nature microbiology

Article

https://doi.org/10.1038/s41564-026-02302-w

Evolution and spillover dynamics of yellow fever at the forest–urban interface in Brazil

Received: 4 July 2025

Accepted: 19 February 2026

Published online: 11 March 2026

Check for updates

A list of authors and their affiliations appears at the end of the paper

Yellow fever virus (YFV) continues to threaten human and wildlife populations in the Americas, yet its transmission at the forest–urban interface remains unclear. Here we integrate ground- and canopy-level mosquito surveillance, systematic monitoring of non-human primate carcasses and viral metagenomics to describe the dynamics of a sylvatic YFV outbreak in a 186-hectare Atlantic Forest fragment embedded within metropolitan São Paulo, Brazil, between 2017 and 2018. Our analyses reveal that transmission was primarily driven by a single genetic cluster introduced during a period of high abundance of the main vector, *Haemagogus leucocelaenus* mosquitoes. A near-complete hepatitis A virus genome was detected in a YFV-infected howler monkey, suggesting potential co-infections at the human–wildlife interface. Phylogenetic and epidemiological modelling estimated a basic reproduction number, R_0, for sylvatic yellow fever of 8.2 (95% CI 5.1–12.2), substantially higher than previous estimates for urban outbreaks. Our findings demonstrate that multisource surveillance could provide actionable early warnings in regions at risk for zoonotic spillover.

Yellow fever is a mosquito-borne viral haemorrhagic disease with an estimated 47% case fatality rate among non-vaccinated humans[1]. It is caused by the yellow fever virus (YFV, *Orthoflavivirus flavi*), a single-stranded 10.86 kb RNA flavivirus related to dengue, Zika and West Nile viruses. In Africa, YFV circulates in alternating enzootic cycles maintained among non-human primates (NHPs) and small mammals via *Aedes africanus*, and during urban cycles the virus is transmitted by *Aedes aegypti*[2]. In the Americas, urban transmission has not been reported since 1942[3], and the earliest evidence of a sylvatic YFV cycle dates to 1932[4]. During sylvatic cycles, arboreal *Haemagogus* and *Sabethes* mosquitoes transmit YFV among susceptible NHPs that act as amplifying hosts. These mosquitoes occasionally feed at ground level, increasing the risk of spillover to humans at forest–urban interfaces[5].

In the Americas, YFV persists in an enzootic cycle in the Amazon and periodically spreads southeastward to densely populated states[6,7]. In 2015, it reached southeastern Brazil, causing 772 human deaths and over 15,000 NHP deaths between 2016 and 2018[8]. Entomological surveys identified *Haemagogus leucocelaenus* and *Haemagogus janthinomys* mosquitoes as primary vectors[5]. Among Brazilian NHPs, howler monkeys (*Alouatta guariba*, family Atelidae) are particularly susceptible to YFV[9,10]. Because NHP cases often precede human cases by several days[11], tracking NHP cases is critical for guiding ring vaccination programmes, particularly in periods of vaccine shortage.

YFV spillover risk is highest where human activity overlaps with suitable ecologies for vectors and NHPs[12]. Recurrent sylvatic outbreaks, detection of YFV in *Aedes albopictus* and laboratory competence of urban vectors have long raised concerns about renewed urban transmission[13]. Howler monkey populations, once widely distributed across eight states, continue to decline from repeated epizootics, habitat loss and hunting[14]. Tracking YFV in arboreal mosquitoes and NHPs is hampered by logistical constraints, absence of rapid diagnostic tests and challenges in sampling at tree-canopy levels[15]. At broad scales, warmer and wetter conditions and fragmented forest–cropland landscapes rich in susceptible NHPs may accelerate YFV spread and spillover[16]. Yet, at higher resolutions, YFV transmission dynamics remains poorly understood, particularly in high-risk areas such as forest fragments near large urban centres. These gaps reflect challenges in entomological and epidemiological

e-mail: jtelles@pasteur.saude.sp.gov.br; lfmucci@gmail.com; nfaria@ic.ac.uk

https://doi.org/10.1038/s41564-026-02302-w

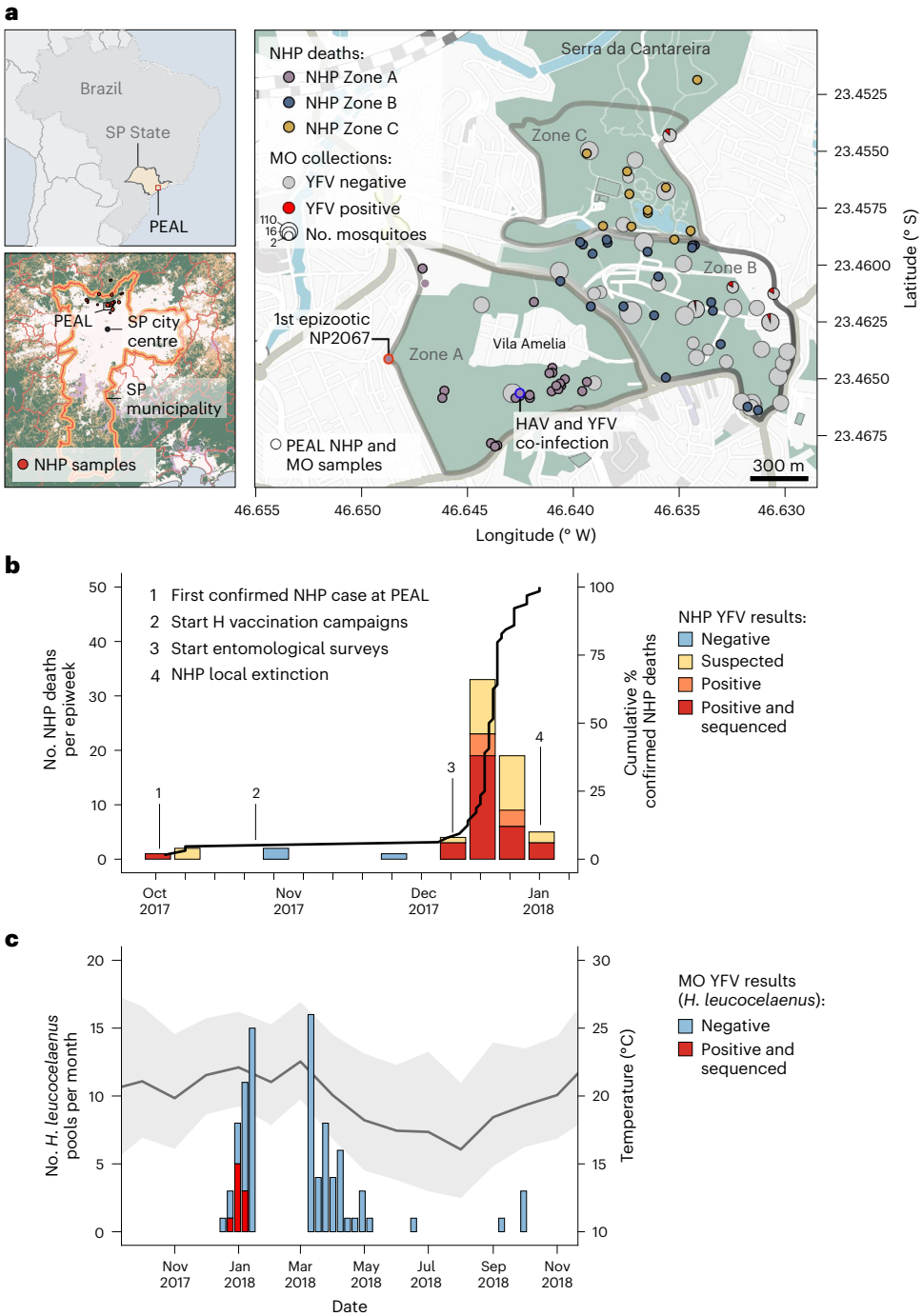


Fig. 1 | Spatial and temporal context of the YFV epizootic in PEAL's State Park (São Paulo, Brazil). a, Location of PEAL within Brazil (top left) and within São Paulo metropolitan area (bottom left; surrounding biomes from MapBiomas[120]; forested areas, green; farming, orange; freshwater, purple; non-forested natural formations, light brown). Right: spatial distribution of non-human primate (NHP) deaths and mosquito sampling across PEAL management zones (A–C). Red-outlined circle marks the first PEAL epizootic case (NP2067); blue-outlined circle marks the NHP co-infected with YFV and HAV (NP2754). Pie charts indicate mosquito pool sites, with the red sector indicating YFV-positive pools and circle area proportional to the number of sampled pools (smallest circle $n = 2$). NHP deaths are scaled by the number of carcasses (smallest $n = 1$). Note that **a** shows the collection of all mosquito taxa, not only *H. leucocelaenus*. Zones A, B and C correspond to different management areas within PEAL State Park, as defined by the official map from the Government of São Paulo. **b**, Epidemic curve of free-ranging *A. guariba* deaths at PEAL by epidemiological week from first epizootic case (NP2067). Bars are stacked by outcome (negative, suspected, positive, positive + sequenced); black line shows the cumulative percentage of deaths among the free-ranging *A. guariba* population at PEAL. Numbers indicate key events. **c**, Timeseries of non-blood-fed *H. leucocelaenus* pools collected at PEAL, coloured by RT–PCR YFV result (negative, positive + sequenced). The dark grey line and light grey shading shows mean temperature (°C) and minimum/maximum ranges from ERA5-Land[121], illustrating seasonal conditions during mosquito surveillance. H, human; MO, mosquito; SP, São Paulo.

data collection, and the absence of frameworks for modelling YFV zoonotic transmission[17].

We integrated field mosquito and NHP sampling with viral metagenomics, using a pathogen-agnostic sequencing approach enabling recovery of complete viral genomes[18], and combined these data with phylogenetic and transmission models to characterize YFV epizootic spread and dynamics during the first YFV outbreak in metropolitan São Paulo.

https://doi.org/10.1038/s41564-026-02302-w

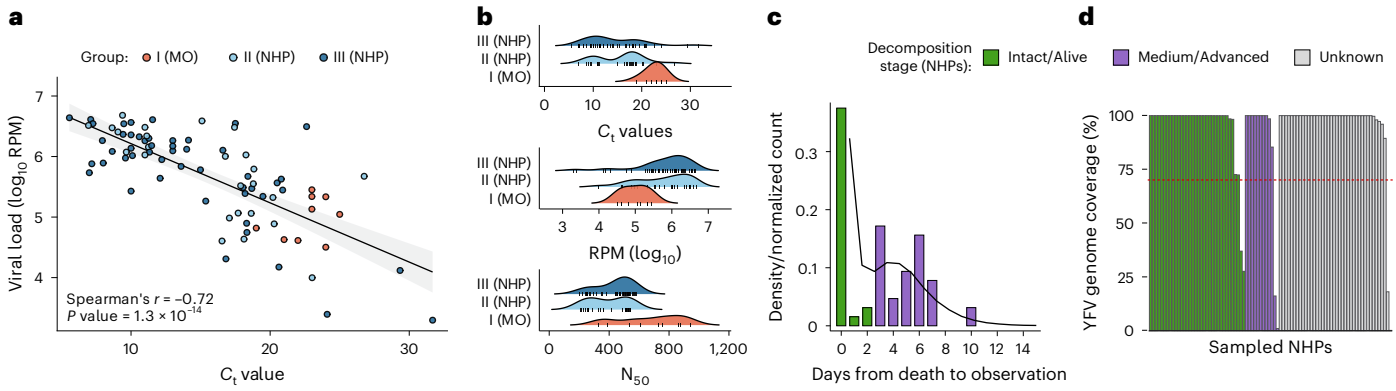


Fig. 2 | Relationship between RT–PCR C_t values, viral load, timing of observations and viral detection. a, Scatterplot of RT–PCR C_t values versus YFV mapped RPM for each sample group (Group I: *H. leucocelaenus*, Group II: PEAL NHPs, Group III: background NHPs), illustrating the inverse relationship between C_t values and viral abundance (Spearman's *r*, two-sided). Fitted line and shaded band show linear trend and 95% confidence interval. **b**, Density ridge plot displaying the distributions of C_t values, viral load (RPM) and N_{50} among the different sample groups. **c**, Empirical and estimated distributions of the delay between *A. guariba* death and notification. The black line shows the fitted probability density (gamma distribution). Bars indicate the empirical distribution of days from death to observation, coloured by decomposition stage: green for intact or recently dead animals (0–2 days post death) and purple for medium or advanced decomposition (≥3 days post death) (see Methods). **d**, YFV genome coverage (% of reference) across 89 samples for which viral metagenomics sequencing was attempted (35 from intact or recently dead NHPs; 12 from NHPs found at medium or advanced decomposition; and 43 NHPs for which stage of decomposition was unknown). Samples were grouped by decomposition stage. The horizontal dashed red line marks the 70% whole-genome coverage threshold used for inclusion in phylogenetic analyses. Spearman's correlations and Wilcoxon tests were two-sided; exact *P* values are reported in main text.

Results

Early outbreak investigations

Between October 2017 and January 2018, an unprecedent yellow fever outbreak occurred at Parque Estadual Alberto Lofgren (PEAL) (Fig. 1a). The first detected epizootic, a juvenile *Alouatta guariba* carcass (NP2067) found on 9 October 2017, marked the first case in the Greater Metropolitan Region of São Paulo (home to >23 million inhabitants) and led to PEAL's closure and initiation of human vaccination campaigns in adjacent neighbourhoods on 20 October 2017 (Fig. 1b). Entomological surveys near NP2067 on 21 October 2017 detected 11 mosquito species, including *A. albopictus*, but none were YFV positive, and *Haemagogus* and *Sabethes* mosquitoes were absent (Fig. 1a).

Before the outbreak, *A. guariba* density at PEAL was 39.08 individuals km^{-2}, the highest reported for this species in the Atlantic Forest[19–22]. By 5 January 2018, the entire population of *A. guariba* became locally extinct (Fig. 1b). Retrospective testing found YFV RNA in 39 of 67 carcasses (58.2%), with 3 YFV RNA-negative and 25 unsampled due to advanced decomposition (Fig. 1c). The case fatality ratio (CFR) at PEAL was 0.96 (95% Bayesian confidence intervals [BCI]: 0.91–1.00) including suspected cases ($n = 64$), and 0.58 (95% BCI: 0.43–0.73) based on confirmed cases ($n = 39$).

H. leucocelaenus as the primary YFV vector at PEAL

We sampled mosquitoes across 39 locations, collecting 2,231 individuals (2,013 females) from 24 species and 10 genera (Supplementary Table 1, Extended Data Fig. 1 and Supplementary Fig. 1). Among 753 pools of non-blood-fed females, YFV RNA was detected exclusively in *H. leucocelaenus*, with 9 of 87 pools collected at ground-level locations (1–5 mosquitoes per pool; median quantitative PCR with reverse transcription (RT–qPCR) cycle threshold (C_t) = 23, range: 19–25) (Supplementary Table 1). To contextualize these findings, we compared C_t values across 118 pools representing 12 distinct species drawn from publicly available datasets (Supplementary Table 2). C_t values differed strongly by species, with *H. leucocelaenus* < *H. janthinomys* < other vectors (global Kruskal–Wallis, $P = 1.1 \times 10^{-11}$; pairwise $q = 0.033$, 3.5×10^{-11} and 1.7×10^{-6}, respectively). A model adjusting for RT–PCR assay and pool size estimated −2.6 C_t values for *H. leucocelaenus* relative to other vectors ($P = 0.034$; d.f. = 112), consistent with higher viral RNA in *H. leucocelaenus* pools (Extended Data Fig. 2). The minimum infection rate (MIR) for *H. leucocelaenus* was 58.8 per 1,000 (5.88%). Adjusting for pool sizes using a maximum-likelihood approach yielded a similar MIR of 6.3% (95% CI: 3.1–11.3%). The MIR at PEAL was high compared to other Brazilian studies/sites[5], and we found no evidence that MIR differs across species (two-sided Kruskal–Wallis $P = 0.058$; all pairwise false discovery rate (FDR)-adjusted tests non-significant; Extended Data Fig. 3 and Supplementary Table 3).

