcription quantitative PCR (RT-
150 qPCR). In rare circumstances (n=12 cases), a case was also considered positive based on its close
151 epidemiological association with tested, RT-qPCR positive cases. Here, a confirmed case is
152 considered to be one that is positive by at least one method.

153

154 ***Indirect immunoflorescence***

155 Samples of NHP blood or serum, and tissue material suspensions obtained from autopsies, were
156 identified using a standardized indirect immunofluorescence technique [16]. An in-house polyclonal
157 anti-flavivirus antibody (anti-DENV 1-4) and an anti-mouse IgG-FITC antibody (Sigma) were used.
158 Positive samples were typed by indirect immunofluorescence with monoclonal antibodies for YFV
159 (Biomanguinhos, Rio de Janeiro).

160

161 ***Histopathology and immunohistochemistry for NHP cases***

162 Samples of brain, heart, lung, liver, spleen and kidney from NHP were fixed in formaldehyde and
163 embedded in paraffin. Histological sections of these tissues were stained with hematoxylin and
164 eosin and examined on a microscope. Where damage consistent with YFV infection was observed,
165 indirect immune histochemistry was used to detect the presence of the yellow fever virus antigen.
166 Sections of liver (0.3uM) were placed on slides coated with silane and treated with an in-house
167 polyclonal anti-YFV antibody (produced in mice, and used at 1/30,000 dilution). The slides were
168 treated with anti-mouse secondary antibodies, linked to either horseradish peroxidase (Reveal HRP
169 Spring polymer System, Spring) or to alkaline phosphate (Link MACH4 universal AP Polymer
170 and Polymer MACH4 universal AP, Biocare). Chromogenic detection of the presence of YFV was

171 subsequently conducted using the substrates 3,3'-diaminobenzidine or Fast Red (Warp Red,
172 Biocare), respectively.
173
174 ***RT-qPCR***
175 Total RNA was extracted from tissue and serum samples using two commercial kits:
176 QIAamp® RNA Blood for tissues and QIAamp® Viral RNA Kit for serum (Qiagen Inc., Germany)
177 according to the manufacturer's instructions. Viral RNA was detected using two previously
178 published RT-qPCR techniques [17, 18].
179
180 ***MinION genome sequencing***
181 A selection of positive samples was sequenced using a rapid whole-genome sequencing protocol
182 that has been previously validated and successfully applied in Brazil [19, 20]. In brief, cDNA was
183 produced from viral RNA using random hexamers and the Protoscript II First Strand cDNA
184 synthesis kit (NEB). The genome was amplified using a multiplex PCR scheme designed to produce
185 overlapping 500bp amplicons across the whole coding region of the recent South American
186 genotype I outbreak clade. PCR products were quantified, barcoded using the Oxford Nanopore
187 Technologies (ONT) Native Barcoding Kit (NBD103), and pooled in an equimolar fashion.
188 Sequencing libraries consisting of 10-12 samples per library were constructed using the Ligation
189 Sequencing kit (ONT, SQK-LSK108). Sequencing was performed on a ONT flow cell for up to 48h
190 as described previously [12].
191
192 ***Generation of consensus sequences***
193 Consensus sequences for each barcoded sample were generated following a previously published
194 approach [21]. Briefly, raw files were basecalled using Albacore, demultiplexed and trimmed using
195 Porechop, and then mapped with *bwa* to a reference genome (GenBank accession number
196 JF912190). Nanopolish variant calling was applied to the assembly to detect single nucleotide

197 variants to the reference genome. Consensus sequences were generated; non-overlapped primer
198 binding sites, and sites for which coverage was <20X were replaced with ambiguity code N.
199 Sequencing statistics can be found in **Table S2**. Accession numbers of newly generated sequences
200 can be found in **Table S1**.
201
202
203 ***Phylogenetic analyses***
204 Consensus sequences were aligned using MAFFT v.7 [22]. Maximum likelihood (ML)
205 phylogenetic trees were estimated using RAxML v.8 [23] under a GTR + Γ_4 nucleotide substitution
206 model. Statistical support for phylogenetic nodes was estimated using a ML bootstrap approach
207 with 100 replicates. The genome sequences generated here were combined with a previously
208 released alignment of genomes from the 2016-2018 YFV epidemic sampled elsewhere in Brazil
209 [12]. The full dataset analysed here contains 98 YFV sequences, 46 of which were isolated from SP
210 and generated in this study (alignments and ML phylogenetic tree of full dataset can be found in
211 GitHub repository (to be added).
212
213 ***Phylogeographic analyses***
214 To investigate the spread of YFV in São Paulo, we analysed in more detail Lineage 2, a
215 phylogenetic lineage that includes the 46 new genomes from SP and two basal sequences from
216 Minas Gerais. These 48 sequences are shown in a ML phylogeny in **Fig. 2A**. Georeferenced and
217 time-stamped sequences were analysed in BEAST v.1.8.4 [24] using the BEAGLE library to
218 enhance computational speed [25]. We used a skygrid coalescent tree prior [26] and a continuous
219 phylogeographic model that uses a relaxed random walk to model the spatial diffusion of lineages.
220 Dispersal velocity variation among lineages was represented using a Cauchy distribution [27].
221 Virus diffusion through time and space was summarised using 1000 phylogenies sampled at
222 regular intervals from the posterior distribution (after exclusion of burn-in). The R package

223 “seraphim” [28] was used to extract the spatio-temporal information contained in these phylogenies,
224 whose branches can be considered as movement vectors (each having start and end spatial
225 coordinates, and start and end dates, in decimal units). The package “seraphim” was also used to
226 estimate statistics of spatial dissemination, such as lineage dispersal velocity, and change inthe
227 maximal wavefront distance from epizootic origin (**FigureS2**) [28]. The results shown in Figure 2
228 were generated using the “spreadGraphic” R function [28].

229 ***Data availability***

230 Epidemiological data, genomic data and XML files analysed in this study are available in the
231 GitHub repository (XXX). New sequences have been deposited in GenBank under accession
232 numbers MH018064, MH018067, MH030049- MH030091 and MH193173- MH193175.

