

1
2 Understanding transmission risk and predicting environmental suitability for Mayaro Virus in the
3 Americas

4
5 Michael Celone¹, Sean Beeman¹, Barbara A. Han², Alexander M. Potter^{3,4,5}, David B. Pecor^{3,4,5}, Bernard Okech¹,
6 Simon Pollett^{6,7}

7
8
9 ¹ Uniformed Services University of the Health Sciences, F. Edward Hébert School of Medicine, Department of
10 Preventive Medicine & Biostatistics, Bethesda, Maryland

11 ² Cary Institute of Ecosystem Studies, Millbrook, NY, USA

12 ³ One Health Branch, Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA

13 ⁴ Walter Reed Biosystematics Unit, Smithsonian Museum Support Center, Suitland, MD 20746, USA

14 ⁵ Department of Entomology, Smithsonian Institution—National Museum of Natural History (NMNH), Washington,
15 DC 20560, USA

16 ⁶ Infectious Disease Clinical Research Program, Department of Preventive Medicine and Biostatistics, Uniformed
17 Services University of the Health Sciences, Bethesda, MD, USA

18 ⁷ Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, MD, USA

19
20 *Corresponding author
21 Michael Celone
22 Michael.celone@usuhs.edu
23 401-297-1875

24 **Abstract (250 words):**

25 **Background**

26 Mayaro virus (MAYV) is a mosquito-borne *Alphavirus* that is widespread in South America. MAYV
27 infection often presents with non-specific febrile symptoms but may progress to debilitating chronic
28 arthritis or arthralgia. Despite the pandemic threat of MAYV, its true distribution remains unknown. The
29 objective of this study was to clarify the geographic distribution of MAYV using an established risk
30 mapping framework. This consisted of generating evidence consensus scores for MAYV presence,
31 modeling the potential distribution of MAYV across the Americas, and estimating at-risk population
32 residing in areas suitable for MAYV transmission.

33 **Methods**

34 We compiled a georeferenced compendium of MAYV occurrence in humans, animals, and arthropods.
35 Based on an established evidence consensus framework, we integrated multiple information sources to
36 assess the total evidence supporting ongoing transmission of MAYV within each country in our study
37 region. We then developed high resolution maps of the disease's estimated distribution using a boosted
38 regression tree approach. Models were developed using ten climatic and environmental covariates that are
39 related to the MAYV transmission cycle. Using the output of our boosted regression tree models, we
40 estimated the total population living in regions suitable for MAYV transmission.

41 **Findings**

42 The evidence consensus scores revealed high or very high evidence of MAYV transmission in Brazil
43 (especially the states of Mato Grosso and Goiás), Venezuela, Peru, Trinidad and Tobago, Colombia,
44 Bolivia, and French Guiana. According to the boosted regression tree models, a substantial region of
45 South America is suitable for MAYV transmission, including north and central Brazil, French Guiana,
46 and Suriname. Some regions (e.g., Guyana) with low or moderate evidence of transmission were
47 identified as highly suitable for MAYV. We estimate that approximately 77 million people in the
48 Americas live in areas that may be suitable for MAYV transmission, including 43.4 million people in
49 Brazil. Our results can assist public health authorities in prioritizing high-risk areas for vector control,
50 human disease surveillance and ecological studies.

51 **Funding**

52 This work was financially supported by the Armed Forces Health Surveillance Division—Global
53 Emerging Infections Surveillance (AFHSD-GEIS) under awards P0065_22_WR and P0050_23_WR. The
54 activities undertaken at WRBU were performed in part under a Memorandum of Understanding between
55 the Walter Reed Army Institute of Research (WRAIR) and the Smithsonian Institution, with institutional
56 support provided by both organizations.

57 **Research in context**

58 *Evidence before this study*

59 We searched PubMed on January 27, 2021 using the search term “Mayaro virus”. This search yielded 274
60 results including two systematic reviews of Mayaro virus (MAYV) occurrence and one study that
61 modeled the distribution of MAYV risk. Although prior systematic reviews included some spatial
62 information, they did not attempt to georeference MAYV occurrence with a high level of spatial precision
63 and did not quantify uncertainty in geolocated records. Furthermore, the one published study that modeled
64 environmental suitability for MAYV included only a limited set of MAYV occurrence locations. No
65 additional studies have been conducted to model the risk of MAYV occurrence across space or to
66 quantify the total population which may be at risk of MAYV infection in the Americas.