Environmental drivers of vector abundance

To understand the environmental context in which the outbreak unfolded, we examined the climatic drivers of mosquito abundance. Mean temperature was the primary correlate of *H. leucocelaenus* abundance at PEAL (negative-binomial generalized linear model (GLM); pseudo-R^2 = 0.83–0.86; Fig. 1c and Supplementary Tables 4–6). In genus-specific models including temperature and cumulative rainfall, temperature remained strongly associated with *Haemagogus* abundance (rate ratio, RR per 1 °C = 2.82, 95% CI: 1.73–4.73), while rainfall contributed little (RR = 0.96, 95% CI: 0.87–1.05). Because temperature and rainfall were strongly correlated (Supplementary Fig. 2), rainfall effects should be interpreted cautiously. Similar temperature associations were observed for *Aedes* (RR = 2.81, 95% CI: 1.81–4.59; pseudo-R^2 = 0.75) and *Limatus* (RR = 2.71, 95% CI: 1.63–4.69; pseudo-R^2 = 0.58), with no clear added effect of rainfall. For *Culex*, associations were not significant (temperature RR = 0.81, 95% CI 0.42–1.56; rainfall RR = 1.08, 95% CI 0.91–1.27; pseudo-R^2 = 0.11) (Supplementary Table 7).

Viral metagenomics and co-infection with hepatitis A virus (HAV)

To characterize the performance of our sequencing approaches and to identify any additional viral signals present in the samples, we examined viral metagenomic data and its concordance with tiled-amplicon sequencing. First, we generated viral metagenomic sequencing data from mosquito pools and NHP tissues[18] and compared its performance with a tiled-amplicon approach[9]. Metagenomics achieved higher coverage across 56 paired samples (median 99.89% vs 85.99%; mean paired difference 10.75 percentage points, 95% CI 6.70–14.80; two-sided Wilcoxon signed-rank test $P = 1.18 \times 10^{-6}$) (Supplementary Table 8). Coverage differences widened at lower C_t values ($n = 52$, $r = -0.78$, $P = 1.25 \times 10^{-11}$) (Supplementary Fig. 3), partly reflecting primer–template mismatches at 6 sites in the YFV500/V1 scheme (6R, 10L, 12L/12R, 20R, 23R, 27L) and consistent with sequence divergence (Extended Data Fig. 4). In addition, sample pre-treatment (centrifugation, filtration and Turbo

https://doi.org/10.1038/s41564-026-02302-w

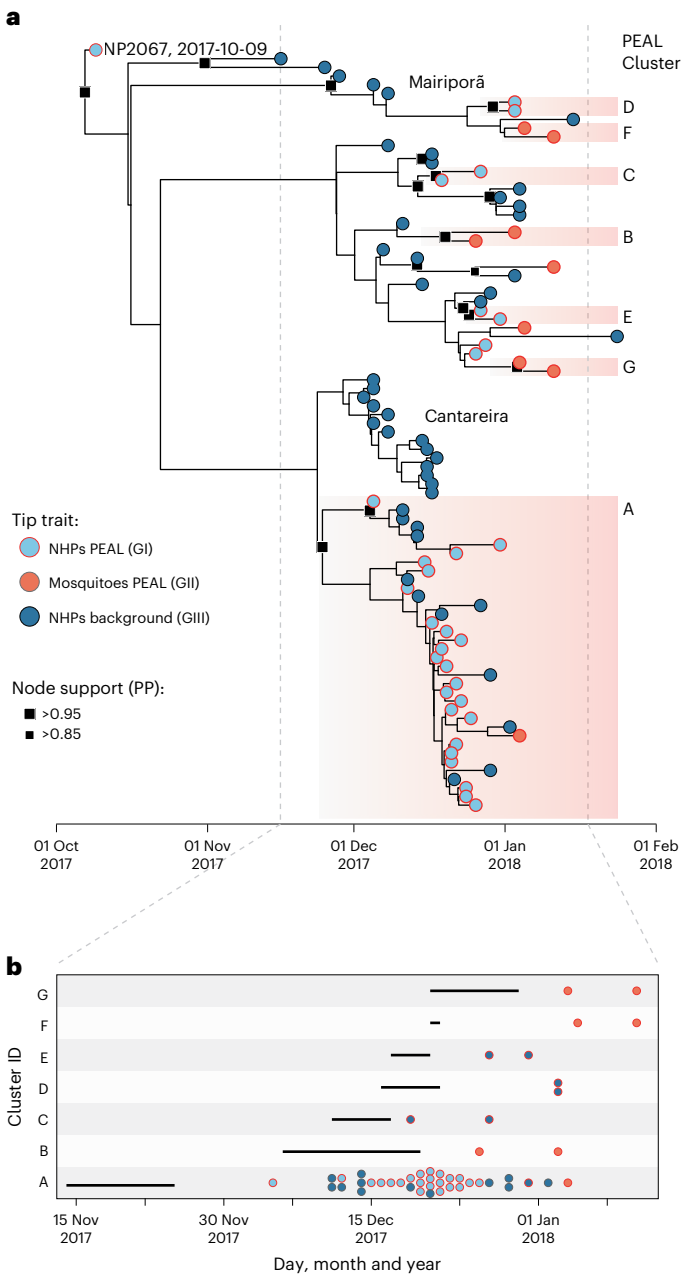


Fig. 3 | Temporal dynamics of YFV cluster introductions in PEAL State Park. **a**, Time-calibrated phylogeny of PEAL including local background genomes ($n = 88$), spanning a 107-day sampling window. Tips are annotated by host (NHP or mosquito) and whether they were sampled inside the park (ring). Squares mark node support (posterior probability, PP; large >0.95, small >0.85). Orange boxes highlight the seven PEAL clusters (A–G). The grey vertical dashed lines correspond to the timeframe expanded in **b**. **b**, Seeding and growth of each PEAL cluster. For each cluster, the left endpoint of the solid horizontal line represents the median time of the parent node of the most recent common ancestor (TMRCA), while the right endpoint marks the median TMRCA, corresponding to the onset of sustained local transmission. Circles correspond to dated samples in each cluster (same tip annotations as in **a**).

DNase) in the viral metagenomics workflow probably further reduced non-viral nucleic acids, further improving viral metagenomic recovery.

Sequencing of 98 YFV RT–PCR-positive samples yielded 88 complete and near-complete YFV sequences with ≥70% coverage. This included all 9 mosquito pools (Group I), 32/36 PEAL NHPs (Group II) and 47/53 background NHPs (Group III) (Supplementary Fig. 4). Overall coverage average, at a sequencing depth of 10× or greater, was high (99.90%, range 72.36–99.98%; Supplementary Tables 8 and 9). Coverage was similarly high in mosquitoes ($n = 9$, median 97.5%, range: 76.2–99.9%) and NHPs ($n = 89$, median 99.90%, range: 0–99.9%). Our metagenomic workflow also generated sufficient mitochondrial reads for cytochrome *c* oxidase subunit I (COI) metabarcoding, enabling species confirmation for 32 *A. guariba* and 4 *H. leucocelaenus* pools.

We also recovered a near-complete HAV genome (73.75% coverage at 10×; lineage HAV I.A_ab) from a YFV-positive female *A. guariba* (NP2754). As *A. guariba* are primarily canopy dwelling, enzootic HAV has not been reported in free-ranging Neotropical primates[23], and HAV is not handled in our laboratory, this finding most probably reflects environmental exposure to human faecal contamination near the collection site (~100 m from the nearest residence, Fig. 1a). The NP2754 HAV sequence showed 97.1% identity and clustered with strong statistical support with a contemporaneous human strain from São Paulo (Extended Data Fig. 5), consistent with environmental exposure rather than laboratory contamination.

Association between host and vector viral loads and cycle thresholds

RT–PCR C_t values were inversely correlated with YFV reads per million (RPM, $\log_{10}$) ($n = 84$, two-sided Spearman's $r = -0.72$, $P = 1.3 \times 10^{-14}$; Fig. 2a). NHPs had lower C_t values than *H. leucocelaenus* pools ($n = 75$, median 14, range: 5.57–31.68 vs $n = 9$, 23, range: 19.00–25.00; Wilcoxon rank-sum $P = 3.5 \times 10^{-5}$; Fig. 2b), although mosquitoes may partly contribute to higher C_t values in vectors. NHPs also had higher median RPM compared with mosquitoes (median RPM = 5.88, range: 2.99–6.68, $n = 89$ vs median RPM = 5.04, range: 4.50–5.45, $n = 9$; Wilcoxon rank-sum $P = 0.003$). Mosquito pools showed higher N_{50} values (732 bp, range: 331–945) than NHP tissues (448 bp, range: 201–585; Wilcoxon rank-sum $P = 9.1 \times 10^{-4}$; Fig. 2b and Supplementary Table 8), indicating longer viral fragments in vector libraries.

YFV sequencing performance across NHP carcass decomposition stages

Current epizootic investigation guidelines in Brazil recommend sampling NHPs within 24 h of death[24]. At PEAL, reporting delays were short and right skewed, with most carcasses reported and sampled within 0–2 days (early-stage, intact/alive) and with progressively fewer reports at ≥3 days (later-stage, medium/advanced decomposition) (Fig. 2c). C_t, N_{50} and RPM did not differ significantly between NHPs sampled as intact/alive ($n = 34$) and those in medium or advanced decomposition ($n = 12$; Fig. 2d); decomposition stage was unavailable for 43 NHPs (Supplementary Table 8). Coverage was similarly high across stages (early-stage: median 99.90%, range 27.54–99.91%; later-stage: median 99.90%, range 0.96–99.91%), with >70% genome recovery in 94.12% of early-stage and 83.33% of later-stage samples (Fig. 2d).

Phylogenetic and transmission dynamics

Maximum likelihood (ML) and Bayesian phylogenetic analyses placed the PEAL genomes within the previously identified YFV_{SP}[25] clade. Across three datasets (Brazil-wide, $n = 1{,}063$, sampling interval 15.2 years; São Paulo, $n = 450$, 3.7 years; PEAL, $n = 88$ 0.3 years), ML trees consistently identified a dominant PEAL cluster (hereafter named 'Cluster A') with strong support (bootstrap 89–98) (Fig. 3a, and Supplementary Figs. 5 and 6). Molecular clock analyses dated the emergence of YFV in São Paulo state to mid-May 2016 (95% BCI: Feb 2016–Jul 2016; Supplementary Fig. 6). Within PEAL, 18/23 NHP genomes were identical, and one mosquito genome (ID81) differed by a single nucleotide (Fig. 3a). Focusing on the PEAL dataset, which showed adequate temporal signal (root-to-tip divergence $R = 0.39$), the within-outbreak evolutionary rate was 3.7×10^{-3} substitutions per site per year (95% BCI: 9.6×10^{-4} to 7.3×10^{-3}), similar to previous estimates[9]. This corresponds to ~11 days of waiting time between mutations (95% BCI: 4–20 days).

https://doi.org/10.1038/s41564-026-02302-w

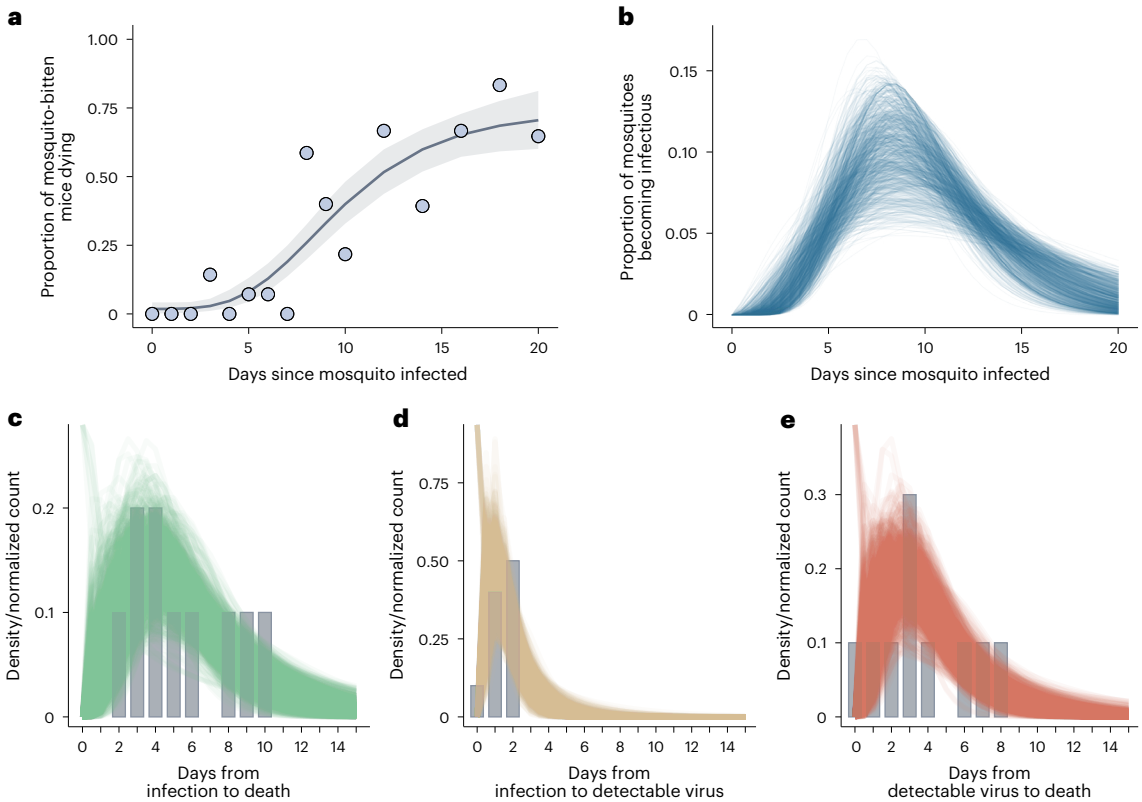


Fig. 4 | Empirical and estimated distributions informing key temporal parameters of YFV epidemic dynamics in a sylvatic cycle. a, Observed (points, data from ref. 116) and posterior predictive estimates (grey line and shaded ribbon show posterior median and 95% credible interval) of the proportion of mice that died following feeding by *H. leucocelaenus* mosquitoes inoculated with YFV; datapoints for days 1 and 2 were excluded since they probably reflected residual inoculum rather than a completed extrinsic incubation process. **b**, Estimated EIP of *H. leucocelaenus* mosquitoes based on data reported in ref. 116, with coloured lines representing draws from the posterior gamma distribution fitted to that data. **c**, Empirical and estimated distributions for the number of days from *A. guariba* exposure to YFV and subsequent death. **d**, Empirical and estimated distributions for the number of days from *A. guariba* exposure to YFV and becoming infectious. **e**, Empirical and estimated distributions for the number of days from *A. guariba* becoming infectious and death. In **c**, **d** and **e**, bars represent the empirical normalized frequencies based on data presented in ref. 114, and coloured lines represent draws from the posterior gamma distribution fitted to that data.

Across PEAL, we identified 7 epizootic clusters (2–36 sequences) and 5 singletons within a single transmission season (Fig. 3, Methods). Interestingly, the early NP2067 case was a singleton outside all clusters, demonstrating that this case did not contribute to onward transmission at PEAL. Given the strong time dependency of evolutionary rates across broader timescales (see Supplementary Fig. 7), dating analyses were based on the PEAL and background dataset. Cluster A was the earliest introduction (around 14–25 Nov 2017), dominated the outbreak ($n = 36$; 24 group II, 12 group III) and persisted for 40–51 days, whereas later introductions generated only small, shorter-lived clusters (Fig. 3b).

Transmission modelling and key epidemiological parameters

To characterize the transmission dynamics of the PEAL outbreak and assess the epidemic potential of YFV in NHPs, we developed an individual-based transmission model (IBM) that tracked each *A. guariba* at PEAL through infection, infectiousness and death, explicitly incorporating carcass-detection delays (Fig. 2c) and uncertainty in the extrinsic incubation period (EIP) of *H. leucocelaenus*. Five empirical distributions underline our framework: (1) delay from mosquito feeding to host death, capturing downstream effects of vector-mediated transmission (Fig. 4a); (2) EIP of *H. leucocelaenus* (median: 10 days, 95% CI: 8–12 days) (Fig. 4b); (3) delay from *A. guariba* death to carcass detection (median: ~2.1 days, 95% CI: 1.5–3.0 days) (Fig. 4c); (4) intrinsic incubation period (IIP) in *A. guariba* (median: 5.2 days, 95% CI: 4.3–6.2 days) (Fig. 4d); and (5) *A. guariba* infectious period (time from detectable viremia until death) (median: 3.8 days, 95% CI: 3.0–4.5 days) (Fig. 4e). The model parameterization and data sources are detailed in Extended Data Table 1.

Estimation of reproduction numbers and outbreak dynamics

To estimate the basic reproduction (R_0) and characterize the transmission dynamics of the PEAL outbreak, we calibrated our IBM using phylogenetic data, anchoring Clade A's introduction around 19 Nov 2017 and accounting for the 5-week absence of confirmed cases after NP2067 (Fig. 1b). This allowed us to constrain the outbreak start and focus on the period of sustained transmission.