233

234

235

236 **Results**

237 Liver, brain or blood samples from 591 NHP tested positive for YFV during the period from
238 epidemiological week 29 of 2016 to week 4 of 2018. An additional 138 NHP were observed dead at
239 the same time and location as identified positive cases, but were not tested for YFV. Samples were
240 confirmed positive using immunohistochemistry (*n*=466) and/or RT-qPCR (*n*=311). Most
241 confirmed cases (at least one positive diagnostic result) in NHP were from animals of the *Alouatta*
242 genus (88%; 403 of 459 cases for which genus information was available), followed by *Callithrix*
243 (8%; 35/459), *Callicebus* (2%; 9/459), *Cebus* (2%; 9/459) and *Sapajus* (0.7%; 3/459).

244 Our results reveal two main epizootic waves, in an area non-vaccinated against YFV, with
245 no notified cases of YFV since the 1940s, preceded by a 'pre-epizootic' phase of low case counts,
246 occurring in an area that have been previously vaccinated. The pre-epizootic phase runs from July
247 2016 to Jan 2017 (6%, 33/591 cases), phase 1 runs from Feb to Jul 2017 (20%, 119/591 cases) and
248 phase 2 from Jul 2017 to Feb 2018 (74%, 439/591 cases) (**Fig. 1A**).The peaks of epizootic phases 1
249 and 2 were mid Apr 2017 and late Nov 2017, respectively.

250 YF cases until November 2018 were geographically widespread across 57 municipalities of
251 SP, with most cases concentrated around the southeast region of the state (**Fig. 2**). The greatest
252 number of confirmed cases were observed in the municipalities of Mairiporã (*n*=88), Jundiaí
253 (*n*=68), São Paulo (*n*=62), Bragança Paulista (*n*=48), and Atibaia (*n*=36). There is a clear distinction
254 between the geographic location of cases in the pre-epizootic and in the epizootic phases 1 and 2
255 (**Fig. 1B**). Specifically, almost all earlier cases that form part of the pre-epizootic wave occur in a
256 geographically distinct cluster of municipalities in the north of São Paulo state (**Fig. 1B**; **Fig. 2**),
257 where YFV has also been detected in 2000 and 2008 [29]. The two waves of epizootic transmission
258 begin at the time that the virus began to be detected in municipalities in the south of São Paulo state
259 (within 200km of São Paulo municipality) (**Figs. 1B and 2**).

260 To investigate the source and transmission of YFV during the two epizootic waves, and the
261 genetic diversity of the virus circulating in NHP populations across SP state, we generated whole

262 genome for forty-one RT-qPCR positive NHP samples (median Ct-values of 13, range 9-25) from
263 14 municipalities using a previously described MinION sequencing protocol [12, 19]. We also
264 sequenced 6 whole genomes from human samples collected in São Paulo state (median Ct-values of
265 35, range 32-37), representing an additional four municipalities. The most recent samples chosen
266 for genome sequencing were from Jan-Feb 2018: isolate SA130 is from a human case in Mairiporã
267 municipality, and isolates SA129 and SA131 are from *Alouatta* monkeys, from Guarulhos
268 municipality and from the south of São Paulo city, respectively (**Fig. S1)**.

269 Newly generated genome sequences (*n*=46) were aligned to available outbreak genome data
270 (*n*=52) and maximum likelihood methods were used to estimate phylogenetic relationships. Our
271 genetic analyses indicate a strong phylogenetic spatial structure of the ongoing epizootic in SP state,
272 with almost all (42/46) sequences from SP clustering in a single well-supported monophyletic clade
273 (bootstrap score=91%). The most closely related sequences to the main SP outbreak clade are from
274 Santa Rita de Caldas (M11), and Caldas (M5) (**Fig. 3**, see **Fig. 2** for locations). These data indicate
275 that the majority of the enzootic infections of YFV in SP may have arisen from a single
276 introduction, possibly from south of Minas Gerais. This single introduction was subsequently
277 followed by sustained transmission within SP state giving rise to epizootic waves 1 and 2 (**Fig. 1**).

278 Finally we used a continuous diffusion model to investigate how rapidly the virus has been
279 spreading over space and time. We find that that YFV disseminated rapidly from the south of Minas
280 Gerais state through São Paulo municipality in the direction to the Atlantic Forest of São Paulo
281 (**Fig. 4**), with the epidemic wavefront moving on average 0.84 (0.50-2.19) km/day (**Fig. S2**). To
282 investigate whether virus lineages moved faster during the first epizootic season, we estimate mean
283 phylogeny branch velocity for phase 1 of the epizootic to that of phase 2. Mean branch velocity for
284 phase 1 was estimated to be around 0.97 (BCI: 0.5-3.0) km/day. For the second epizootic phase we
285 find that the mean branch velocity was slightly slower at 0.78 (BCI: 0.46-2.22) km/day, although
286 the difference is not statistically significant.

287

288 ## Discussion

289 In this study we present data on the temporal and geographic distribution of detected YFV
290 cases in NHPs in São Paulo state, Brazil, during 2016-2018. Our results demonstrate the existence
291 of several distinct phases of the outbreak. Specifically, we observe a 'pre-epizootic' period in which
292 a small number of cases were identified in the northern region of SP state during late 2016. This
293 pre-epizootic period was followed by two larger epizootic phases, the first from July 2016 – July
294 2017, and the second from August 2017 to at least Feb 2018 (the data of the most recent data
295 available for this study). During these two phases, a large number of cases were observed in the
296 south of SP state. We generated 46 novel genomes of YFV from samples collected in 18 different
297 municipalities in São Paulo state. These majority of sequenced viruses were sampled from *Alouatta*
298 monkeys during the late 2017 and early 2018, and are therefore largely representative of the second
299 epizootic phase (**Figure S1, Fig. 1**). Phylogenetic and phylogeographic analyses of these data
300 indicate that virus lineages move diffused southwards from Minas Gerais into São Paulo state with
301 an invasion speed of<1km per day (**Fig. 4**).