67 *Added value of this study*

68 We used a comprehensive, georeferenced compendium of MAYV occurrence to model the suitability for
69 MAYV occurrence in the Americas. Our boosted regression tree model incorporated 203 MAYV
70 occurrence locations and 10 gridded environmental covariates to generate a 5 x 5 km continuous surface
71 of MAYV suitability across the Americas. Using this distribution model, we estimated the total
72 population residing in areas that are suitable for MAYV transmission. Furthermore, we developed
73 evidence consensus scores for each country in our study region that synthesized a variety of sources to
74 assess the overall evidence of MAYV transmission.

75 *Implications of all the available evidence*

76 Our study provides a contemporary estimate of MAYV distribution using a well-established disease
77 mapping framework. This information provides an evidence base that can guide disease surveillance
78 (including human cases and ecological studies) and vector control efforts in the Americas. This is
79 especially useful in regions with high MAYV suitability but little or no evidence of MAYV transmission
80 (e.g., Guyana and Suriname).

81 Introduction

82 Mayaro virus (MAYV) is a mosquito-borne *Alphavirus* that was first detected in Trinidad in
83 1954.¹ MAYV has caused periodic outbreaks throughout Latin America² and serological surveys
84 and syndromic surveillance studies suggest widespread circulation in the region.³ Some
85 researchers have hypothesized that MAYV has broader epidemic potential and raised alarm
86 about its increased geographic spread.^{4,5}

87 MAYV can cause debilitating arthralgia or arthritis that can persist for months after initial
88 infection.⁶ However, most MAYV patients present with non-specific febrile symptoms that are
89 clinically indistinguishable from other vector borne diseases such as dengue or Zika.⁷ Therefore,
90 clinical diagnosis is often difficult and accurate estimates of disease burden remain elusive. This
91 is further complicated by the many limitations of serological diagnostics including the cross-
92 reactivity of antigenically similar viruses.⁸ Supportive care remains the current standard of
93 clinical treatment for MAYV as no licensed vaccine or antiviral treatment currently exists.

94 Limited studies on MAYV ecology suggest that this virus is maintained in a sylvatic
95 transmission cycle involving arboreal mosquito vectors and non-human animal reservoirs. High
96 seroprevalence among non-human primates (NHPs)⁹ suggests they may be involved in the
97 MAYV transmission cycle, although their precise role is inconclusive. In addition, MAYV
98 antibodies have been detected in other mammals including rodents and marsupials¹⁰. Risk factors
99 including residing near forested areas¹¹ and hunting in the rainforest¹² are associated with
100 MAYV infection in humans, highlighting the importance of the sylvatic transmission cycle and
101 the potential for spillover events. However, the occurrence of MAYV in the city of Manaus has
102 also led to concerns about the involvement of *Aedes* (*Ae.*) mosquitoes in an urban transmission
103 cycle.¹³ Though studies of wild-caught mosquito populations implicated the canopy-dwelling
104 *Haemagogus* (*Hg.*) *janthinomys* mosquito as the primary vector during a major outbreak in
105 Brazil,⁹ *Ae. aegypti* and *Ae. albopictus* have demonstrated the potential to transmit MAYV in
106 laboratory settings.¹⁴

107 A recent epidemiological alert by the Pan American Health Association (PAHO) emphasized the
108 need for increased awareness of and extended surveillance for MAYV in the Americas.¹⁵ Ideally,
109 MAYV spillover and outbreak prevention in the Americas would be guided by granular maps of
110 MAYV risk enabling targeted febrile surveillance and ecological surveillance efforts and better
111 tailored risk communications to endemic populations and travelers. However, the precise areas of
112 risk from MAYV throughout the Americas remain unclear due to limited data on MAYV
113 occurrence¹⁶, underscoring a fundamental need for a more comprehensive and georeferenced
114 dataset on MAYV occurrence in the Americas.

115 Here, we provide a critical update to the current state of knowledge on MAYV transmission risk
116 across the Americas. We adopted a well-established machine-learning based disease mapping
117 approach originally developed by ecologists to model species distributions but has since been
118 successfully applied to several medically-relevant vector-borne pathogens¹⁷⁻¹⁹. These methods
119 are particularly powerful for leveraging biological and ecological information underpinning a
120 sylvatic disease system to generate biologically realistic and spatially explicit predictions when
121 epidemiological data are still sparse. Many of these models rely on machine learning techniques
122 including boosted regression trees (BRT)²⁰ to develop a multivariate relationship between

123 disease occurrence locations and relevant climatic or environmental covariates that impact
124 disease transmission.