Integrating phylogenies with epidemiological modelling yielded $R_0 = 8.2$ (95% CI: 5.1–12.2) (Fig. 5a,b). This estimate is higher than estimates for urban outbreaks[26,27], but falls within the range reported for Brazilian sylvatic settings[28]. Our posterior exploration across a two-dimensional grid of outbreak starting dates and reproduction numbers identified high-likelihood regions centred on introductions in the first weeks of November, with values typically above 6 (Fig. 5a). Our results were robust to alternative outbreak start dates, including scenarios with later epidemic onset (Supplementary Fig. 8). Because vectors are not modelled explicitly, R_0 represents the product of the number of vectors infected by infectious *A. guariba* and the number of *A. guariba* subsequently infected by infectious vectors (that is, it integrates both vector and NHP transmission).

The posterior trajectories of the time-varying effective reproduction number, R_e, which accounts for depletion of susceptible hosts, show a rapid decline in transmission over the course of the epidemic

https://doi.org/10.1038/s41564-026-02302-w

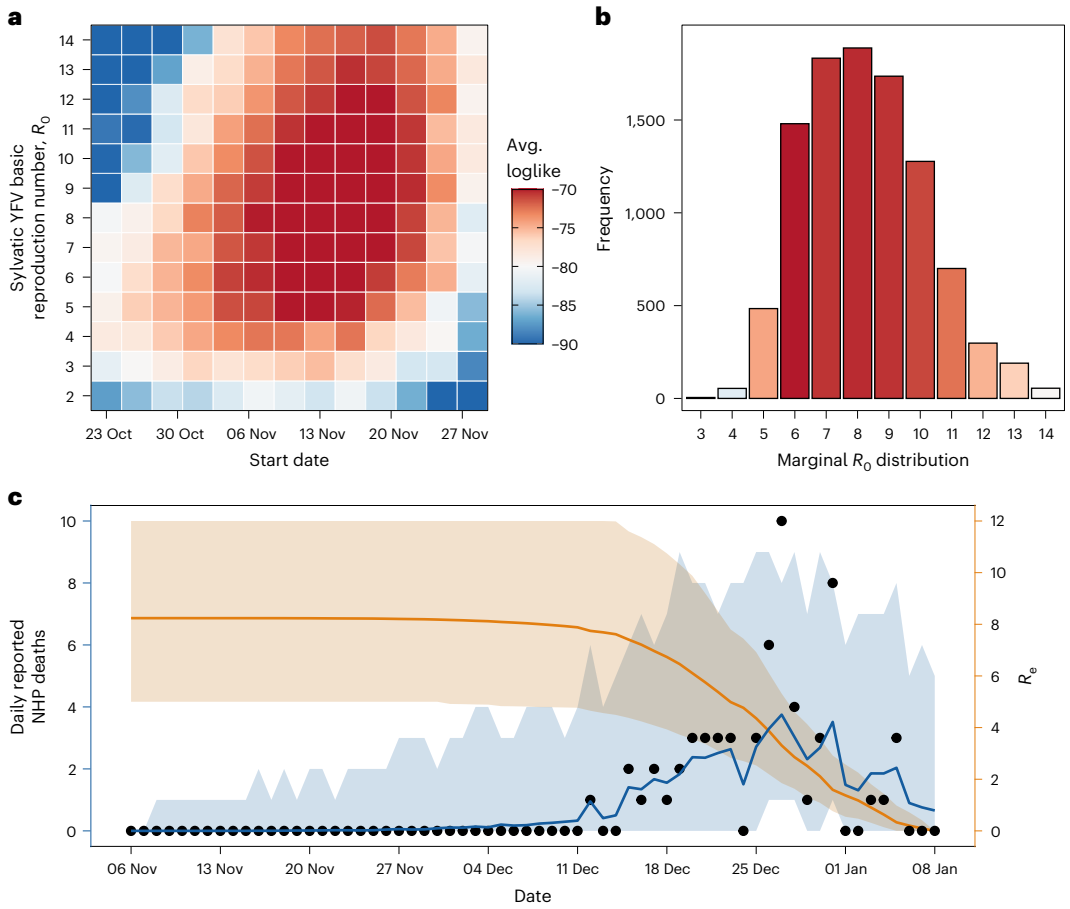


Fig. 5 | Transmission dynamics and epidemic potential of sylvatic YFV in PEAL. a, Average log-likelihood (colour scale) for each joint combination of the basic reproduction number (R_0, *y* axis) and outbreak start date (*x* axis). Warmer tones indicate parameter pairs with higher posterior support. **b**, Marginal posterior distribution of R_0 integrating over all start-date hypotheses shown in **a**. Bar height denotes posterior sample frequency; colours match the log-likelihood scale used in **a**. **c**, Model fit to daily reported non-human primate deaths (black circles, left axis) and the corresponding effective reproduction number over time (R_e, orange line, right axis). Blue line and shaded ribbon give the posterior median and 95% credible interval for model-predicted deaths; the orange ribbon shows the 95% credible interval for R_e.

(Fig. 5c). This decline is consistent with the extensive infection and mortality that ultimately resulted in local extinction of *A. guariba* at PEAL.

Repeating seeding but rare take-off of viral introductions

To assess the impact of different viral introductions, we used a branching-process model incorporating the IIP in *A. guariba* and EIP in *H. leucocelaenus*[29]. For Clade A (23/32, 71.9% of PEAL's sequenced NHP cases), an estimated 1.5 host–vector–host generations (95% CI: 1–3) were required to generate the observed outbreak from a single infected mosquito. In contrast, the three minor introductions (6/32, 18.1% of sequenced cases) each contributed minimally, producing <0.4 generations of onward transmission.

Discussion

We integrated ground and canopy entomology, systematic NHP surveillance, viral metagenomics and phylodynamics with transmission modelling to describe a sylvatic yellow fever outbreak at the forest–urban interface. Despite multiple introductions into PEAL, transmission was dominated by a single lineage introduced during a period of high *H. leucocelaenus* abundance, leading to rapid local extinction of *A. guariba*. Limited genetic divergence within this lineage and a fast within-outbreak evolutionary rate indicate a brief, intense transmission chain. Our modelling yielded an $R_0 \approx 8.2$ (95% CI 5.1–12.2), substantially higher than estimates from *Aedes*-mediated urban outbreaks[26,27].

Temperature was the strongest correlate of *H. leucocelaenus* abundance, consistent with effects on development, survival and biting frequency[30], while rainfall added little explanatory power due to collinearity (Supplementary Fig. 2). The multibite, multihost feeding behaviour across vertical strata of *Haemagogus*[31] aligns with our detection of YFV-positive *H. leucocelaenus* at ground level. Higher YFV viremia in *H. leucocelaenus* compared with other vectors, together with high MIRs at PEAL (among the highest reported for the Atlantic Forest[5,32]) (Extended Data Fig. 2) and high *A. guariba* density[19], probably accelerated NHP decline. Other primate taxa (*Callithrix*, *Sapajus*) showed no deaths or evident infections, consistent with lower susceptibility in shared habitats[33]. After *A. guariba* became locally extinct, vector infection rates quickly declined, and we found no evidence for onward YFV circulation despite reports of vertical transmission in *H. leucocelaenus*[34] and the presence of *A. albopictus* and *A. aegypti* at PEAL.

Metagenomic sequencing provided several advantages during the investigation. It outperformed tiled-amplicon sequencing, resolved vector and host identities, and revealed an unexpected HAV co-infection in *A. guariba*. RPM tracked viral load and may serve as a surrogate when RT–PCRs are unavailable, or when molecular diagnostic performance is affected by primer–template mismatches. Recovery of YFV genomes from carcasses ≥3 days post mortem indicates that sampling windows could be extended to reduce missed epizootics. These results also reinforce the need to update tiled-amplicon schemes with contemporaneous genomes, as is standard for SARS-CoV-2 (ref. 35).

Ecological context and stochasticity determined whether introductions resulted in sustained transmission or self-limited events. The dominant cluster probably arose through founder effects and ecological bottlenecks, with lineage fitness shaped by local host–vector interactions[36]. This aligns with reservoir–host spillover theory, which predicts frequent fade-outs unless ecological and epidemiological conditions support onward transmission[17]. NP2067 was a self-limited event, occurring in a sector of PEAL with no detectable *H. leucocelaenus* during thermally unfavourable conditions (Fig. 1c). In contrast, later introductions during warmer periods and higher vector abundance generated several distinct transmission clusters. Given the absence of earlier cases in Greater São Paulo, short-range primate movement is unlikely to explain NP2067. Long-distance seeding into PEAL and elsewhere[37,38]—via asymptomatic carriers or viraemic individuals[6], transported vectors or NHPs[39], or windborne dispersal of infected mosquitoes[40]—remains plausible. *H. leucocelaenus* live for ~21 days and can disperse >5 km (refs. 41,42), and wind speeds up to 6.5 m s^{-1} during the outbreak could have transported mosquitoes tens of kilometres within hours[43], well within ~25 km of earlier foci (Atibaia, Jundiaí). Purpose-built aerial sampling[44] could help determine the contribution of windborne movement of YFV-competent mosquitoes.

Recent YFV resurgence near urban areas in 2024/2025 highlights the need for sustained NHP and vector surveillance, even after apparent fade-outs. At forest–urban boundaries, where spillover risk is elevated, viral metagenomics could detect YFV and other arboviruses (for example, St Louis encephalitis virus, Mayaro virus[45–47]) while providing insight into host and vector feeding patterns[48]. Detection of YFV in *Callithrix* in urban Minas Gerais further underscores ongoing spillover risk in densely populated areas of southeast Brazil[49]. With expanding *Callithrix* populations, proximity to the Cantareira Forest (PEC) and São Paulo metropolis, and with >1.6 million annual visitors, PEAL remains a high-risk area for YFV.

These considerations also have implications for the conservation and management of susceptible primates. *A. guariba* populations in Atlantic Forest fragments are endangered[50], and immunization strategies for selected NHPs (including golden-lion and golden-headed lion tamarins[51]) may warrant evaluation[52]. Under conditions similar to PEAL ($R_0 \approx 8.2$) and assuming 99% vaccine efficacy, ~89% coverage would be required to prevent outbreaks. This illustrates both the potential value and the ethical, ecological and logistical challenges of intervening in wild primate populations, including achieving high coverage in fragmented habitats and engaging local communities and stakeholders.

Our study has several limitations. The host-death delay distribution was derived from mouse data and may differ in NHPs. We assumed that mid-November cases reflected a single introduction, although this clade comprises about two-thirds of sequenced NHP PEAL genomes. Partial carcass recovery and modest sample sizes may bias cluster inference, although cross-dataset concordance mitigated this. Gaps in mosquito sampling, particularly at the canopy level, highlight the need for automated trapping with integrated meteorological logging[53]. Moreover, the absence of IgM serology and HAV immunohistochemistry prevented confirmation of HAV infection or exposure. Finally, R_0 estimates are sensitive to ecological variation in primate density, contact rates and *Haemagogus* abundance, which can vary widely; previous studies report wide variation in R_0 globally[26] and within Brazil (2.7–7.2)[28]. Comparative, multisite studies integrating ecological, behavioural and genomic data are needed, particularly in light of the recent 2024–2025 outbreaks[54–56].

Addressing these gaps will require complementary and scalable monitoring approaches. Low-cost and field-deployable diagnostics[57], community-based carcass reporting systems (for example, SISS-Geo[58]) and non-invasive monitoring tools—including improved canopy-level entomological surveillance[59], areal-imagery and acoustic tracking of NHPs[60], and the use of NHP saliva and mosquito-blood-meal analyses[48]—could together accelerate outbreak detection and response. These approaches could also refine estimates of NHP density and help assess susceptibility and exposure across multiple primate families, both at urban–forest edges in the Atlantic Forest and in undersampled biomes such as the Amazon Forest and inland Cerrado where YFV persists in an enzootic cycle[61].

Overall, our findings show that explosive yellow fever outbreaks at the forest–urban interface arise from interactions between vector ecology, host susceptibility and environmental drivers. With YFV resurging across the Americas[56], sustained multisource collaborative surveillance, rapid public health action and alignment with WHO's elimination strategy remain essential to protect at-risk human and wildlife populations. Sentinel NHP deaths occurring alongside rising *Haemagogus* abundance mark a brief window for intervention, with transmission able to intensify within only a few host–mosquito–host generations. Keeping detection-to-action intervals as short as possible will therefore be critical for future yellow fever mitigation strategies at the forest–urban interface.

Methods

Ethics approval

Only naturally deceased NHP carcasses were sampled for YFV surveillance. The surveillance protocol for dead NHPs was approved by the Ethics Committee for the use of Animals in Research, Instituto Adolfo Lutz (nos. 0135D/2012 and 020 G/2014), and includes work in protected environmental areas.

Study site

Parque Estadual Alberto Löfgren (PEAL; 186 hectares, 40% public use) in an Atlantic Forest fragment in northern São Paulo municipality (775–850 m elevation) contiguous with Cantareira State Park (PEC). It sits within São Paulo Green Belt Biosphere Reserve (UNESCO), part of the Atlantic Forest Biosphere Reserve. Public-use sectors include Horto Florestal, Olaria, Polo Ecocultural and Arboreto Vila Amália, which also contains ~256 residential dwellings and staff housing. Annual visitation exceeded 1.6 million in 2019. For spatial analyses, we grouped management sectors into three zones (A–C, Fig. 1a).

Entomological surveys

Following YFV confirmation at PEAL, we conducted targeted mosquito surveys in 43 days from 22 Dec 2017 to 02 Oct 2018 (40 epidemiological weeks) at 39 points near NHP carcasses. Using personal protective equipment, teams applied a standard protocol with: (1) hand nets and (2) Nasci vacuum aspirators at ground level and (3) CDC light traps baited with CO_2 (dry ice) at ground and canopy levels. Traps operated from 09:00 to 15:00 and aspirator collections were performed for a fixed duration per point within the same period. To capture mosquito–host interface zones, forest edges, peridomestic areas and forest trails were sampled.

Captured mosquitoes were frozen alive in liquid nitrogen and stored at −70 °C in labelled cryotubes. For each collection we recorded method, stratum (ground/canopy), coordinates, date, number of tubes and collection period. Specimens were identified on a cold table with standard taxonomic keys[62–64]. Females from the same place/date were pooled (1–10) by species (or genus if needed). Raw, georeferenced records are provided in Supplementary Table 1. Temporal and spatial sampling layouts, genus composition and relative abundance are shown in Supplementary Fig. 1. Abundance summaries stratified by collection method and height are shown in Extended Data Fig. 1.

All statistical analyses, data wrangling and visualization were performed in R (v.4.3.2)[65] using tidyverse, ggplot2, lubridate and related packages unless otherwise specified. Additional packages used for specialized tasks are cited at first mention.

Meteorological data for São Paulo municipality

Hourly total precipitation (tp), 2-m air temperature (T) and 2-m dewpoint temperature (d) for 1 Jan 2017–31 Dec 2018 were retrieved from the ERA5-Land reanalysis dataset provided by Copernicus Climate

Change Service (C3S) (0.1° resolution)[66]. ERA5 temperatures (Kelvin) were converted to °C and precipitation (m) to mm. Relative humidity (rh) was computed using the Magnus approximation (parameters $a = 17.625$, $b = 243.04$; temperatures in °C):

$$\mathrm{rh} = 100 \times \exp(d \times a)/(d + b) \exp(T \times a)/(T + b) \quad (1)$$

ERA5-Land was overlaid on LandScan annual population rasters (1/120° resolution[67,68]) using geobr (v.1.8)[69] municipality boundaries (majority-area tule when borders intersected a cell). Each LandScan cell inherited the nearest ERA5-Land grid-point values (nearest-neighbour mapping), yielding a per-cell table with administration unit, population and hourly meteorological variables for each pixel. Population-weighted hourly series were computed for São Paulo municipality and summarized into daily and then monthly indices (temperature minima/means/maxima; relative humidity minima/means/maxima; precipitation sum of daily totals) (Supplementary Table 4). These indices approximate meteorological conditions experienced across São Paulo municipality (including PEAL) during the study period.

Association between vector abundance and climate factors

Monthly counts of *H. leucocelaenus* were modelled with negative-binomial GLMs. Covariates were *z* standardized, and predictors included contemporaneous and 1-month-lagged temperature (min/mean/max), relative humidity (min/mean/max) and cumulative rainfall. Unless specified, all hypotheses were two-sided and interpreted at alpha = 0.05. Model selection used the Akaike information criterion (AIC). Results are reported as rate ratios (RR) with two-sided 95% confidence intervals (CIs) and pseudo-R^2. Collinearity (particularly between temperature and rainfall) (Supplementary Fig. 2) was assessed and, where indicated, addressed via interaction/quadratic terms. Models were fitted using MASS::glm.b, and correlations were summarized with corrplot[70] and regression tables generated with stargazer[71] (Supplementary Tables 5 and 6). An identical framework was applied to the four most frequent genera at PEAL (*Haemagogus*, *Aedes*, *Limatus* and *Culex*) (Supplementary Table 7). Coefficients are reported as RR per 1 °C and per unit rainfall, with 95% CIs and pseudo-R^2.