302 Understanding the processes that led to the initiation of the two main epizootic waves in SP
303 is important for preventing future YFV outbreaks. In this study were not able to sequence samples
304 from cases identified in northern SP during the pre-epizootic phase. We are therefore unable to use
305 genetic and evolutionary analysis to evaluate the spatial and epidemiological relationship between
306 cases observed during this pre-epizootic phase in the north of SP state, and cases observed during
307 the subsequent, larger outbreaks in the south of SP state. Sequencing of samples from the pre-
308 epizootic wave would be important to determine whether the northern, pre-epidemic cases directly
309 initiated the epizootic observed in Minas Gerais and southern São Paulo around February 2017, or
310 whether the northern YFV lineage was successfully extinguished before onwards spread. We
311 contend that longitudinal entomological data across different climatic regions in Sao Paulo state
312 will help to disentangle seasonal and spatial differences in the availability of mosquito vectors and
313 to better anticipate future YFV transmission.

314 As of January 2019, human and NHP cases of YFV have been reported in the south western
315 regions of São Paulo state, and in neighbouring Paraná state(PAHO 2019). These cases may
316 represent the beginning of a third wave of YFV in the region, and the continued expansion of YFV
317 down the coastal region of southern Brazil. Interestingly, historical data from sylvatic YFV
318 outbreaks in Brazil suggest a similar geographic pattern of viral spread during sylvatic outbreaks
319 that occurred during the 1930s and 1950s [30, 31]. Detailed comparison of data from the current
320 outbreak to historical outbreaks, or to YFV genome sequences recovered from archival samples,
321 may help to identify key ecological corridors of YFV transmission in Brazil that have remained
322 constant over time.

323 Using YFV genetic data, we estimate that the virus lineages moved at a mean rate of <1 km
324 per day during the first and second epizootic phases in SP state. We observe no difference between
325 the rate of viral lineage dissemination during the first (Feb 2017 – July 2017) and second (August
326 2017 - Feb 2018) phases (**Fig. S2**). Our estimate of the rate of YFV lineage movement in SP state is
327 slightly lower than that previously estimated by ourselves and others for the states of Minas Gerais,
328 Espírito Santo and Rio de Janeiro (~4km/day) [12]. Additional analyses of larger datasets than span
329 state borders are necessary to test whether ecological drivers, such as NHP density and the degree
330 of habitat fragmentation, or whether differences in surveillance efforts and sampling or differences
331 in viral load found in NHP species might explain differences in the rate of YFV spread among
332 different areas.

333 The vaccination strategies employed to combat the YFV in SP state during the ongoing
334 outbreak have been designed on the basis of the georeferenced positive NHP occurrences (Adriano
335 Pinter, personal communication). Human activities may have the potential to accelerate the
336 dispersal of YFV virus lineages via the travel of asymptomatic patients, illegal trade in wildlife, or
337 by encouraging the 'hitchhiking' of infected mosquitos in cargo or vehicles during agricultural or
338 forestry work. A better understanding of the contribution of these factors is needed to achieve
339 WHO's goals to eliminate YFV epidemics before 2026.

340 Understanding how YFV is introduced into highly urbanised areas, such as São Paulo city,
341 is important for designing strategies that can effectively interrupt these introductions. Isolate SA131
342 was collected from an *Alouatta* sp. individual at the Parque Estadual das Fontes do Ipiranga (PEFI),
343 a public park that forms part of the São Paulo city zoo and which comprises a 53 km^2 fragmented
344 piece of Atlantic forest. The zoo, the largest in Brazil, attracts tens of thousands of visitors per week
345 and was closed to visitors on the 22 Jan 2018 after YFV confirmation, and reopened on the 15 Mar
346 2018.

347 The area has ecological conditions for the maintenance of mosquitoes of the genus *Sabethes*
348 and *Haemagogus*, vectors of forest mosquitoes commonly involved in the wild transmission cycle
349 of the YFV, however entomological collections performed in the park did not show the presence of
350 *Haemagogus* mosquitoes, two specimens of the genus *Sabethes* were collected, *Sa. albiprivus* and
351 *Sa. chloropterus*. Among the most frequent mosquitoes were *Aedes scapularis* and *Aedes*
352 *albopictus*. Phylogenetic analyses show that the YFV genome recovered from sample SA131
353 clusters together (bootstrap score=76%) with isolate Y37-244 collected in Piracaia, 120 km from
354 the park, and not with more local isolates that were detected in the outskirts of SP or in surrounding
355 municipalities. This result could be explained by: (i) incomplete sampling of a wave of continuous
356 transmission among NHPs that live between the two locations; (ii) human-mediated transport of
357 YFV infected NHP, (iii) human-mediated transport of mosquitos [32] or (iv) introduction of the
358 virus by an asymptomatic human visitor, who carried the virus from Piracaia to PEFI. Scenario (i)
359 seems unlikely since PEFI is an isolated forested fragment to the south of SP city with limited
360 connectivity to other forested fragments; at this stage, scenarios (ii), (iii) and (iv) remain possible.

361 Our study sheds light on the spatial and temporal dynamics of yellow fever in different
362 animal hosts across Sao Paulo state. A better understanding of the vectors and host species involved
363 in the persistence of the virus both in epidemic seasons and in non-epidemic periods is critical to
364 understand the drivers of yellow fever transmission and anticipate future outbreaks.

365

366 **Biographical Sketch**

367 Renato Pereira de Souza has a PhD in Epidemiology from the University of São Paulo, Brazil. He is
368 currently a Scientific Researcher at the Instituto Adolfo Lutz, reference laboratory for arboviruses in
369 São Paulo, where he combines epidemiological surveillance with studies on the molecular evolution
370 of arboviruses, hantavirus and arenavirus. Sarah Hill has a PhD in viral genomic epidemiology from
371 the University of Oxford. She is currently a postdoctoral researcher at the University, working on
372 understanding the dynamics of mosquito-borne arboviruses.

373

374 **Acknowledgments**

375 We thank the IAL and the USPTM staff, in particular the crew from Núcleo de Doenças de
376 Transmissão Vetorial and from the Centro de Patologia. This study was developed as a
377 collaborative effort between the IAL, FIOCRUZ, and the University of Oxford, UK. The research
378 was supported by a Wellcome Trust and Royal Society Sir Henry Dale Fellowship (grant
379 204311/Z/16/Z), internal HEFCE GCRF grant 005073, John Fell Research Fund Grant 005166,
380 Medical Research Council and FAPESP CADDE partnership award (MR/S0195/1), CNPq
381 #400354/2016-0 and FAPESP# 2016/01735-2, and by the Oxford Martin School. SD is supported
382 by the Fund for Scientific Research (FWO) Flanders (Fonds voor Wetenschappelijk Onderzoek,
383 Flanders, Belgium). This work was also supported by CNPq/MCTI
384 Decit/SCTIE/MoH (440685/2016-8) and CAPES (88887.130716/2016-00).