125 In order to develop a contemporary estimate of MAYV risk in the Americas, we applied a
126 predictive mapping approach with three components: (1) scoring the total evidence supporting
127 ongoing MAYV transmission within each country (i.e., evidence consensus scores); (2)
128 modeling the likely distribution of MAYV suitability throughout the Americas with high
129 predictive accuracy; (3) estimating the total population residing in areas with a high suitability
130 for MAYV transmission. Compared to previous estimates, these updated datasets and analyses
131 suggest that MAYV poses a substantial and possibly underestimated threat to the Americas.

132 **Methods**

133 **Evidence Consensus**

134 We used a well-established framework²¹ to generate an evidence consensus score for each
135 country in the study region. This approach integrates several information sources to generate a
136 score that characterizes the evidence for disease transmission. The data sources that we
137 considered for MAYV evidence consensus scores included health organization status, date of
138 most recent human occurrence, validity of diagnostic tests, recency of outbreaks or clinical cases,
139 and recency of occurrence in non-human animals or arthropods. Scores ranged from 0 (“*No*
140 *evidence of MAYV presence*”) to 21 (“*Complete evidence of MAYV presence*”). A more
141 comprehensive description of the evidence consensus scoring process is presented in the
142 Supplementary Materials (pp 3-5).

143 **Occurrence records**

144 We collected and georeferenced MAYV occurrence records in humans, non-human animals, and
145 arthropods^{22,23}. These records were extracted from peer-reviewed and grey literature sources.
146 Occurrences with a high level of spatial precision were designated point locations, while less
147 precise records (e.g., administrative units) were designated polygon locations. We recorded the
148 coordinates of all point locations and polygon centroids as well as the uncertainty associated with
149 each record. We included occurrence records with ≤ 75 km of uncertainty in our current analysis.
150 Overall, our dataset comprised 203 MAYV occurrence records in humans, animals, and
151 arthropods. This was subsequently reduced to 170 records after we thinned the dataset to reduce
152 spatial autocorrelation. Details on the assembly of the occurrence records are available in the
153 Supplementary Materials (p 3).

154 **Description of covariates**

155 We considered 10 ecologically relevant gridded variables (i.e., raster data) for inclusion in our
156 model. These variables included various measures of topography, climate, land cover and
157 vegetation that likely influence the MAYV transmission cycle and the distribution of MAYV
158 risk throughout the region. (See Supplementary Materials for a description of the covariates and
159 rationale for their inclusion in the model.) Values for each gridded covariate were extracted at
160 each presence/pseudoabsence location and used in the modeling procedure described below.
161 Climatic variables included rainfall, nighttime land surface temperature (LST Night), and