NHP carcass collection and notification

On Oct 9, 2017, the first confirmed YFV case at PEAL was detected in a southern brown howler monkey (*A. guariba*). During the ensuing outbreak, carcasses were collected under routine wildlife surveillance and samples were collected for virological testing. For each carcass, we recorded detection date, coordinates and decomposition stage (early-stage ≤48 h, medium/advanced >48 h post mortem) based on external morphology and soft-tissue preservation; no carcasses in desiccated/remains stages were collected (≥144 h post mortem) (Supplementary Table 10). Notifications were filed by the Wild Animal Management Center (CeMaCas) and reported to the São Paulo State Health Secretariat via Information System for Notifiable Diseases (SINAN), following Brazilian Ministry of Health guidelines[24].

Nucleic acid isolation and YFV RT–qPCR

Specimens were tested for YFV by RT–qPCR in accordance with the YFV Brazilian National Surveillance Program[24]. For NHPs, liver fragments were homogenized with magnetic beads (Magna Lyser, Roche), and RNA was extracted from 700 µl of homogenate (QIAamp RNA Blood mini kit, QIAGEN) following manufacturer instructions. Mosquito pools were homogenized in 700 µl phosphate-buffered saline (PBS, 0.75% bovine albumin, penicillin 100 U ml^{-1} and streptomycin 100 µg ml^{-1}) using the MagNA Lyser (Roche) and RNA extracted (QIAamp Viral RNA mini kit, QIAGEN). Viral detection was conducted using the GoTaq 1-Step RT–qPCR (Promega) at Instituto Adolfo Lutz Virology Centre as previously described[72], with standard RT–qPCR cycling conditions: reverse transcription (45 °C for 10 min), enzyme activation (95 °C for 10 min), followed by 40 cycles (95 °C for 15 s and 60 °C for 45 s for annealing/extension) on an ABI7500 Real-Time PCR (Thermo Fisher).

Species-specific C_t analysis across mosquito species

To compare viral RNA abundances across mosquito species, we analysed C_t values from YFV-positive pools of non-blood-fed females, combining 9 observations from this study with 109 observations from three Brazilian datasets[5,73,74]. Species names were harmonized (grouping *H. janthinomys* and *H. capricornii* as females are morphologically indistinguishable[75]); assay protocols[72,76,77], date, municipality and pool size were recorded. We compared *H. leucocelaenus* against other species within each assay using two-sided Mann–Whitney/Wilcoxon tests with Benjamini–Hochberg FDR adjustment across pairs. To account for assay and pool size, we fitted a linear model with $C_t \approx$ pecies + assay + pool size (Extended Data Fig. 2). Statistical tests were conducted in R and results tidied with broom[78]. The source data (species, assay, pool size, pool-level C_t, references) are provided as Supplementary Table 2.

H. leucocelaenus YFV natural infection rates

We estimated mosquito infection rates using (1) minimum infection rate (MIR; per 1,000) and (2) a pooled-binomial maximum-likelihood approach accounting for variable pool sizes[79]. To contextualize PEAL, we compiled Brazilian studies testing ≥10 pools per species and extracted MIRs for *H. leucocelaenus* ($n = 7$), *H. janthinomys* ($n = 5$) and 'other species' ($n = 4$, including *Sabethes soperi, Sabethes chloropterus, Aedes scapularis* and *Aedes taeniorhynchus*)[5,32,80] (Extended Data Fig. 3). Species-level MIR distributions were compared using Kruskal–Wallis followed by Benjamini–Hochberg FDR-adjusted two-sided pairwise Wilcoxon rank-sum tests. The source data (species, municipality, number of pools, total mosquitoes and pool sizes (when available), MIR, dates and references) are provided as Supplementary Table 3.

Viral metagenomics

A schematic of laboratory and phylogenetic analyses is shown in Supplementary Fig. 4. YFV-positive NHP tissues and mosquito pools were sequenced with a validated SMART-9N metagenomic sequencing protocol[9,10,18]. Briefly, 44 µl of extracted viral RNA were DNase treated (TURBO DNase, Thermo Fisher), cleaned up and concentrated to 10 µl (RNA Clean & Concentrator-5, Zymo). Extraction blanks and no-template controls were included at each step to monitor contamination. MinION libraries were prepared from 50 ng of double-tagged cDNA per sample, barcoded and pooled equimolarly using the EXP-NBD104 (1–12) and EXP-NBD114 (13–24) Native Barcoding kits (Oxford Nanopore Technologies (ONT)), and sequenced with the SQK-LSK109 sequencing kit on fresh FLO-MIN106 (R.9.4.1) flowcells (ONT). Three 24-plex libraries were run on a GridION (ONT) under MinKNOW 1.15.1 (ONT) using a standard 48-h script.

Tiled-amplicon sequencing and cross-method benchmarking

A subset of 56 YFV-positive NHP and mosquito samples was sequenced with a validated multiplex tiled-amplicon protocol (YFV500/V1)[81] designed during the early stages of the outbreak (April 2017)[9]. Overlapping 500 bp amplicons spanning the coding region of the YFV South American genotype I outbreak clade were generated and sequenced on GridION (ONT). We compared C_t values with paired differences in consensus sequence coverage between tiled-amplicon vs viral metagenomics, fitting a linear model (stats v.3.6.2 package) to calculate the C_t at which the methods perform equivalently (Supplementary Fig. 3). Source data can be found in Supplementary Table 8. Coverage profiles and recurrent amplicon dropouts for the YFV500/V1 multiplex primer scheme are shown in Extended Data Fig. 4.

Consensus genome generation

FASTQ files were demultiplexed and adapter trimmed (Guppy v.5.0.16, ONT), mapped to YFV South American genotype I (BeH655417; GenBank

accession number JF912190) with minimap2 (v.2.28)[82], then processed with SAMtools (v.1.20)[83]. Raw read counts and fragment N_{50} were summarized with NanoStat (v.1.1.2)[84]. Genome visualization, mapped-read counts and depth profiles were inspected with Tablet (v.1.17.08.17)[85] and SAMtools (v.1.20)[83]. Variants were called using Medaka v.1.12.1 (ONT) and consensus sequences generated with margin_cons Medaka v.1.12.1 (ONT); regions with <10× depth were masked. Sequencing metrics, C_t values and metadata are provided in Supplementary Tables 8 and 9.

Host and vector identification and detection of other viral pathogens from metagenomic sequencing data

Despite DNAse treatment to remove background DNA in the SMART-9N protocol, sufficient host and vector reads remained to confirm species. Demultiplexed FASTQs were mapped with minimap2 (v.2.28)[82] to *H. leucocelaenus* (GenBank accession number NC057212.1) and *A. guariba* (GenBank accession number NC_064186.1). Alignments were processed with SAMtools (v.1.20)[83] and consensus fragments were generated as described above. Species identities were verified using BLASTn v.216.0 against NCBI GenBank[86]. To screen for non-YFV viruses, reads were classified with Kraken2 (v.2.1.3)[87] (RefSeq complete viral genomes[88]), inspected in pavian[89] and validated by reference mapping (NCBI Viral Genome Resource[90]). We considered a detection as corroborated when mapping showed coherent coverage across the expected genomic region (breadth and depth consistent with library yield) and the signal was absent from concurrent negative controls.

HAV consensus and phylogeny

Reads from NP2754 were mapped to GenBank accession number MG049743.1 and a consensus was generated using the same pipeline (positions with depth <10× masked). Genotype assignment used the Hepatitis A virus genotyping tool (v.1.0)[91]. For context, we retrieved the closest 100 sequences to NP2754 HAV via BLASTn (v.216.0)[86] and NCBI Virus (Brazilian sequences >500 bp, Oct 2025)[92]. After excluding outliers with very long branch lengths (AF268396.1, MG181943.1, MZ557007.1) and removing records without country information, sequences were aligned with MAFFT v.7 (–auto) and end trimmed to the longest region shared with the NP2754 HAV consensus. After additional QC with TempEst (v.1.5.3)[93], a 97-sequence alignment (6,638 bp, sampling range = 68 years) was used to estimate a maximum-likelihood tree using IQ-TREE2 (v.2.3.66)[94] with ModelFinder Plus[95] and node support from UFBoot2 (1,000) and SH-aLRT (1,000) replicates[96]. The final tree was midpoint rooted and is shown in Extended Data Fig. 5.

YFV nucleotide data collation, curation and ML phylogenies

Public YFV sequences were retrieved from GenBank and filtered to include Brazil-origin samples with >70% genome coverage (based on submitter metadata or computed coverage) and exclude laboratory strains, chimaeras and records lacking a verifiable collection date/location. For samples represented by multiple assemblies, we kept the highest coverage or earliest submission and removed duplicates. PEAL sequences from this study were added after the same QC. Coding regions were aligned with MAFFT v.7 (–auto)[97] and manually curated in AliView (v.1.28)[98]. Sequences with >30% ambiguous sites or obvious frame disruptions were excluded.

From the curated alignment (10,221 bp) we defined three datasets: (1) Brazil (2008–2023; n = 1,063, 13 states, including PEAL; Supplementary Fig. 5); (2) São Paulo (a subset focused on the previously described YFV_{SP} clade[25], including sequences from Goiás, Minas Gerais, São Paulo and Rio de Janeiro collected between 2015–2019; n = 450, including PEAL; Supplementary Fig. 6); and (3) PEAL (n = 88 near-complete and complete PEAL genomes from groups I–III; Fig. 3a). Per-sequence QC metrics and accession numbers of the PEAL dataset are provided in Supplementary Table 9. We inferred ML phylogenies using IQ-TREE (v.2.2.2.6)[94] with the ModelFinder Plus model selection[95] and 1,000 ultrafast bootstrap replicates[96]. Clock-like signal was assessed with TempEst (v.1.5.3)[93]. For the PEAL dataset, we additionally applied Bayesian Evaluation of Temporal Signal (BETS)[99], computing four independent analyses: strict vs relaxed molecular clock (with log-normally distributed rate variation among branches[100]), and exponential vs Bayesian skygrid tree priors[101]. Including sampling dates in BETS provided strong support for the presence of temporal signal in the PEAL dataset (log Bayes factors > 18) when including sampling dates vs the null model (no sampling dates) in all four clock and tree prior model combinations. Moreover, the relaxed clock with a skygrid tree prior outperformed the relaxed clock with an exponential tree prior (log Bayes factors = 15.5). Alignments, trees and XMLs are provided in Data availability.

Bayesian phylogenetic inference

Time-scaled phylogenies for the São Paulo and PEAL datasets were inferred using BEAST X[102] with BEAGLE v.3 acceleration[103] under an autocorrelated relaxed clock (log-normal among-branch rate variation[100]), Bayesian skygrid demographic model (using 47 grid points and cut-off of 4 years for the São Paulo dataset, and 35 grid points and a cut-off of 3 years for the PEAL dataset)[101], HKY substitution with among-site heterogeneity[104,105] and a CTMC reference prior on the clock rate[106]. Tip dates were set at the reported precision (day, month or year). Two independent MCMC chains (100 million steps, sampling every 10,000 steps) were run per dataset. Convergence and mixing of the MCMC chains were assessed in Tracer (v.1.7.2)[107] (effective sample size > 200 for all key parameters and visual agreement across duplicate runs). After 10–25% burn-in, runs were combined (LogCombiner v.1.10.5) and summarized as maximum credibility (MCC) trees (TreeAnnotator (v.1.10.5)[108]). Time-dependence of evolutionary rates in YFV South American genotype 1 was explored by comparing posterior rates inferred from the PEAL and São Paulo datasets with published data[9,109–112] (Supplementary Fig. 7).

Epizootic cluster definition

On PEAL phylogenies, a cluster was defined as a monophyletic group of ≥2 PEAL genomes (NHP or mosquito) with strong support (ML bootstrap >90 and/or posterior probability >0.90). We verified monophyly in the São Paulo and Brazil-wide ML trees; all remained strongly supported (bootstrap >95) except Cluster D (2 *H. leucocelaenus* sequences), which we retained on the basis of support within the PEAL analysis (bootstrap = 81). Given the observation that YFV evolutionary rates can be time dependent (Supplementary Fig. 7), cluster dating used the single-season PEAL dataset (root-to-tip r = 0.39). For each cluster we recorded the (1) median date of the cluster most recent common ancestor (MRCA); (2) median date of the parent of the MRCA; (3) midpoint between (1) and (2); and (4) cluster duration, defined as the number of days from the last sampling date to (3) (Fig. 4b).

Modelling YFV transmission dynamics

Model framework. We simulated the PEAL epizootic in *A. guariba* using an IBM implemented in R (package 'individual' (v.1.1.17)[113]). Each NHP transitions through susceptible (S), exposed (E), infectious (I) and dead (D) states, assuming 100% lethality in *A. guariba* at PEAL (approximating observed lethality; Fig. 1b). Upon death, individuals are detected and sampled with sampling probability p_{obs}, reflecting incomplete ascertainment of all deceased *A. guariba*. A summary of the model parameterization and sources for the parameters can be found in Extended Data Table 1.

The infection process and the force of infection (FOI). We assumed density-dependent transmission, such that each susceptible *A. guariba* experiences a force of infection (FOI) proportional to the number of infectious individuals at time t, $I(t)$. Thus:

$$\lambda(t) = \beta I(t) \quad (2)$$

where β is the transmission rate. If γ is the average duration of infectiousness, the basic reproduction number R_0 is:

$$R_0 = \beta S(t)\gamma \quad (3)$$

representing the total number of infected *A. guariba* that a single infected *A. guariba* infects during the course of their infection, in a population the size of the susceptible *A. guariba* population in PEAL. Because the model does not explicitly represent mosquito populations, R_0 reflects the combined host–vector–host process (that is, the product of the number of vectors infected by an infectious *A. guariba* and the number of *A. guariba* subsequently infected by infectious vectors).

IIP in *A. guariba*. Following infection, individuals enter an exposed (E) state before becoming infectious. The time spent in the E state is drawn from a gamma distribution and reflects both the IIP of YFV in NHPs which was estimated from ref. 114 (Fig. 4c,d) and the EIP in the vector (see below). In the case of the incubation period of YFV in *A. guariba*, a gamma distribution was fitted to this data using the Stan (v.2.36.0) probabilistic programming language[115]. Weakly informative gamma priors were placed on both the shape (a) and scale (b) parameters of the gamma distribution for the IIP, as follows: $a \approx$ Gamma(1, 0.5); $b \approx$ Gamma(1, 2). The model was run with 4 chains of 2,000 iterations each, of which the first 1,000 iterations were used as warm up for adaptation, resulting in a total of 4,000 posterior samples. Convergence was confirmed by standard diagnostics (R-hat < 1.02).

EIP in *H. leucocelaenus*. We reanalysed the experimental data of ref. 116 (Fig. 4a,b) to estimate the EIP for *H. leucocelaenus*. Briefly, *Haemagogus* mosquitoes were inoculated with YFV and then made to feed on mice some number of days following inoculation. The proportion of mice that go on to die following *Haemagogus* feeding is therefore indicative of the proportion of mosquitoes that have completed the EIP and are thus able to successfully transmit YFV upon feeding. A gamma distribution was fitted to the data presented in the paper for the 25 °C experiment (representing the most complete experiment that is nearest to the temperatures measured in PEAL as shown in Fig. 1c). Uniform, minimally informative priors were placed on both the shape (a) and scale (b) parameters of the gamma distribution, with each specified as uniform(0.1, 10). We excluded the results associated with days 0 and 1, which are associated with high mortality and probably represent infection occurring from residual viable YFV from injection into the vector, rather than new infectious YFV virions generated by replicative cycles inside the vector. As above, the model was run with 4 chains of 2,000 iterations each, of which the first 1,000 iterations were used as warm up for adaptation, resulting in a total of 4,000 posterior samples, and convergence was confirmed across chains (R-hat < 1.01). Although vectors were not explicitly modelled, this empirically derived EIP was included in the delays governing host–vector–host transmission, ensuring that expected delays between an *A. guariba* becoming infected and (indirectly, through infectious *Haemagogus* vectors) generating subsequent infections in other NHPs were accurately captured. Following this intrinsic and extrinsic incubation period, the infected *A. guariba* were then presumed to begin contributing to onward transmission (that is, the FOI term described above).