385

387 **References**

388 1. Monath TP, Vasconcelos PF. Yellow fever. Journal of clinical virology : the official
389 publication of the Pan American Society for Clinical Virology. 2015;64:160-73. Epub 2014/12/03.
390 doi: 10.1016/j.jcv.2014.08.030. PubMed PMID: 25453327.

391 2. Wang E, Weaver SC, Shope RE, Tesh RB, Watts DM, Barrett AD. Genetic variation in
392 yellow fever virus: duplication in the 3' noncoding region of strains from Africa. Virology.
393 1996;225(2):274-81. Epub 1996/11/15. doi: 10.1006/viro.1996.0601. PubMed PMID: 8918913.

394 3. Mutebi JP, Wang H, Li L, Bryant JE, Barrett AD. Phylogenetic and evolutionary
395 relationships among yellow fever virus isolates in Africa. Journal of virology. 2001;75(15):6999-
396 7008. Epub 2001/07/04. doi: 10.1128/JVI.75.15.6999-7008.2001. PubMed PMID: 11435580;
397 PubMed Central PMCID: PMCPMC114428.

398 4. Bryant JE, Barrett AD. Comparative phylogenies of yellow fever isolates from Peru and
399 Brazil. FEMS Immunol Med Microbiol. 2003;39(2):103-18. Epub 2003/11/20. doi: 10.1016/S0928-
400 8244(03)00238-4. PubMed PMID: 14625093.

401 5. Shearer FM, Moyes CL, Pigott DM, Brady OJ, Marinho F, Deshpande A, et al. Global
402 yellow fever vaccination coverage from 1970 to 2016: an adjusted retrospective analysis. The
403 Lancet infectious diseases. 2017;17(11):1209-17. Epub 2017/08/22. doi: 10.1016/S1473-
404 3099(17)30419-X. PubMed PMID: 28822780; PubMed Central PMCID: PMCPMC5666204.

405 6. Abreu FVS, Ribeiro IP, Ferreira-de-Brito A, Santos A, Miranda RM, Bonelly IS, et al.
406 Haemagogus leucocelaenus and Haemagogus janthinomys are the primary vectors in the major
407 yellow fever outbreak in Brazil, 2016-2018. Emerg Microbes Infect. 2019;8(1):218-31. Epub
408 2019/03/15. doi: 10.1080/22221751.2019.1568180. PubMed PMID: 30866775; PubMed Central
409 PMCID: PMCPMC6455131.

410 7. Vasconcelos PF. Yellow fever in Brazil: thoughts and hypotheses on the emergence in
411 previously free areas. Revista de saude publica. 2010;44(6):1144-9. Epub 2010/11/27. PubMed
412 PMID: 21109907.

413 8. Vasconcelos PF, Sperb AF, Monteiro HA, Torres MA, Sousa MR, Vasconcelos HB, et al.
414 Isolations of yellow fever virus from Haemagogus leucocelaenus in Rio Grande do Sul State,
415 Brazil. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2003;97(1):60-2.
416 Epub 2003/08/02. PubMed PMID: 12892055.
417 9. Souza RP, Petrella S, Coimbra TL, Maeda AY, Rocco IM, Bisordi I, et al. Isolation of
418 yellow fever virus (YFV) from naturally infected Haemagogus (Conopostegus) leucocelaenus
419 (diptera, cukicudae) in Sao Paulo State, Brazil, 2009. Revista do Instituto de Medicina Tropical de
420 Sao Paulo. 2011;53(3):133-9. Epub 2011/07/15. PubMed PMID: 21755235.
421 10. Camara FP, Gomes AL, Carvalho LM, Castello LG. Dynamic behavior of sylvatic yellow
422 fever in Brazil (1954-2008). Revista da Sociedade Brasileira de Medicina Tropical. 2011;44(3):297-
423 9. Epub 2011/05/04. PubMed PMID: 21537794.
424 11. Vasconcelos PF. [Yellow Fever]. Revista da Sociedade Brasileira de Medicina Tropical.
425 2003;36(2):275-93. Epub 2003/06/14. PubMed PMID: 12806465.
426 12. Faria NR, Kraemer MUG, Hill SC, Goes de Jesus J, Aguiar RS, Iani FCM, et al. Genomic
427 and epidemiological monitoring of yellow fever virus transmission potential. Science.
428 2018;361(6405):894-9. Epub 2018/08/25. doi: 10.1126/science.aat7115. PubMed PMID: 30139911.
429 13. PAHO/WHO. Epidemiological Update Yellow Fever
430 https://www.paho.org/hq/index.php?option=com_docman&view=download&category_slug=yellow
431 -fever-2194&alias=47954-6-march-2019-yellow-fever-epidemiological-
432 update&Itemid=270&lang=en 6 Mar 2019. 2019.
433 14. CVE. Boletim Epidemiologico da febre amarela (12/04/2019).
434 http://www.saude.sp.gov.br/resources/cve-centro-de-vigilancia-epidemiologica/areas-de-
435 vigilancia/doencas-de-transmissao-por-vetores-e-
436 zoonoses/doc/famarela/2019/fa19_boletim_epid_1204.pdf: Governo Estado de Sao Paulo, 2019.
437 15. Team QD. QGIS Geographic Information System. Open Source Geospatial Foundation
438 Project. http://qgis.osgeo.org. 2018.