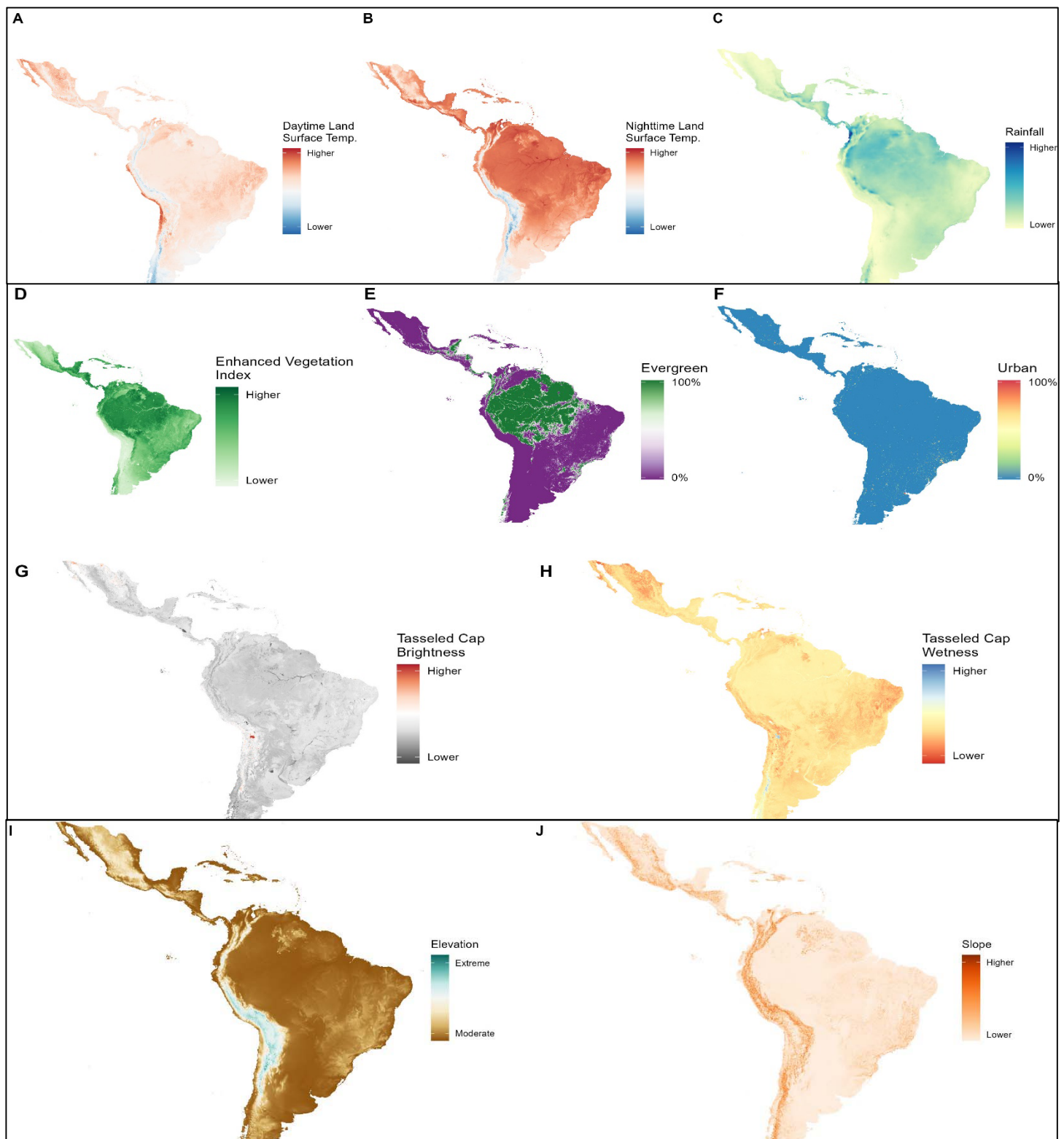
162 daytime land surface temperature (LST Day) (Figures 1A – 1C). Measures of land cover and
163 vegetation included two proportional land cover classes (evergreen broadleaf and urban/built-
164 up), enhanced vegetation index (EVI), tasseled cap wetness (TCW) and tasseled cap brightness
165 (TCB) (Figures 1D – 1H). Lastly, we included two topographic covariates: slope and elevation
166 (Figures 1I and 1J). The covariate layers are presented as maps in Figure 1 and additional details
167 on the covariates are presented in the Supplementary Materials (pp 5-6). We did not detect any
168 collinearity between the covariates based on a Spearman correlation coefficient of 0.8, and thus
169 included all 10 covariates in the final model.

170 **Mapping risk for MAYV occurrence**

171 Although first developed by ecologists to model the potential distribution of plants and animals,
172 ecological niche models (also known as species distribution models or environmental suitability
173 models) are now common in the epidemiological literature to model human disease risk. A
174 robust disease mapping framework has been established in the last decade to model the zoonotic
175 niche of medically-relevant pathogens including dengue virus (DENV)¹⁷ and Zika virus
176 (ZIKV)¹⁸, among others. We used BRT to model the environmental suitability for MAYV
177 occurrence across our study region. BRT is a machine learning approach that uses regression or
178 classification trees to partition the dataset and combines many simple models (i.e., boosting) to
179 improve overall model accuracy.²⁰ We fitted 100 BRT ‘sub-models’ to separate bootstraps of the
180 dataset using the `gbm.step` procedure in the `dismo` R package.²⁴ The predictive accuracy of the
181 model was assessed using the area under the receiver operator characteristic curve (AUC),
182 sensitivity, specificity, Kappa statistic, and percent correctly classified (PCC). Metrics were
183 calculated for each sub-model using 10-fold cross validation. The final prediction map represents
184 the mean MAYV suitability at each 5 x 5 km pixel across our ensemble of 100 models. We also
185 generated a map of the model uncertainty, represented by the per-pixel standard deviation.