NHP infectious period, time to death and observation process. We assumed that YFV is 100% lethal in *A. guariba* and thus, once infectious, an individual NHP eventually goes on to die due to YFV. In our IBM, this delay between infection and death was drawn from a gamma distribution, the parameters of which were estimated from reanalysis of experimental *A. guariba* infection data presented in ref. 114. As previously, gamma distributions were fitted to this data using the Stan (v.2.36.0) probabilistic programming language. Weakly informative gamma priors were placed on both the shape (a) and scale (b) parameters of the gamma distribution, as follows: $a \approx$ Gamma(1, 0.5); $b \approx$ Gamma(1, 2). The model was run with 4 chains of 2,000 iterations each, of which the first 1,000 iterations were used as warm up for adaptation, resulting in a total of 4,000 posterior samples. Convergence of the chains was confirmed (R-hat < 1.02).

Upon death, each NHP was reported with probability p_{obs}, which represents the probability that a dead NHP was observed, and a notification sent that enabled enumeration and sampling to occur. In practice, we implemented this by drawing a Bernoulli random variable at the time of death for each individual, with parameter p_{obs} representing the chance the death was recorded:

$$D_{reported} \sim \text{Binomial}(D_{obs}, p_{obs}) \quad (4)$$

Given confirmation of local extinction of the NHP population of PEAL (see also Fig. 1b), this probability was estimated empirically from the collated data for PEAL, specifically by dividing the number of enumerated *A. guariba* deaths over the modelled time period by the total *A. guariba* population of PEAL immediately preceding the modelled time period. Detection and reporting of deaths occurred with some delay following death; this delay was estimated empirically from local veterinarian team estimates of the stage of corpse decomposition at detection (that is, the number of days since the NHP had died, see also Supplementary Table 8), and its estimation is described in further detail below. Specifically, a mixture of exponential and gamma distributions was fitted to the data using the Stan (v.2.36.0) probabilistic programming language[115]. Weakly informative gamma priors were placed on both the shape (a) and scale (b) parameters of the gamma distribution as follows: $a \approx$ Gamma(1, 1); $b \approx$ Gamma(2, 2). For the exponential distribution, we placed a uniform(0.1, 10) prior on the rate parameter. As above, the model was run with 4 chains of 2,000 iterations each, of which the first 1,000 iterations were used as warm up for adaptation, resulting in a total of 4,000 posterior samples. Convergence of the chains was confirmed (R-hat < 1.02). See Data availability for full details of the IBM and its implementation. The empirical and estimated distributions for the delay between *A. guariba* death and that death being reported and enumerated are shown in Fig. 2c.

Inference of R_0 and fitting to PEAL timeseries data. To infer the basic reproduction number (R_0) of YFV-infected *A. guariba* in PEAL during the outbreak, we fitted the IBM to the timeseries of NHP YFV deaths in PEAL. We placed an informative prior on the timing and intensity of importations based on the dated phylogenetic analysis of YFV PEAL dataset (see Fig. 3), constraining possible start dates using the posterior distribution of the parental node of Cluster A (sensitivity analyses using the midpoint of Clade A's parental node and the MRCA of Clade A as the basis for the prior on the start data are shown in Supplementary Fig. 8). Because the first detected case (NP2067) was phylogenetically unrelated to subsequent transmission (a scenario that was consistent in all three phylogenetic datasets) and to the subsequent sustained outbreak that led to local extinction, we excluded deaths before 15 Nov 2017 (Fig. 1b) from model fitting, treating this early event as a separate, self-limited introduction.

Daily incidence of reported NHP deaths $y(t)$ were modelled as the Poisson realizations of the expected number of detected deaths, $D_{reported}(t)$, reported on day t:

$$Y(t) \sim \text{Poisson}(D_{reported}(t)) \quad (5)$$

The likelihood for the entire timeseries is therefore given by:

$$\text{Poisson}(Y(t)|_{D\text{reported}}(t)) \quad (6)$$

A closed-form expression of the likelihood of the observed data given the model and its parameters was not analytically tractable, so we used particle filtering methods to obtain unbiased likelihood estimates[117]. To

generate an estimate of the marginal model likelihood for each parameter combination, we conducted a parameter scan across different parameter combinations, utilizing a bootstrap particle filter with 250 particles. If the expected values of count distributions derived from the modelling framework are equal to 0 when empirically observed data are not 0, this results in particles of 0 weight, which can lead to the particle filter estimating the marginal likelihood to be 0 for all particles, and prevent the bootstrap particle filter from sampling efficiently or appropriately. To mitigate this, we followed the approach of ref. 118 and added a small but non-zero weight for each particle at every observation. This was achieved by adding a small amount of noise (exponentially distributed with mean 10^{-2}) to modelled incident death values of 0. For each parameter combination, we ran 10 independent replicates with a bootstrap particle filter (to generate 10 independent estimates of the model likelihood for that parameter combination) and calculated the mean likelihood from these independent replicates. Model fitting was carried out within a Bayesian framework, with weakly informative priors used for the R_0 and epidemic start date. For the prior on R_0, we compiled 11 field-based estimates of R_0 reported in a recent review[26]. A truncated normal was fitted to the data reported in this review to construct a prior distribution for R_0 (mean ≈ 4.6, s.d. ≈ 2.7). The prior on the epidemic start date was constructed on the basis of the PEAL molecular clock analysis. Specifically, the start date of the outbreak was bounded by Clade A's MRCA posterior distribution, fitting a Weibull distribution to this posterior and using it as the prior in the transmission model inferential framework.

Transmission generations and epidemic amplification. To interpret cluster sizes in terms of underlying transmission dynamics, we simulated a negative-binomial branching process (offspring mean R, dispersion k) with combined generation time G = EIP (*H. leucocelaenus*) + EIP (*A. guariba*). This allowed us to summarize the number of host–mosquito–host generations required to reach each observed cluster size. For the dominant Cluster A, the inferred size was consistent with 1–3 generations. Full details of the inferential framework and its implementation are available in Data availability.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The YFV genomic data generated in this study are available in NCBI GenBank under accession numbers OQ714241–OQ714328. The near-complete hepatatis A virus (HAV) genome is available under accession number PV702359. Raw metagenomic sequencing reads (FASTQ files) have been deposited under the NCBI Sequence Read Archive (SRA) BioProject PRJNA1269522, with associated BioSample accessions SAMN48792130–SAMN48792218. Detailed laboratory protocols, including the viral metagenomic sequencing workflow and SMART-9N primer sequences, are available at https://protocols.io ref. 46. The phylogenetic trees, genomic datasets and XML files are publicly accessible in our dedicated GitHub repository at https://github.com/CADDE-CENTRE/YFV_horto (ref. 119).

Code availability

The IBM YFV model is publicly accessible via GitHub at https://github.com/cadde-centre/YFV_horto (ref. 119), along with the scripts for the statistical analyses reported in this study. Tools, packages and software used in this study are publicly available.

References

1. Johansson, M. A., Vasconcelos, P. F. C. & Staples, J. E. The whole iceberg: estimating the incidence of yellow fever virus infection from the number of severe cases. *Trans. R. Soc. Trop. Med. Hyg.* **108**, 482–487 (2014).
2. Mutebi, J.-P. & Barrett, A. D. T. The epidemiology of yellow fever in Africa. *Microbes Infect.* **4**, 1459–1468 (2002).
3. Cavalcante, K. R. L. J., Tauil, P. L., Cavalcante, K. R. L. J. & Tauil, P. L. Risco de reintrodução da febre amarela urbana no Brasil. *Epidemiol. Serv. Saúde* **26**, 617–620 (2017).
4. Soper, F. L. et al. Yellow fever without *Aëdes aegypti*. Study of a rural epidemic in the Valle Do Chanaan, Espirito Santo, Brazil, 1932. *Am. J. Epidemiol.* **18**, 555–587 (1933).
5. Abreu, F. V. S. D. et al. *Haemagogus leucocelaenus* and *Haemagogus janthinomys* are the primary vectors in the major yellow fever outbreak in Brazil, 2016–2018. *Emerg. Microbes Infect.* **8**, 218–231 (2019).
6. Vasconcelos, P. F. C. et al. Genetic divergence and dispersal of yellow fever virus, Brazil. *Emerg. Infect. Dis.* **10**, 1578–1584 (2004).
7. Bernardi, V. et al. Yellow fever virus in mosquitoes from rainforest bordering Manaus, Brazil, 2022. *Emerg. Infect. Dis.* **31**, 851–854 (2025).
8. *Boletim Epidemiológico – Situação Epidemiológica Da Febre Amarela No Monitoramento 2019/2020* Vol. 51 https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/edicoes/2020/boletim-epidemiologico-vol-51-no-01 (Secretaria de Vigilância em Saúde, Ministério da Saúde Brasil, 2020).
9. Faria, N. R. et al. Genomic and epidemiological monitoring of yellow fever virus transmission potential. *Science* **361**, 894–899 (2018).
10. Hill, S. C. et al. Genomic surveillance of yellow fever virus epizootic in São Paulo, Brazil, 2016–2018. *PLoS Pathog.* **16**, e1008699 (2020).
11. Lacerda, A. B., del Castillo Saad, L., Ikefuti, P. V., Pinter, A. & Chiaravalloti-Neto, F. Diffusion of sylvatic yellow fever in the state of São Paulo, Brazil. *Sci. Rep.* **11**, 16277 (2021).
12. Sacchetto, L. et al. Neighbor danger: yellow fever virus epizootics in urban and urban-rural transition areas of Minas Gerais state, during 2017–2018 yellow fever outbreaks in Brazil. *PLoS Negl. Trop. Dis.* **14**, e0008658 (2020).
13. Couto-Lima, D. et al. Potential risk of re-emergence of urban transmission of yellow fever virus in Brazil facilitated by competent *Aedes* populations. *Sci. Rep.* **7**, 4848 (2017).
14. Oklander, L. I. et al. Restoration of *Alouatta guariba* populations: building a binational management strategy for the conservation of the endangered brown howler monkey of the Atlantic Forest. *Front. Conserv. Sci.* **5**, 1401749 (2024).
15. De Abreu, F. V. S. et al. Capture of *Alouatta guariba clamitans* for the surveillance of sylvatic yellow fever and zoonotic malaria: Which is the best strategy in the tropical Atlantic Forest? *Am. J. Primatol.* **81**, e23000 (2019).
16. Hamlet, A. et al. Seasonality of agricultural exposure as an important predictor of seasonal yellow fever spillover in Brazil. *Nat. Commun.* **12**, 3647 (2021).
17. Lloyd-Smith, J. O. et al. Epidemic dynamics at the human–animal interface. *Science* **326**, 1362–1367 (2009).
18. Claro, I. M. et al. Rapid viral metagenomics using SMART-9N amplification and nanopore sequencing. *Wellcome Open Res.* **6**, 241 (2023).
19. Culot, L. et al. ATLANTIC - PRIMATES: a dataset of communities and occurrences of primates in the Atlantic Forests of South America. *Ecology* **100**, e02525 (2019).
20. Almeida-Silva, B., Cunha, A. A., Boubli, J. P., Mendes, S. L. & Strier, K. B. Population density and vertical stratification of four primate species at the Estação Biológica de Caratinga/RPPN-FMA, Minas Gerais, Brazil. *Neotrop. Primates* **13**, 25–29 (2005).
21. Jung, L., Mourthe, I., Grelle, C. E. V., Strier, K. B. & Boubli, J. P. Effects of local habitat variation on the behavioral ecology of two sympatric groups of brown howler monkey (*Alouatta clamitans*). *PLoS ONE* **10**, e0129789 (2015).

22. Mello, I., Alvarenga de Oliveira, C. & Sobral, G. Temporal variation and age influence activity budget more than sex and reproductive status in wild brown howler monkeys (*Alouatta guariba clamitans*) inhabiting a large, continuous forest. *Primates* **65**, 411–419 (2024).
23. Svoboda, W. K. et al. Serological detection of hepatitis a virus in free-ranging neotropical primates (*Sapajus* spp., *Alouatta caraya*) from the Paraná River Basin, Brazil. *Rev. Inst. Med. Trop. S. Paulo* **58**, 9 (2016).
24. *Guia de Vigilância de Epizootias Em Primatas Não Humanos e Entomologia Aplicada à Vigilância Da Febre Amarela* (Ministério da Saúde do Brasil, Secretaria de Vigilância em Saúde, Departamento de Vigilância das Doenças Transmissíveis, 2017).
25. Hill, S. C. et al. Climate and land-use shape the spread of zoonotic yellow fever virus. Preprint at *medRxiv* https://doi.org/10.1101/2022.08.25.22278983 (2022).
26. Liu, Y. & Rocklöv, J. What is the reproductive number of yellow fever? *J. Travel Med.* **27**, taaa156 (2020).
27. Fraser, K. et al. Assessing yellow fever outbreak potential and implications for vaccine strategy. *PLOS Glob. Public Health* **4**, e0003781 (2024).
28. Ferreira, F. C. D. S. L., Bastos Camacho, L. A. & Villela, D. A. M. Occurrence of yellow fever outbreaks in a partially vaccinated population: an analysis of the effective reproduction number. *PLoS Negl. Trop. Dis.* **16**, e0010741 (2022).
29. Fraser, C. et al. Pandemic potential of a strain of influenza A (H1N1): early findings. *Science* **324**, 1557–1561 (2009).
30. Hamlet, A., Gaythorpe, K. A. M., Garske, T. & Ferguson, N. M. Seasonal and inter-annual drivers of yellow fever transmission in South America. *PLoS Negl. Trop. Dis.* **15**, e0008974 (2021).
31. Dias, R., de Mello, C. F., Santos, G. S., Carbajal-de-la-Fuente, A. L. & Alencar, J. Vertical distribution of oviposition and temporal segregation of arbovirus vector mosquitoes (Diptera: Culicidae) in a fragment of the Atlantic Forest, State of Rio de Janeiro, Brazil. *Trop. Med. Infect. Dis.* **8**, 256 (2023).
32. Vasconcelos, P. F. C. et al. An epidemic of sylvatic yellow fever in the southeast region of Maranhao State, Brazil, 1993–1994: epidemiologic and entomologic findings. *Am. J. Trop. Med. Hyg.* **57**, 132–137 (1997).
33. de Azevedo Fernandes, N. C. C. et al. Differential yellow fever susceptibility in New World nonhuman primates, comparison with humans, and implications for surveillance. *Emerg. Infect. Dis.* **27**, 47–56 (2021).
34. Waddell, M. B. Comparative efficacy of certain South American *Aëdes* and *Haemagogus* mosquitoes as laboratory vectors of yellow fever. *Am. J. Trop. Med. Hyg.* **29**, 567–575 (1949).
35. Quick, J. & Lansdowne, L. ARTIC SARS-CoV-2 sequencing protocol v4 (LSK114) v2. *protocol.io* https://doi.org/10.17504/protocols.io.bp2l6n26rgqe/v4 (2024).
36. Weaver, S. C., Forrester, N. L., Liu, J. & Vasilakis, N. Population bottlenecks and founder effects: implications for mosquito-borne arboviral emergence. *Nat. Rev. Microbiol.* **19**, 184–195 (2021).
37. Goes de Jesus, J. et al. Yellow fever transmission in non-human primates, Bahia, Northeastern Brazil. *PLoS Negl. Trop. Dis.* **14**, e0008405 (2020).
38. Oliosi, E. et al. Yellow fever in two unvaccinated French tourists to Brazil, January and March, 2018. *Eurosurveillance* **23**, 1800240 (2018).
39. Gubler D. J. *Aedes aegypti* and *Aedes aegypti*-borne disease control in the 1990s: top down or bottom up. *Am. J. Trop. Med. Hyg.* **40**, 571–578 (1989).
40. Bamou, R. et al. Pathogens spread by high-altitude windborne mosquitoes. *Proc. Natl. Acad. Sci. USA* https://doi.org/10.1073/pnas.2513739122 (2025).
41. Causey, O. R., Kumm, H. W. & Laemmert, H. W. Dispersion of forest mosquitoes in Brazil: further studies. *Am. J. Trop. Med.* **30**, 301–312 (1950).
42. Dégallier, N. et al. Release–recapture experiments with canopy mosquitoes in the genera *Haemagogus* and *Sabeihes* (Diptera: Culicidae) in Brazilian Amazonia. *J. Med. Entomol.* **35**, 931–936 (1998).
43. Davis, N. N. et al. The Global Wind Atlas: a high-resolution dataset of climatologies and associated web-based application. *Bull. Am. Meteorol. Soc.* **104**, E1507–E1525 (2023).
44. Huestis, D. L. et al. Windborne long-distance migration of malaria mosquitoes in the Sahel. *Nature* **574**, 404–408 (2019).
45. Caleiro, G. S. et al. Molecular epidemiology of St. Louis encephalitis virus, São Paulo State, Brazil, 2016–2018. *Emerg. Infect. Dis.* **31**, 1052–1054 (2025).
46. Lorenz, C., Freitas Ribeiro, A. & Chiaravalloti-Neto, F. Mayaro virus distribution in South America. *Acta Trop.* **198**, 105093 (2019).
47. Forato, J. et al. Molecular epidemiology of Mayaro virus among febrile patients, Roraima State, Brazil, 2018–2021. *Emerg. Infect. Dis.* **30**, 1013–1016 (2024).
48. Mirza, J. D. et al. Tracking arboviruses, their transmission vectors and potential hosts by nanopore sequencing of mosquitoes. *Microb. Genom.* **10**, 001184 (2024).
49. Garcia-Oliveira, G. et al. YELLOW ALERT: persistent yellow fever virus circulation among non-human primates in urban areas of Minas Gerais State, Brazil (2021–2023). *Viruses* **16**, 31 (2023).
50. Buss, G. et al. *Southern Brown Howler Monkey Alouatta guariba ssp. clamitans* The IUCN Red List of Threatened Species 2021: e.T39918A190419216. https://doi.org/10.2305/IUCN.UK.2021-1.RLTS.T39918A190419216.en (IUCN, 2016).
51. Dietz, J. M. et al. Yellow fever in Brazil threatens successful recovery of endangered golden lion tamarins. *Sci. Rep.* **9**, 12926 (2019).
52. Massad, E., Miguel, M. M. & Coutinho, F. A. B. Is vaccinating monkeys against yellow fever the ultimate solution for the Brazilian recurrent epizootics? *Epidemiol. Infect.* **146**, 1622–1624 (2018).
53. Pastusiak, A. et al. A metagenomic analysis of the phase 2 *Anopheles gambiae* 1000 genomes dataset reveals a wide diversity of cobionts associated with field collected mosquitoes. *Commun. Biol.* **7**, 667 (2024).
54. Cunha, M. S. et al. Yellow fever virus resurgence in São Paulo State, Brazil, 2024–2025. *Rev. Inst. Med. Trop. Sao Paulo* **68**, e13 (2026).
55. Fernandes, N. C. C. et al. Phylogenetic analysis reveals a new introduction of Yellow Fever virus in São Paulo State, Brazil, 2023. *Acta Trop.* **251**, 107110 (2024).
56. *Epidemiological Alert – Yellow Fever in the Americas Region – 26 March 2025* (PAHO, 2025).
57. Cavuto, M. L. et al. Portable molecular diagnostic platform for rapid point-of-care detection of mpox and other diseases. *Nat. Commun.* **16**, 2875 (2025).
58. Chame, M. et al. SISS-Geo: leveraging citizen science to monitor wildlife health risks in Brazil. *J. Healthc. Inform. Res.* **3**, 414–440 (2019).
59. Le Breton, C. et al. Advancing canopy-level entomological surveillance to monitor vector-borne and zoonotic disease dynamics. *Trends Parasitol.* **41**, 150–161 (2025).
60. Lawson, J. et al. Automated acoustic detection of Geoffroy's spider monkey highlights tipping points of human disturbance. *Proc. R. Soc. B.* **290**, 20222473 (2023).
61. Possas, C. et al. Yellow fever outbreak in Brazil: the puzzle of rapid viral spread and challenges for immunisation. *Mem. Inst. Oswaldo Cruz* **113**, e180278 (2018).