439 16. Gubler DJ, Kuno G, Sather GE, Velez M, Oliver A. Mosquito cell cultures and specific
440 monoclonal antibodies in surveillance for dengue viruses. The American journal of tropical
441 medicine and hygiene. 1984;33(1):158-65. Epub 1984/01/01. PubMed PMID: 6364855.
442 17. Drosten C, Gottig S, Schilling S, Asper M, Panning M, Schmitz H, et al. Rapid detection
443 and quantification of RNA of Ebola and Marburg viruses, Lassa virus, Crimean-Congo hemorrhagic
444 fever virus, Rift Valley fever virus, dengue virus, and yellow fever virus by real-time reverse
445 transcription-PCR. Journal of clinical microbiology. 2002;40(7):2323-30. Epub 2002/06/29.
446 PubMed PMID: 12089242; PubMed Central PMCID: PMCPMC120575.
447 18. Domingo C, Patel P, Yillah J, Weidmann M, Mendez JA, Nakoune ER, et al. Advanced
448 yellow fever virus genome detection in point-of-care facilities and reference laboratories. Journal of
449 clinical microbiology. 2012;50(12):4054-60. Epub 2012/10/12. doi: 10.1128/JCM.01799-12.
450 PubMed PMID: 23052311; PubMed Central PMCID: PMCPMC3503008.
451 19. Quick J, Grubaugh ND, Pullan ST, Claro IM, Smith AD, Gangavarapu K, et al. Multiplex
452 PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from
453 clinical samples. Nat Protoc. 2017;12(6):1261-76. Epub 2017/05/26. doi: 10.1038/nprot.2017.066.
454 PubMed PMID: 28538739.
455 20. Faria NR, Quick J, Claro IM, Theze J, de Jesus JG, Giovanetti M, et al. Establishment and
456 cryptic transmission of Zika virus in Brazil and the Americas. Nature. 2017;546(7658):406-10.
457 Epub 2017/05/26. doi: 10.1038/nature22401. PubMed PMID: 28538727.
458 21. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, et al. Real-time, portable
459 genome sequencing for Ebola surveillance. Nature. 2016;530(7589):228-32. doi:
460 10.1038/nature16996. PubMed PMID: 26840485; PubMed Central PMCID: PMCPMC4817224.
461 22. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7:
462 improvements in performance and usability. Molecular biology and evolution. 2013;30(4):772-80.
463 Epub 2013/01/19. doi: 10.1093/molbev/mst010. PubMed PMID: 23329690; PubMed Central
464 PMCID: PMCPMC3603318.

465 23. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
466 phylogenies. Bioinformatics. 2014;30(9):1312-3. doi: 10.1093/bioinformatics/btu033. PubMed
467 PMID: 24451623; PubMed Central PMCID: PMCPMC3998144.
468 24. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and
469 the BEAST 1.7. Molecular biology and evolution. 2012;29(8):1969-73. doi:
470 10.1093/molbev/mss075. PubMed PMID: 22367748; PubMed Central PMCID: PMC3408070.
471 25. Ayres DL, Darling A, Zwickl DJ, Beerli P, Holder MT, Lewis PO, et al. BEAGLE: an
472 application programming interface and high-performance computing library for statistical
473 phylogenetics. Systematic biology. 2012;61(1):170-3. Epub 2011/10/04. doi:
474 10.1093/sysbio/syr100. PubMed PMID: 21963610; PubMed Central PMCID: PMCPMC3243739.
475 26. Gill MS, Lemey P, Faria NR, Rambaut A, Shapiro B, Suchard MA. Improving Bayesian
476 population dynamics inference: a coalescent-based model for multiple loci. Molecular biology and
477 evolution. 2013;30(3):713-24. doi: 10.1093/molbev/mss265. PubMed PMID: 23180580; PubMed
478 Central PMCID: PMC3563973.
479 27. Lemey P, Rambaut A, Welch JJ, Suchard MA. Phylogeography takes a relaxed random walk
480 in continuous space and time. Molecular biology and evolution. 2010;27(8):1877-85. doi:
481 10.1093/molbev/msq067. PubMed PMID: 20203288; PubMed Central PMCID: PMC2915639.
482 28. Dellicour S, Rose R, Faria NR, Lemey P, Pybus OG. SERAPHIM: studying environmental
483 rasters and phylogenetically-informed movements. Bioinformatics. 2016. doi:
484 10.1093/bioinformatics/btw384. PubMed PMID: 27334476.
485 29. Mucci LF, Junior RP, de Paula MB, Scandar SA, Pacchioni ML, Fernandes A, et al. Feeding
486 habits of mosquitoes (Diptera: Culicidae) in an area of sylvatic transmission of yellow fever in the
487 state of Sao Paulo, Brazil. J Venom Anim Toxins Incl Trop Dis. 2015;21:6. Epub 2015/03/27. doi:
488 10.1186/s40409-015-0005-z. PubMed PMID: 25810711; PubMed Central PMCID:
489 PMCPMC4373060.

30. Possas C, Lourenco-de-Oliveira R, Tauil PL, Pinheiro FP, Pissinatti A, Cunha RVD, et al. Yellow fever outbreak in Brazil: the puzzle of rapid viral spread and challenges for immunisation. Memorias do Instituto Oswaldo Cruz. 2018;113(10):e180278. Epub 2018/11/15. doi: 10.1590/0074-02760180278. PubMed PMID: 30427974; PubMed Central PMCID: PMCPMC6135548.

31. Soper FL. The eliminationo of urban yellow fever in the Americas through the eradication of Aedes aegypti. American journal of public health. 1963;53(1):1-16.

32. Nascimento RA, Montano, R. A. M. An assessment of illegal capuchin monkey trade in Bahia State, Brazil. Neotropical Biology and Conservation. 2013;8(2):79-87.

Supporting Information Legends

Technical Appendix. Contains additional Figure S1, Figure S2, Table S1 and Table S2. Figure S1: date of sample collection and number of genomes generated from each host genus. Figure S2. evolution of the wavefront distance from epidemic origin over time. Table S1: details of the YFV genomes generated in this study. Table S2: non-human primate yellow fever virus genome sequences from São Paulo, by municipality.

510

511 ## Figure Legends

512

513 **Figure 1**. Distribution of weekly YFV cases. **A**. NHP YFV cases diagnosed by RT-qPCR per week.
514 **B**. Distance in kilometres from the municipality of NHP YFV occurrence to São Paulo municipality
515 is plotted against date of sample collection.

516

517 **Figure 2**. Choropleth map of the distribution of confirmed NHP cases per municipality in São
518 Paulo state between July 2016 and February 2018. Triangles depict locations of human sequenced
519 cases; diamonds depict 14 municipalities in São Paulo state from which non-human primate
520 genome sequences have been generated. Red symbols correspond to earlier cases between July
521 2017 and February 2018. Names for key samples/isolates are shown (see also Figure 3).

522

523 **Figure 3**. Maximum likelihood phylogenetic tree of YFV in São Paulo, Brazil. The tree includes 46
524 newly generated isolates, and 2 previously published ones (M5 and M11 isolates). Bootstrap scores
525 are provided for well-supported nodes.

526

527 **Figure 4**. Reconstructed spatiotemporal diffusion of the São Paulo YFV clade. Phylogenetic
528 branches are mapped in space according to the location of phylogenetic nodes (circles). Arrows
529 show the cross-state movement of the virus from Minas Gerais followed by movement to the
530 Atlantic forest closer to present. Shaded regions show 95% credible regions of internal nodes.
531 Nodes and uncertainty regions are coloured according to time.

532

533

534

535

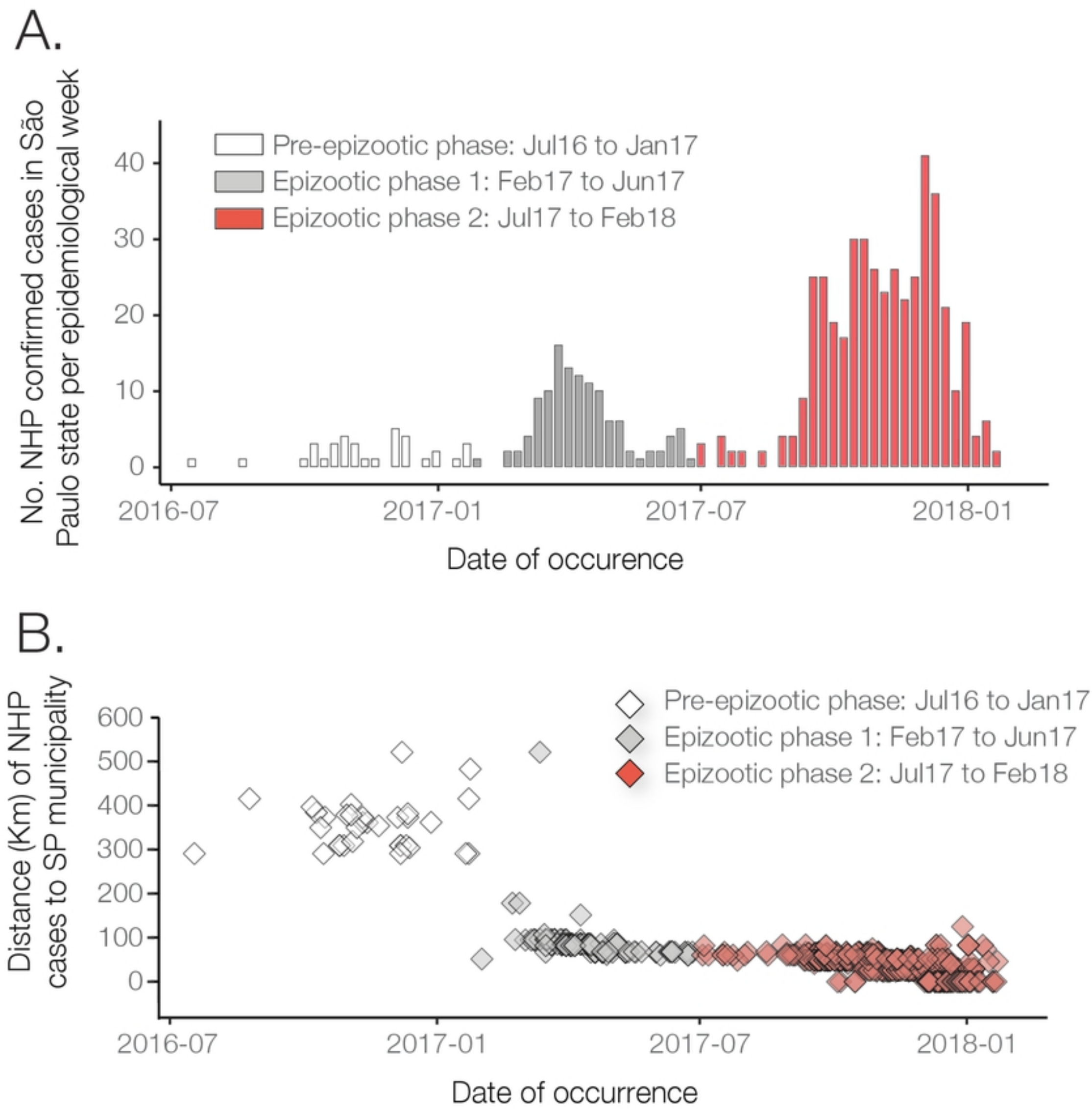
A.
No. NHP confirmed cases in São Paulo state per epidemiological week
Pre-epizootic phase: Jul16 to Jan17
Epizootic phase 1: Feb17 to Jun17
Epizootic phase 2: Jul17 to Feb18
0
10
20
30
40
2016-07
2017-01
2017-07
2018-01
Date of occurence
B.
Distance (Km) of NHP cases to SP municipality
Pre-epizootic phase: Jul16 to Jan17
Epizootic phase 1: Feb17 to Jun17
Epizootic phase 2: Jul17 to Feb18
0
100
200
300
400
500
600
2016-07
2017-01
2017-07
2018-01
Date of occurrence