62. Forattini, O. P. *Medical Culicidology* Vol 2. Identification, Biology, Epidemiology (Sao Paulo, 2002).
63. Consoli, R. A. G. B. & de Oliveira, R. L. *Principais Mosquitos de Importância Sanitária no Brasil* (Fundação Oswaldo Cruz, 1994).
64. Lane, J. *Neotropical Culicidae* Vols I and II (Univ. Sao Paulo, 1953).
65. R Core Team *R: A Language and Environment for Statistical Computing* (R Core Team, 2023).
66. Sabater, J. M. ERA5-land hourly data from 1950 to present. *Copernicus Climate Change Service (C3S) Climate Data Store (CDS)* https://doi.org/10.24381/cds.e2161bac (2019).
67. Rose, A., McKee, J., Urban, M. & Bright, E. LandScan Global 2017. *Oak Ridge National Laboratory* https://doi.org/10.48690/1524212 (2018).
68. Rose, A., McKee, J., Urban, M. & Bright, E. LandScan Global 2018. *Oak Ridge National Laboratory* https://doi.org/10.48690/1524213 (2019).
69. Pereira, R. H. & Goncalves, C. N. geobr: download official spatial data sets of Brazil. R package version 1.9.1. *GitHub* https://ipeagit.github.io/geobr/ (2024).
70. Wei, T. & Simko, V. R Package corrplot: visualization of a correlation matrix (v.0.95). *GitHub* https://github.com/taiyun/corrplot (2024).
71. Hlavac, M. stargazer: well-formatted regression and summary statistics tables. R package version 5.2.2. *CRAN* http://CRAN.R-project.org/package=stargazer (2022).
72. Domingo, C. et al. Advanced yellow fever virus genome detection in point-of-care facilities and reference laboratories. *J. Clin. Microbiol.* **50**, 4054–4060 (2012).
73. de Oliveira, C. H. et al. Yellow fever virus maintained by *Sabethes* mosquitoes during the dry season in Cerrado, a semiarid region of Brazil, in 2021. *Viruses* **15**, 757 (2023).
74. Caleiro, G. S. et al. Yellow fever virus (YFV) detection in different species of culicids collected during an outbreak in southeastern Brazil, 2016–2019. *Trop. Med. Infect. Dis.* **10**, 118 (2025).
75. Telles-de-Deus, J. et al. COI DNA barcoding to differentiate *Haemagogus janthinomys* and *Haemagogus capricornii* (Diptera: Culicidae) mosquitoes. *Acta Trop.* **259**, 107377 (2024).
76. Moreira-Soto, A. et al. Evidence for multiple sylvatic transmission cycles during the 2016–2017 yellow fever virus outbreak, Brazil. *Clin. Microbiol. Infect.* **24**, 1019.e1–1019.e4 (2018).
77. Bonaldo, M. C. et al. Genome analysis of yellow fever virus of the ongoing outbreak in Brazil reveals polymorphisms. *Mem. Inst. Oswaldo Cruz* **112**, 447–451 (2017).
78. Robinson, D., Hayes, A. & Couch, S. broom: convert statistical analysis objects into tidy tibbles. R package version 1.0.10. *GitHub* https://github.com/tidymodels/broom/blob/v1.0.10/R/broom-package.R (2025).
79. Williams, C. J. & Moffitt, C. M. A critique of methods of sampling and reporting pathogens in populations of fish. *J. Aquat. Anim. Health* **13**, 300–309 (2001).
80. Pinheiro, G. G., Rocha, M. N., de Oliveira, M. A., Moreira, L. A. & Andrade Filho, J. D. Detection of yellow fever virus in sylvatic mosquitoes during disease outbreaks of 2017–2018 in Minas Gerais State, Brazil. *Insects* **10**, 136 (2019).
81. Quick, J. et al. Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. *Nat. Protoc.* **12**, 1261–1276 (2017).
82. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
83. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
84. De Coster, W., D'Hert, S., Schultz, D. T., Cruts, M. & Van Broeckhoven, C. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666–2669 (2018).
85. Milne, I. et al. Using Tablet for visual exploration of second-generation sequencing data. *Brief. Bioinform.* **14**, 193–202 (2013).
86. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinformatics*. **10**, 421 (2009).
87. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
88. Goldfarb, T. et al. NCBI RefSeq: reference sequence standards through 25 years of curation and annotation. *Nucleic Acids Res.* **53**, D243–D257 (2025).
89. Breitwieser, F. P. & Salzberg, S. L. Pavian: interactive analysis of metagenomics data for microbiome studies and pathogen identification. *Bioinformatics* **36**, 1303–1304 (2020).
90. Brister, J. R., Ako-adjei, D., Bao, Y. & Blinkova, O. NCBI viral genomes resource. *Nucleic Acids Res.* **43**, D571–D577 (2015).
91. Vennema, H., Kroneman, A. & Koopmans, M. *Hepatitis A Virus Automated Genotyping Tool* (RIVM, 2024).
92. *NCBI Virus* https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/ (NCBI, 2024).
93. Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016).
94. Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
95. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A. & Jermiin, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
96. Minh, B. Q., Nguyen, M. A. T. & von Haeseler, A. Ultrafast approximation for phylogenetic bootstrap. *Mol. Biol. Evol.* **30**, 1188–1195 (2013).
97. Katoh, K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
98. Larsson, A. AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* **30**, 3276–3278 (2014).
99. Duchene, S. et al. Bayesian evaluation of temporal signal in measurably evolving populations. *Mol. Biol. Evol.* **37**, 3363–3379 (2020).
100. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).
101. Gill, M. S. et al. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol. Biol. Evol.* **30**, 713–724 (2013).
102. Baele, G. et al. BEAST X for Bayesian phylogenetic, phylogeographic and phylodynamic inference. *Nat. Methods* **22**, 1653–1656 (2025).
103. Ayres, D. L. et al. BEAGLE 3: improved performance, scaling, and usability for a high-performance computing library for statistical phylogenetics. *Syst. Biol.* **68**, 1052–1061 (2019).
104. Hasegawa, M., Kishino, H. & Yano, T. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174 (1985).
105. Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* **39**, 306–314 (1994).
106. Ferreira, M. A. R. & Suchard, M. A. Bayesian analysis of elapsed times in continuous-time Markov chains. *Can. J. Stat.* **36**, 355–368 (2008).
107. Rambaut, A. Tracer v1.6 http://beast.bio.ed.ac.uk/Tracer (2014).
108. Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018).
109. Bryant, J. E., Holmes, E. C. & Barrett, A. D. T. Out of Africa: a molecular perspective on the introduction of yellow fever virus into the Americas. *PLoS Pathog.* **3**, e75 (2007).

110. Gómez, M. M. et al. Genomic and structural features of the yellow fever virus from the 2016–2017 Brazilian outbreak. *J. Gen. Virol.* **99**, 536–548 (2018).
111. Delatorre, E. et al. Distinct YFV lineages co-circulated in the central-western and southeastern Brazilian regions from 2015 to 2018. *Front. Microbiol.* **10**, 1079 (2019).
112. Mir, D. et al. Phylodynamics of yellow fever virus in the Americas: new insights into the origin of the 2017 Brazilian outbreak. *Sci. Rep.* **7**, 7385 (2017).
113. Charles, G., Wu, S. L., Winskill, P. & Lietar, P. individual: an R package for individual-based epidemiological models. *J. Open Source Softw.* **6**, 3539 (2021).
114. Laemmert, H. W. & Kumm, H. W. The susceptibility of howler monkeys to yellow fever virus. *Am. J. Trop. Med.* **30**, 723–731 (1950).
115. Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1–32 (2017).
116. Roca-García, M. & Bates, M. The development of the virus of yellow fever in *Haemagogus* mosquitoes. *Am. J. Trop. Med. Hyg.* **26**, 585–605 (1946).
117. Andrieu, C., Doucet, A. & Holenstein, R. Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. B Stat. Methodol.* **72**, 269–342 (2010).
118. Knock, E. S. et al. Key epidemiological drivers and impact of interventions in the 2020 SARS-CoV-2 epidemic in England. *Sci. Transl. Med.* **13**, eabg4262 (2021).
119. Faria, N. R. & Whittaker, C. YFV_horto: analysis code for 'Evolution and spillover dynamics of yellow fever at the forest–urban interface' (Version 1.0.0) [Computer software]. *GitHub* https://github.com/cadde-centre/YFV_horto (2026).
120. *Collection Use and Land Coverage* (MapBiomas, 2023); https://brasil.mapbiomas.org/codigos-de-legenda/
121. Muñoz-Sabater, J. et al. ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth Syst. Sci. Data* **13**, 4349–4383 (2021).

Acknowledgements

We thank the São Paulo State Secretariat for the Environment (2017–2018), including M. Brusadin, L. A. Bucci (Director of PEAL) and P. R. da Silva, for logistical support at PEAL; the Environmental Military Police of São Paulo State and the Metropolitan Civil Guard for operational assistance, particularly during the CADDE Workshop on Viral Metagenomics for Public Health (March 2023) held at PEAL and at the Institute of Tropical Medicine (USP), where preliminary findings were presented; R. Cardoso de Paula (Municipal Health Department of São Paulo) and R. Spinola and H. Sato (State Health Department of São Paulo) for leadership in YFV surveillance and vaccination; D. Fonseca Jr., A. Nepomuceno Duarte (SUCEN, São Paulo State Health Department) and the SUCEN field team for their essential contributions to entomological collections, and P. R. Urbinatti and R. M. Marques Sá de Almeida (Entomology Laboratory of the School of Public Health - USP) for support with taxonomic identification. We acknowledge the use of ERA5-Land meteorological data (Muñoz Sabater, 2019) accessed via the Copernicus Climate Change Service Climate Data Store. The results incorporate modified Copernicus Climate Change Service information (2023); neither the European Commission nor ECMWF is responsible for any use of this information. This work was supported by the Wellcome Trust Digital Technology Development Award (226075/Z/22/Z) (W.M.d.S., N.R.F.); the UK Medical Research Council (MRC) and FAPESP (MRC MR/S0195/1; FAPESP 18/14389-0) (I.M.C., M.B., E.C.S., N.R.F.); the International Pathogen Surveillance Network Catalytic Grant Fund (RT-MeTA, FRMM, ICL, N.R.F.); and the Wellcome Trust Dengue and Zika Immunology and Genomics Multi-Country Network (DeZi Network) (316633/Z/24/Z) (N.R.F., E.C.S.). We also acknowledge support from the MRC Centre for Global Infectious Disease Analysis (MR/X020258/1), funded by the UK MRC (N.R.F.); this UK-funded award is delivered within the Global Health EDCTP3 Joint Undertaking.

Author contributions

J.T.-d.-D., I.M.C., M.B., C.W., M.P.-C., A.P., L.F.M. and N.R.F. conceived the study. J.T.-d.-D., I.M.C., M.B., E.C.R., T.M.C., C.A.M.d.S., I.N.V., J.G.d.J., P.d.S.A., A.C.F., A.V.d.P., R.H.d.A.E., M.d.S.L., F.M.d.O., V.R.d.S., L.A.M.F., M.S.N., T.C.S., E.T.B.C.d.S., A.A.C.C., P.R.d.S., T.N.L.-C., M.B.d.P., M.S.C., J.L.S., A.P.P.T., E.S.B., M.P., D.A.G.O., H.M.P., T.G.-L., C.M.G.A., G.C.d.R., K.L.d.G. and F.E.S.P.V. performed fieldwork, sample collection, laboratory work and sequencing. C.W., V.C., N.C.C.F.d.A., J.M.G., F.R.R.M., S.C.H., L.A., W.M.d.S., B.L., R.P.d.S., A.P., E.C.S., L.F.M. and N.R.F. performed data curation and phylogenetic, epidemiological and statistical analyses. C.W. developed the IBM model. J.T.-d.-D., I.M.C., M.B., C.W., L.F.M. and N.R.F. interpreted results. N.R.F., L.F.M., J.T.-d.-D., I.M.C., M.B. and C.W. drafted the manuscript with input from all authors. N.R.F., L.F.M. and E.C.S. supervised the project. All authors reviewed and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41564-026-02302-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41564-026-02302-w.

Correspondence and requests for materials should be addressed to Juliana Telles-de-Deus, Luis Filipe Mucci or Nuno R. Faria.

Peer review information *Nature Microbiology* thanks Katherine Laiton-Donato, Georgina Limon and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Juliana Telles-de-Deus[1,23] ✉, Ingra M. Claro[2,3,4,5,23], Mayara Bertanhe[2,3,4,23], Charles Whittaker[4,6,23], Márcio Port-Carvalho[7,8], Esmenia C. Rocha[2,3,9], Thaís M. Coletti[2,3], Camila A. M. da Silva[2,3], Ian Nunes Valença[2,3,4], Tamara N. Lima-Camara[9], Márcia Bicudo de Paula[9], Mariana S. Cunha[3,10], Jaqueline G. de Jesus[2,3], Pâmela dos Santos Andrade[2,3,9], Victoria Cox[4], Natalia C. C. F. de Azevedo[11], Juliana M. Guerra[11], Juliana L. Summa[12], Ana Paula P. Teixeira[13], Eduardo S. Bergo[1], Mariza Pereira[1], Filipe R. R. Moreira[4,14], Alvina Clara Felix[2,3], Anderson V. de Paula[2,3], Raissa H. de Araujo Eliodoro[2,3], Marissa da Silva Lima[2,3], Franciane M. de Oliveira[2,3], Valquíria R. de Souza[2,3], Lucas A. M. Franco[2,3], Marcelo S. Nardi[12], Thais C. Sanches[12], Eric T. B. C. da Silva[12], Amanda A. C. Coimbra[12], Paulo R. dos Santos[12], Katherine Lima de Gouveia[15], Francisco E. S. P. Vilela[6], Sarah C. Hill[16], Dilmar A. G. Oliveira[17], Hélia M. Piedade[17], Thaís Guimarães-Luiz[17], Camila M. G. Abreu[17], Guilherme Casoni da Rocha[17], Leandro Abade[18], William M. de Souza[5], Ben Lambert[19,20], Renato Pereira de Souza[21], Adriano Pinter[22], Ester C. Sabino[2,3], Luis Filipe Mucci[1] ✉ & Nuno R. Faria[2,3,4] ✉

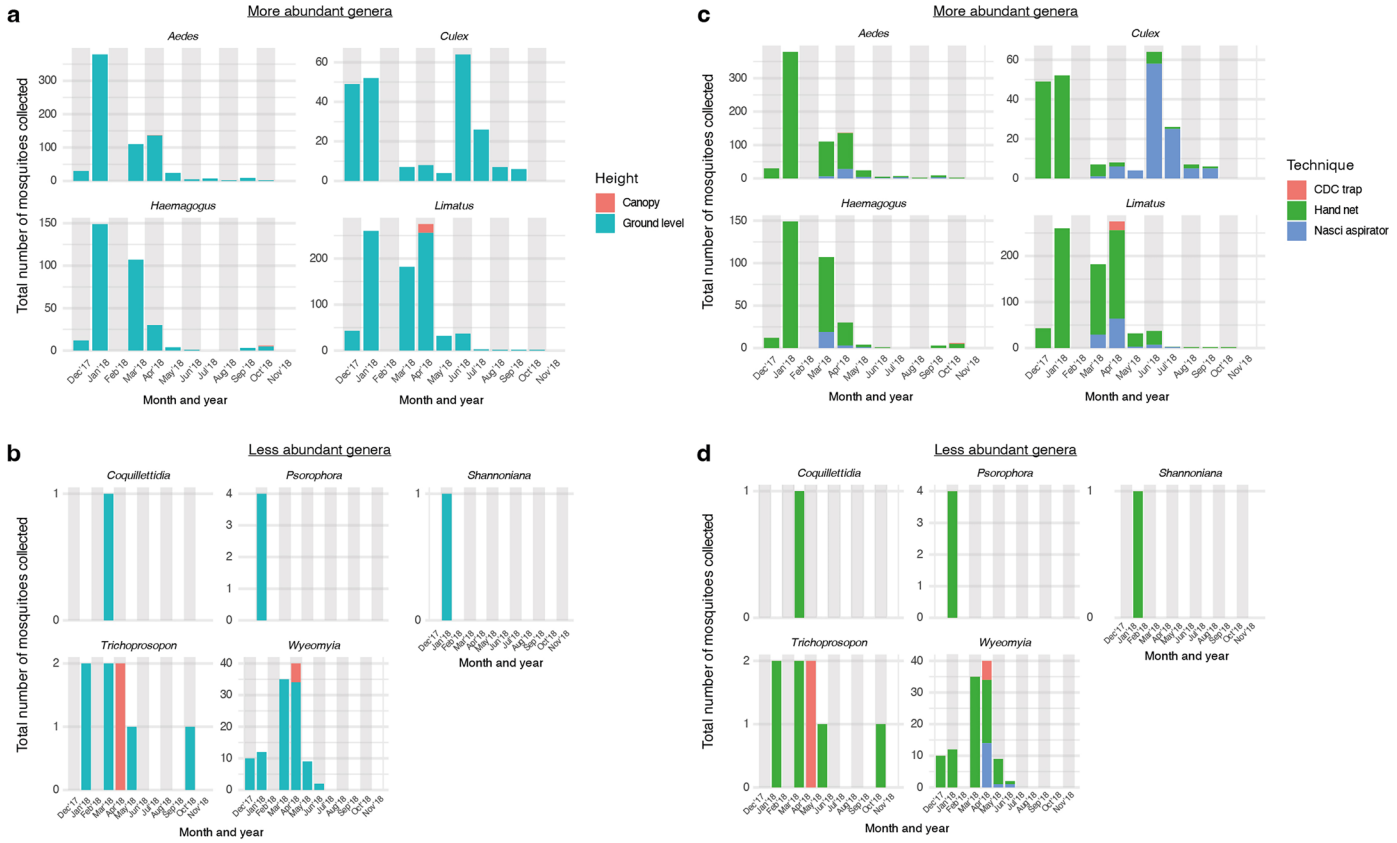
[1]Secretaria de Estado da Saúde de São Paulo, Instituto Pasteur, São Paulo, Brazil. [2]Departamento de Moléstias Infecciosas e Parasitárias, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil. [3]Instituto de Medicina Tropical, Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil. [4]MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College London, London, UK. [5]Department of Microbiology, Immunology, and Molecular Genetics, College of Medicine, University of Kentucky, Lexington, KY, USA. [6]Division of Infectious Diseases & Vaccinology, School of Public Health, University of California Berkeley, Berkeley, CA, USA. [7]Conservation Biodiversity Nucleus, Environmental Research Institute, Secretariat for the Environment, Infrastructure and Logistics of São Paulo, São Paulo, Brazil. [8]Post-Graduation Program in Biodiversity of Conservations Units, National School of Tropical Botanic, Rio de Janeiro Botanical Garden, Rio de Janeiro, Brazil. [9]Departamento de Epidemiologia, Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, Brazil. [10]Centro de Virologia, Núcleo de Doenças de Transmissão Vetorial, Instituto Adolfo Lutz, São Paulo, Brazil. [11]Núcleo de Anatomia Patológica, Instituto Adolfo Lutz, São Paulo, Brazil. [12]Divisão da Fauna Silvestre, Secretaria do Verde e Meio Ambiente da Prefeitura de São Paulo, São Paulo, Brazil. [13]São Paulo Municipal Department of Health, São Paulo, Brazil. [14]Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil. [15]Parque Estadual Alberto Löfgren, Secretaria do Meio Ambiente do Governo do Estado de São Paulo, São Paulo, São Paulo, Brazil. [16]Department of Pathobiology and Population Sciences, Royal Veterinary College, Hatfield, UK. [17]Departamento de Gestão da Fauna Silvestre, Coordenadoria de Fauna Silvestre, Secretaria de Meio Ambiente, Infraestrutura e Logística do Estado de São Paulo, São Paulo, Brazil. [18]St Mary's University, Twickenham London, Twickenham, UK. [19]Department of Statistics, University of Oxford, Oxford, UK. [20]Pandemic Sciences Institute, University of Oxford, Oxford, UK. [21]Centro de Laboratório Regional XII, Núcleo de Ciências Biológicas, Instituto Adolfo Lutz, Taubaté, São Paulo, Brazil. [22]Departamento de Medicina Veterinária Preventiva e Saúde Animal, Faculdade de Medicina Veterinária e Zootecnia, Universidade de São Paulo, São Paulo, Brazil. [23]These authors contributed equally: Juliana Telles-de-Deus, Ingra M. Claro, Mayara Bertanhe, Charles Whittaker. ✉e-mail: jtelles@pasteur.saude.sp.gov.br; lfmucci@gmail.com; nfaria@ic.ac.uk

https://doi.org/10.1038/s41564-026-02302-w

Extended Data Table 1 | Description of the IBM model parameterisation and sources for model parameters

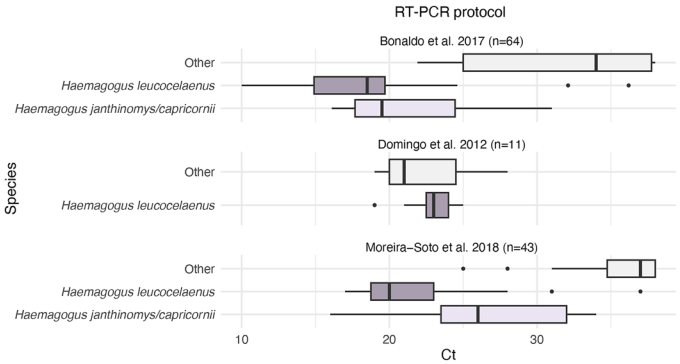
Parameter	Description	Reference
Transmission Rate (β)	Transmission rate parameter. Inferred directly from PEAL time-series data using particle filter-based likelihood estimation approach.	NA
Incubation period duration	Gamma distribution with shape X and rate Y giving average duration of Z days. Delay distribution exposed to infectious (that is the incubation period for YFV) in *A. guariba*. Estimated from analysis of data from Haemmert et al., 1950 (see Fig. 4c–e). Inverse of the expectation of this distribution is γ used in the calculation of R_0	Laemmert et al.,[15]
Duration of infectiousness	Gamma distribution with shape X and rate Y giving average duration of Z days. Delay between *Alouatta* becoming infectious and dying from YFV. Estimated from analysis of data from Haemmert et al., 1950.	Laemmert et al.,[15]
EIP	Gamma distribution with shape X and rate Y giving average duration of Z days. Extrinsic incubation period that is the time taken for infected *Haemagogus* mosquitoes to become infectious. Estimated from analysis of data from Bates & Roca-Gacia, 1946 (see Fig. 4c–e). Gamma distribution with shape X and rate Y.	Bates & Roca-Garcia[16]à
Delay between death and observation	Exponential distribution rate parameter X. The delay between *A. guariba* death occurring, and that death being detected, reported and enumerated. This was estimated empirically from veterinarian estimated stage of corpse decomposition at detection (that is the number of days since the NP had died).	Empirically estimated
p_{obs}	0.7625 (61 / 80) Probability of a dead *A. guariba* belonging to the PEAL population being detected, reported and successfully enumerated during the outbreak. Estimated empirically as the fraction of enumerated deaths relative to the overall known size of the *A. guariba* population in PEAL at the outbreak's start.	Empirically estimated

https://doi.org/10.1038/s41564-026-02302-w



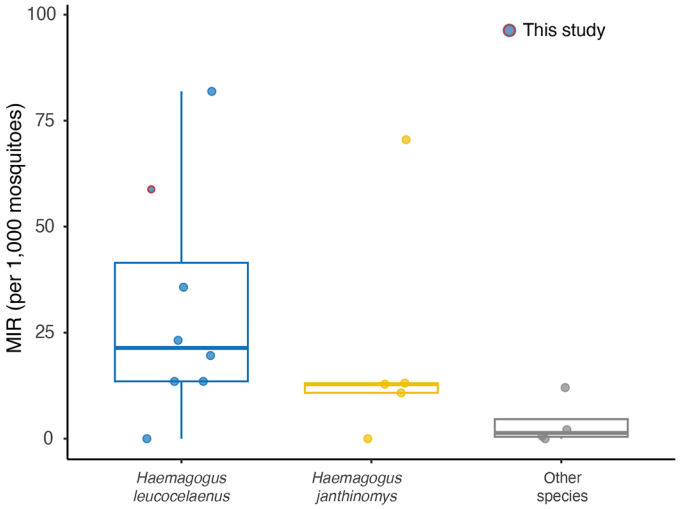
Extended Data Fig. 1 | Most and least abundant mosquito genera collected at PEAL stratified by collection technique (a, b) and collection height (c, d). Total number of mosquitoes tested (number of mosquitoes multiplied by pool size) per epidemiological week. The most abundant genera were *Limatus* (37.6%, n = 838), *Aedes* (31.6%, n = 705) – including *Ae. scapularis* (13.3%, n = 297), *Ae. albopictus* (13.1%, n = 292), and *Ae. aegypti* (1.0%, n = 23); *Haemagogus* (14.0%, n = 312), and *Culex* (10.0%, n = 223), while *Sabethes* was less common (1.4%, n = 31). Among 818 pools, 98.7% (n = 807; 2,022 specimens) were collected at ground level and 1.3% (n = 11; 29 specimens) at the canopy level. Raw data, including height of entomological collections, dates and coordinates of each location, are provided in Supplementary Table 1.

https://doi.org/10.1038/s41564-026-02302-w



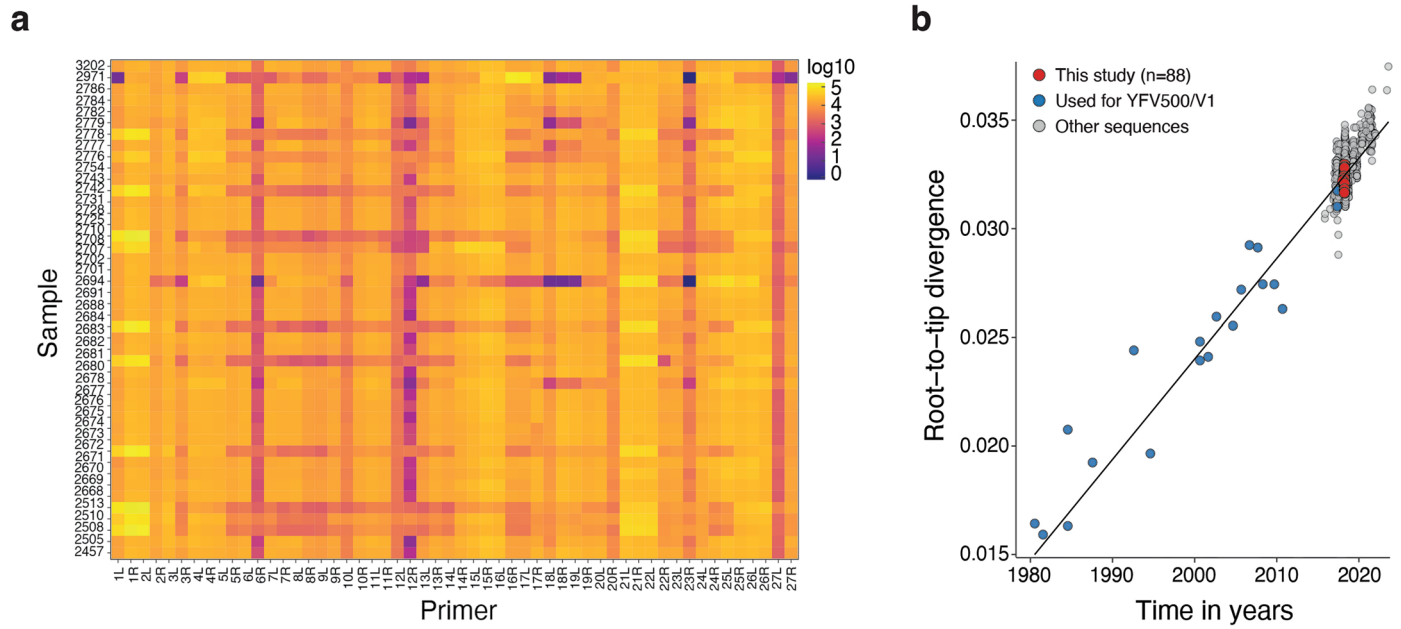
Extended Data Fig. 2 | YFV Ct values across vector species. Ct values from YFV-positive mosquito pools by species and RT-qPCR protocols[1o5–1o7]. Within-protocol comparisons indicate significantly lower Ct values for *Hg. leucocelaenus* (n = 64) compared to *Hg. janthinomys/capricornii* (n = 25) and other species (n = 29). N indicates number of observations. Pools from other species with available YFV cycle threshold values include *Ae. scapularis* (n = 7), *Ae. taeniorhynchus* (n = 1), *Ae. albopictus* (n = 3), *Ae. serratus* (n = 4), *Psorophora ferox* (n = 5), *Sabethes albiprivus* (n = 3), *Sabethes chloropterus* (n = 3), *Sabethes identicus* (n = 1), and *Sabethes purpureus* (n = 2). Raw data, including mosquito species, date of collection, mean pool Ct, study reference, and protocol reference are provided in Supplementary Table 2. Boxplots show the median (centre horizontal line), the 25th–75th percentiles (box), and whiskers extending to 1.5× the interquartile range, with points beyond this plotted as outliers.