Figure 1

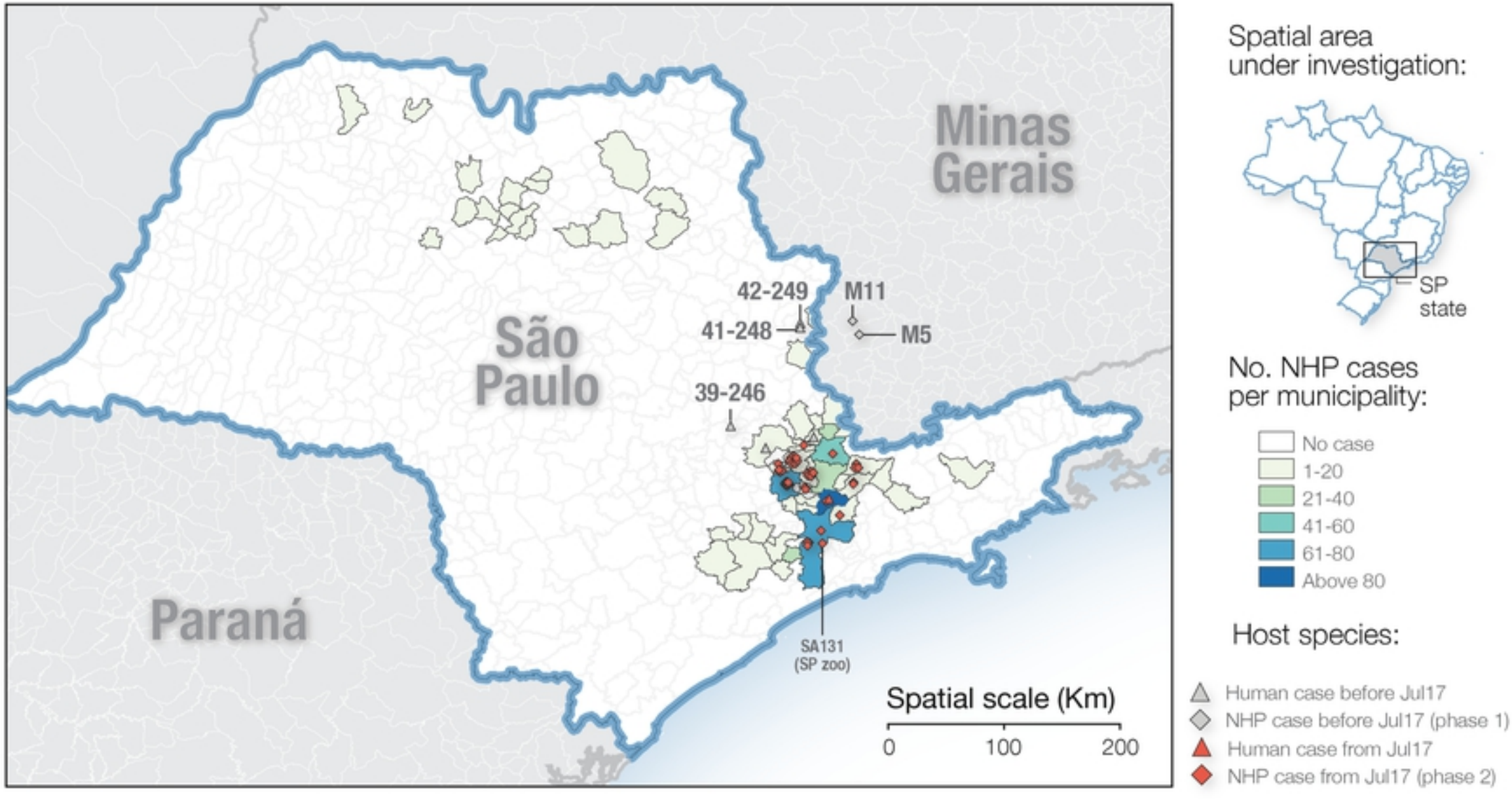
Minas Gerais
São Paulo
Paraná
42-249
M11
41-248
M5
39-246
SA131
(SP zoo)
Spatial scale (Km)
0
100
200
Spatial area under investigation:
SP state
No. NHP cases per municipality:
No case
1-20
21-40
41-60
61-80
Above 80
Host species:
Human case before Jul17
NHP case before Jul17 (phase 1)
Human case from Jul17
NHP case from Jul17 (phase 2)

Figure 2

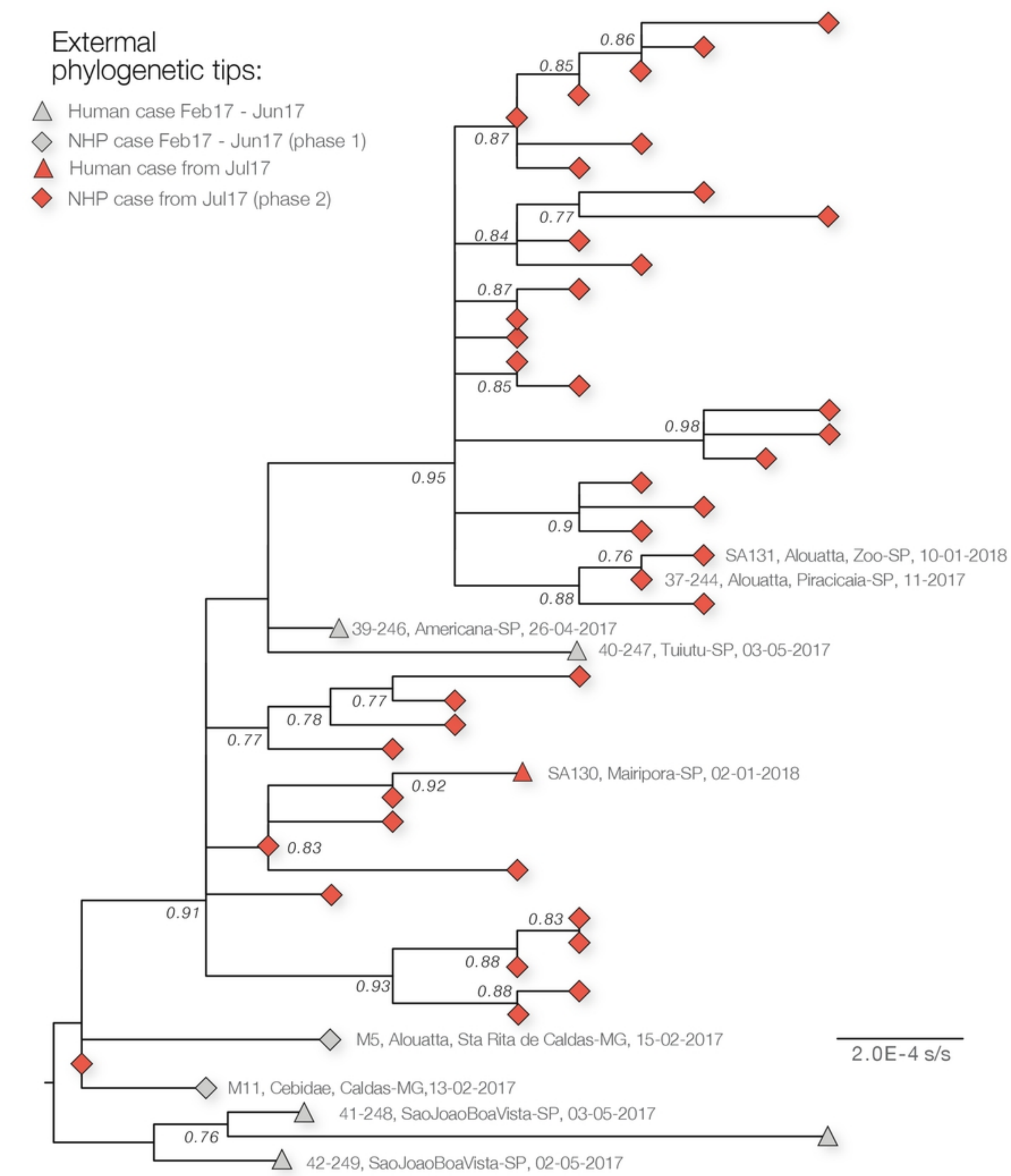
External phylogenetic tips:
Human case Feb17 - Jun17
NHP case Feb17 - Jun17 (phase 1)
Human case from Jul17
NHP case from Jul17 (phase 2)
0.86
0.85
0.87
0.77
0.84
0.87
0.85
0.98
0.95
0.9
0.76
SA131, Alouatta, Zoo-SP, 10-01-2018
37-244, Alouatta, Piracicaia-SP, 11-2017
0.88
39-246, Americana-SP, 26-04-2017
40-247, Tuiutu-SP, 03-05-2017
0.77
0.78
0.77
SA130, Mairipora-SP, 02-01-2018
0.92
0.83
0.91
0.83
0.88
0.93
0.88
M5, Alouatta, Sta Rita de Caldas-MG, 15-02-2017
2.0E-4 s/s
M11, Cebidae, Caldas-MG,13-02-2017
41-248, SaoJoaoBoaVista-SP, 03-05-2017
0.76
42-249, SaoJoaoBoaVista-SP, 02-05-2017