https://doi.org/10.1038/s41564-026-02302-w



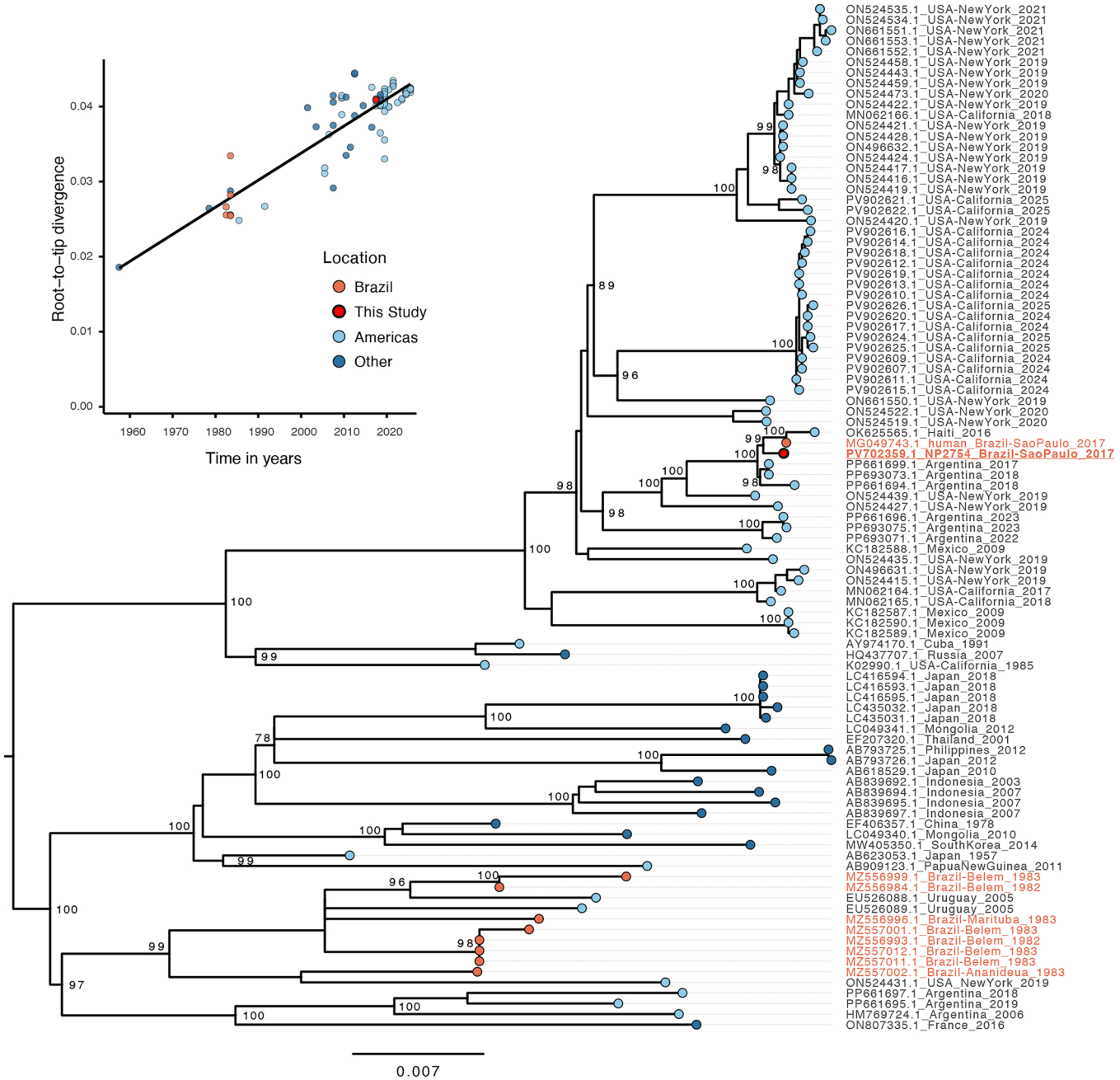
Extended Data Fig. 3 | YFV minimum infection rates across vector species. Minimum infection rates (MIR) for YFV vectors in Brazil. Circles show species-specific MIRs obtained from studies in four states (Espírito Santo, Maranhão, Minas Gerais and Rio de Janeiro), obtained from published data[17,38,39]. Only MIR values obtained from testing ≥10 pools per species were included (*Hg. leucocelaenus*: n = 8, *Hg. janthinomys*: n = 5, "Other": n = 4). Boxplots summarise the distribution for each taxon (median, IQR; whiskers at 1.5× IQR). The MIR for *Hg. leucocelaenus* at PEAL is highlighted with a red-outlined marker. The 'Other' category groups additional sylvatic species with ≥10 pools but too few observations for a separate panel (*Aedes scapularis*, *Ae. taeniorhynchus*, *Sabethes soperi* and *Sa. chloropterus)*. MIRs were computed as (number of YFV-positive pools divided by the total number of adult mosquitoes of the same species) ×1,000, following standard methodology[17,39,110]. Data is available in Supplementary Table 3.

https://doi.org/10.1038/s41564-026-02302-w



Extended Data Fig. 4 | Performance of the YFV500/V1 tiled-amplicon scheme and temporal signal in South American genotype I genomes. (**a**) Primer-binding heatmap across NP samples (n = 42). Each cell shows the number of reads containing a primer (or its reverse complement), normalised by library size (CPM, counts-per-million). Primers are ordered by amplicon number and orientation (L = left; R = right). Lower CPM indicates weaker binding/possible mismatches. (**b**) Root-to-tip genetic divergence against sampling year for South American genotype I genomes ($r^2 = 0.80$); n = 1,080 genomes from NCBI GenBank up until April 2023). Red = genomes generated in this study (n = 88); blue = genomes used to design the original YFV500/V1 primer set (up to April 2017); grey = other South American genotype I genomes.

https://doi.org/10.1038/s41564-026-02302-w



Extended Data Fig. 5 | Maximum-likelihood phylogeny of hepatitis A virus (HAV) including the PEAL NP2754 genome. The red tip marks the HAV near complete genome recovered from an *Alouatta guariba* at PEAL (NP2754). Tip colours: dark blue = Brazil (non-PEAL), light blue = other sample collection in the Americas, grey = other global sequences, red = this study. NP2754 nests within the clade that includes a contemporary 2017 São Paulo strain. Numbers on nodes are bootstrap support shown for selected clades. Branch lengths are in substitutions per site units. The tree is mid-point rooted. The map location of NP2754 is shown in Fig. 1a. Inset on the upper left shows root-to-tip regression of genetic divergence against sampling date for the same dataset; points are coloured as above and the line shows the best-fit regression ($r^2 = 0.80$, n = 97, sampling range = 68 years).

nature portfolio

Corresponding author(s): Juliana Telles de Deus, Luis Filipe Mucci, Nuno R. Faria

Last updated by author(s): Feb 16, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Commercial software / platforms:
MinKNOW v1.15.1 (Oxford Nanopore Technologies): control and run execution for GridION sequencing.
Guppy v5.0.16 (Oxford Nanopore Technologies): basecalling, demultiplexing and adapter trimming for Nanopore reads.
ABI 7500 Real‑Time PCR System (Thermo Fisher Scientific): instrument software for RT‑qPCR YFV detection (standard ABI7500 environment).

Open‑source software:
minimap2 v2.28: initial mapping at data‑collection QC stage (read‑quality and barcoding checks).
SAMtools v1.20: initial alignment inspection and QC.
NanoStat v1.1.2: sequencing run summaries (read counts, N50)
Tablet (v1.17.08.17): visual inspection of raw read alignments.

Custom and protocol‑specific code:
SMART‑9N metagenomics workflow (custom scripts): laboratory protocol–linked scripts for cDNA tagging, sample preparation tracking, barcode assignment.
Custom code related to data collection is deposited in the project's code repository as listed in the manuscript's Code Availability section.

Data analysis

Commercial software:
MinKNOW v1.15.1 (ONT): generation of FAST5/FASTQ output; sequencing metadata exported for downstream analysis.
Guppy v5.0.16: high‑accuracy basecalling and demultiplexing used as inputs for analytical pipelines.

Open‑source software:

Genome processing and consensus generation
minimap2 v2.28: mapping reads to YFV and HAV references.
SAMtools v1.20: sorting, indexing, depth profiling.
NanoStat v1.1.2: long-read QC summaries.
Tablet (1.17.08.17): manual visual QC.
medaka v1.12.1: variant calling.

Taxonomic and host/vector identification:
BLAST+ (2.16.0): host/vector confirmation by BLASTn.
Kraken2 v2.1.3 with RefSeq viral genomes database: pathogen-agnostic classification.
pavian: classification visualisation.

Phylogenetics and evolutionary analysis:
MAFFT v7 (--auto): multiple sequence alignment.
IQ-TREE2 v2.3.6 with ModelFinder Plus, UFBoot2 and SH-aLRT: maximum-likelihood phylogenetics.
TempEst v1.5.3: temporal signal assessment.
BEAST X (BEAGLE v3 acceleration): Bayesian time-scaled phylogenetics; XML configurations included in repository.

Statistical modelling and visualisation (R ecosystem):
R v4.3.2 for all modelling.
tidyverse: wrangling and plotting.
lubridate: temporal handling.
MASS (glm.nb): negative-binomial GLMs.
corrplot v0.95: correlation diagnostics.
stargazer: regression table formatting.
ggplot2: statistical visualisation.
broom v1.0.10: tidy model outputs.

Transmission modelling and inference:
Stan (v2.36.0): Bayesian fitting of incubation-period and delay distributions.
individual R package (v1.1.17): simulation framework for the individual-based model.
Custom R and bash scripts for particle filtering, likelihood estimation, branching-process simulations, model visualisation and reproduction-number estimation, available under Code Availability.

Custom code:
Custom code for the IBM transmission model, particle-filtering code, R analysis pipelines, phylogenetic XMLs and plotting scripts is deposited in a dedicated repository as specified in the Code Availability section.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The YFV genomic data generated in this study are available in NCBI GenBank under accession numbers OQ714241–OQ714328. The near-complete hepatitis A virus (HAV) genome is available in NCBI GenBank under accession number PV702359. Raw metagenomic sequencing reads (FASTQ files) have been deposited in the NCBI Sequence Read Archive under accession numbers SAMN48792130–SAMN48792218 (BioProject PRJNA1269522). Detailed laboratory protocols, including the viral metagenomic sequencing workflow and SMART-9N primer sequences, are available at https://protocols.io46. The phylogenetic trees, genomic datasets, and XML files are publicly accessible in our dedicated GitHub repository (https://github.com/cadde-centre/YFV_horto).

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	N/A.
Reporting on race, ethnicity, or other socially relevant groupings	N/A.
Population characteristics	N/A.
Recruitment	N/A.
Ethics oversight	N/A.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We investigated a sylvatic yellow fever outbreak in a forest fragment in metropolitan São Paulo by integrating mosquito sampling, detection of naturally deceased howler monkeys, and metagenomic sequencing with phylogenetic and epidemiological analyses.
Research sample	Naturally deceased Alouatta guariba found in PEAL, plus a small number of background NHP cases from Cantareira (routine state surveillance). Free‑flying forest mosquitoes, mainly Haemagogus leucocelaenus, were also collected across PEAL.
Sampling strategy	All detected NHP carcasses were sampled. Mosquitoes were collected at fixed sites using hand nets, aspirators and CO_2 traps at ground and canopy levels. Sample sizes reflect natural encounter rates; no sample?size calculations were applicable.
Data collection	Field teams georeferenced NHP carcasses and collected mosquitoes using standard entomological methods. Samples were preserved and processed in accredited public‑health laboratories following standard protocols for RNA extraction and metagenomic sequencing.
Timing and spatial scale	Sampling occurred between Oct 2017 and Oct 2018 in PEAL, with 43 mosquito‑sampling days and continuous NHP monitoring. Additional contextual NHP sequences originated from neighbouring Cantareira State Park.
Data exclusions	Only samples that failed objective quality criteria (e.g., insufficient tissue, <70% genome coverage for phylogenetics) were excluded. Exclusion criteria were pre‑defined.
Reproducibility	Field and laboratory methods followed standardised protocols. All analytical code, parameters and workflows are version‑controlled and publicly available for full reproducibility.
Randomization	Not applicable. Samples reflect natural mortality and natural mosquito abundance; no individuals were allocated to treatment groups.
Blinding	Not applicable. All samples were processed using identical laboratory and analytical workflows regardless of origin or expected YFV status.

Did the study involve field work? ☒ Yes ☐ No

Field work, collection and transport

Field conditions	Field work was conducted in Parque Estadual Alberto Löfgren (PEAL), a 186‑ha Atlantic Forest fragment embedded within metropolitan São Paulo, Brazil. Sampling took place between October 2017 and October 2018 under warm, humid subtropical conditions. Mean daily temperatures ranged from 18–28 °C, with seasonal variation described using ERA5‑Land data. Mosquito surveys were performed between 09:00 and 15:00 across forest interior, forest–urban edges, and canopy strata. Terrain included dense forest cover, steep gradients, and mixed vegetation typical of the Cantareira mountain foothills.
Location	Sampling occurred mostly within PEAL State Park (23.466–23.455° S, 46.655–46.630° W; 775–850 m elevation), in the northern zone of São Paulo city. Field sites covered management Zones A–C, including forest interior, the Vila Amália settlement edge, peridomestic surroundings, and trails near the Serra da Cantareira. Non‑human primate carcasses and mosquito collections were georeferenced and spatially restricted to areas managed by the São Paulo State Secretariat for the Environment. In addition to PEAL field samples, a small number of non‑human primate background cases from the contiguous Cantareira State Park (PEC), collected independently under routine state surveillance, were included for phylogenetic context. No additional fieldwork was carried out by the authors in PEC.
Access & import/export	Field activities were carried out exclusively within Parque Estadual Alberto Löfgren (PEAL) under authorisation from the São Paulo State Secretariat for the Environment and in accordance with Brazilian Ministry of Health guidelines for wildlife surveillance. Sampling of non‑human primate carcasses followed approved protocols from the Instituto Adolfo Lutz Ethics Committee for Animal Use in Research (0135D/2012; 020G/2014). Additional non‑human primate background samples from the contiguous Cantareira State Park (PEC) were not collected by the authors, but obtained from routine state surveillance systems managed by the São Paulo State environmental and health authorities. These samples were provided directly to the reference laboratories responsible for YFV monitoring. All specimens remained in Brazil and were processed in São Paulo, Brazil (Instituto Adolfo Lutz and the University of São Paulo). No international transport or export permits were required.

Disturbance	Disturbance to wildlife and habitat was minimised by restricting all field sampling to naturally deceased non?human primates and by using low?impact mosquito?collection methods (hand nets, aspirators, CO_2?baited traps) along predefined trails and designated access points. No animals were captured, handled, immobilised or euthanised for this study. Work was conducted by trained personnel who avoided sensitive habitat, minimised noise and foot traffic, and adhered to conservation and biosafety guidelines for protected Atlantic Forest areas. Background PEC samples used for phylogenetic context were derived from routine state surveillance, and no field activities were conducted by the authors in PEC, ensuring no additional disturbance in that conservation unit.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☐ Antibodies
☐	☐ Eukaryotic cell lines
☐	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☐	☐ Clinical data
☐	☐ Dual use research of concern
☐	☐ Plants

Methods

n/a	Involved in the study
☐	☐ ChIP-seq
☐	☐ Flow cytometry
☐	☐ MRI-based neuroimaging

Antibodies

Antibodies used	N/A.
Validation	N/A.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	N/A.
Authentication	N/A.
Mycoplasma contamination	N/A.
Commonly misidentified lines (See ICLAC register)	N/A.

Palaeontology and Archaeology

Specimen provenance	N/A.
Specimen deposition	N/A.
Dating methods	N/A.

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	N/A.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	N/A. No laboratory animals were used.

Wild animals	Only naturally deceased NHP carcasses found through routine state surveillance were sampled. No animals were captured, handled or euthanised.
Reporting on sex	Sex was not consistently identifiable for many carcasses due to decomposition, and sex was not part of the study design. Analyses therefore were not sex□stratified.
Field-collected samples	Only naturally deceased NHP carcasses were sampled for YFV surveillance.
Ethics oversight	The surveillance protocol for dead NHPs was approved by the Ethics Committee for the use of Animals in Research, Instituto Adolfo Lutz, under the numbers 0135D/2012 and 020G/2014, which include work in protected environmental areas.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	N/A.
Study protocol	N/A.
Data collection	N/A.
Outcomes	N/A.

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks: N/A.

Novel plant genotypes: N/A.

Authentication: N/A.

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.: N/A.

Files in database submission: N/A.

Genome browser session
(e.g. UCSC): N/A.

Methodology

Replicates: N/A.

Sequencing depth: N/A.

Antibodies: N/A.

Peak calling parameters: N/A.

Data quality: N/A.

Software: N/A.

Flow Cytometry

Plots

Confirm that:

☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☐ All plots are contour plots with outliers or pseudocolor plots.

☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation: N/A.

Instrument: N/A.

Software: N/A.

Cell population abundance: N/A.

Gating strategy N/A.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type N/A.

Design specifications N/A.

Behavioral performance measures N/A.

Acquisition

Imaging type(s) N/A.

Field strength N/A.

Sequence & imaging parameters N/A.

Area of acquisition N/A.

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software N/A.

Normalization N/A.

Normalization template N/A.

Noise and artifact removal N/A.

Volume censoring N/A.

Statistical modeling & inference

Model type and settings *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested *Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.*

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference (See Eklund et al. 2016) *Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*

Correction *Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).*

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity *Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).*

Graph analysis *Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).*

Multivariate modeling and predictive analysis	*Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.*