Figure 3

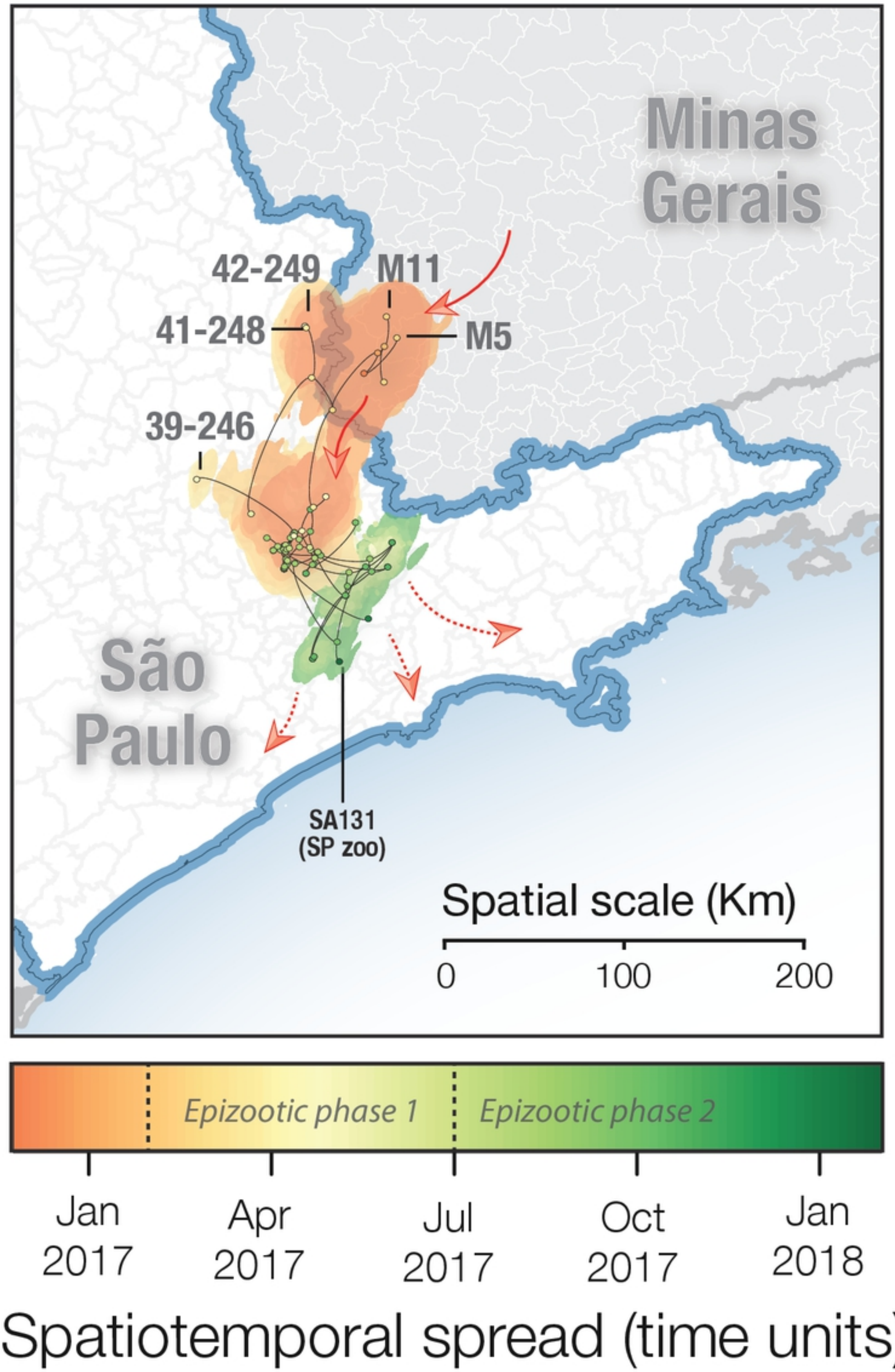
Minas Gerais
42-249
M11
41-248
M5
39-246
São Paulo
SA131
(SP zoo)
Spatial scale (Km)
0
100
200
Epizootic phase 1
Epizootic phase 2
Jan 2017
Apr 2017
Jul 2017
Oct 2017
Jan 2018
Spatiotemporal spread (time units)

Figure 4