186 **Total population at potential risk**

187 We estimated the total human population living in areas of high predicted MAYV suitability. We
188 first transformed the mean prediction map into a binary risk map using a previously established
189 protocol,¹⁸ whereby a suitability threshold value was chosen that encompassed 90% of the
190 MAYV occurrence points. Each 5 x 5 km pixel was classified as 1 (i.e., suitable) if its predicted
191 suitability exceeded the threshold value; otherwise, it was classified as 0 (i.e., not suitable). We
192 then determined the total population residing in suitable areas by multiplying the population
193 count within each grid cell by the binary suitability classification and summing these values
194 across each country.



35

36

37

38

Figure 1. Covariates used to model the environmental suitability of MAYV. A. and B. Land surface temperature (LST) night and LST day, respectively; C. Rainfall; D. Enhanced vegetation index (EVI); E.

09 Evergreen forest; F. Urban/built-up land cover; G. Tasseled cap brightness (TCB); H. Tasseled cap wetness
10 (TCW); I. Elevation; J. Slope

201 **Results**

202 The map of evidence consensus is presented in Figure 2A and the evidence score for each
203 country is presented by category in the Supplementary Materials (pp 11-16). Evidence consensus
204 scores ranged from 0 (no evidence of MAYV transmission) to 19 (very high evidence of
205 transmission). We recorded a very high evidence consensus score for Brazil and Venezuela, with
206 scores of 19 and 16, respectively. Other countries with a high evidence consensus score included
207 Peru with a score of 15, French Guiana with a score of 13, and Trinidad and Tobago, Colombia,
208 and Bolivia (all with scores of 11). We recorded evidence consensus scores ranging from very
209 low to moderate in all Central American and Caribbean countries. Among these countries with
210 low to moderate risk, the highest evidence consensus scores were documented for Haiti with a
211 score of 10, and Panama with a score of 9.

212 In order to provide more granular data throughout Brazil, we also calculated evidence scores by
213 state (i.e., first-administrative division). These results are presented in Figure 2B. Because health
214 organization status was not available for each state, we assigned a baseline score of one to each
215 state, and then calculated the remaining categories according to the methods described above.
216 Evidence of MAYV transmission was highest in the central Brazilian states of Mato Grosso and
217 Goiás, both with very high scores of 16. High evidence of MAYV was also documented in five
218 Northern and Central states, including Pará (Evidence Consensus= 14), Roraima (Evidence
219 Consensus= 13), Bahia and Piauí (Evidence Consensus= 12 for both), and Mato Grosso do Sul
220 (Evidence Consensus= 11).

221 Figure 3 displays the 203 MAYV occurrences (human, animal, and arthropod) that were used to
222 construct our model. The occurrence locations fell in 13 countries, most frequently in Brazil
223 (n=102), French Guiana (n=27) and Peru (n=22). MAYV occurrences were reported between the
224 years 1954 and 2021, with the majority of cases (n=140) occurring since the year 2000. One
225 hundred and sixty (79%) of the 203 occurrence locations were detected in humans while 43
226 (21%) were detected in non-human animals or arthropods. An alternative model using just the
227 reported human occurrence was also constructed using the 160 human occurrence points. This
228 model is presented in the Supplementary Materials along with a map showing the observed
229 difference between the all-host and the human-only models (pp 8-9).

230 The predicted distribution of MAYV environmental suitability is presented in Figure 4A. This
231 risk map represents the average output across the 100 BRT sub-models. The map of model
232 uncertainty (i.e., the per-pixel standard deviation across the 100 model runs) is presented in
233 Figure 4B. High suitability for MAYV transmission was evident across the Amazon rainforest
234 ecoregion in South America. The model predicted very high suitability for MAYV across a large
235 portion of Central and Northern Brazil, especially the states of Amazonas, Acre, Pará, and
236 Tocantins. High suitability was also predicted throughout French Guiana, Guyana, Suriname,
237 and Trinidad and Tobago, as well as Southern portion of Colombia and Venezuela, and the
238 north-eastern region of Peru and northern region of Bolivia.

239 Certain regions of Central America with low or moderate evidence consensus scores were also
240 found to be highly suitable for MAYV transmission, namely the Southern region of Panama and
241 eastern coastal regions of Nicaragua, Honduras, and Belize. The results of the human-only model
242 are presented in the Supplementary Figure 1. Overall, the results of this model were very similar
243 to the all-host model, with only slight differences. The difference in predicted environmental

244 suitability between the all-host model and human-only model is presented in Supplementary
245 Figure 2.

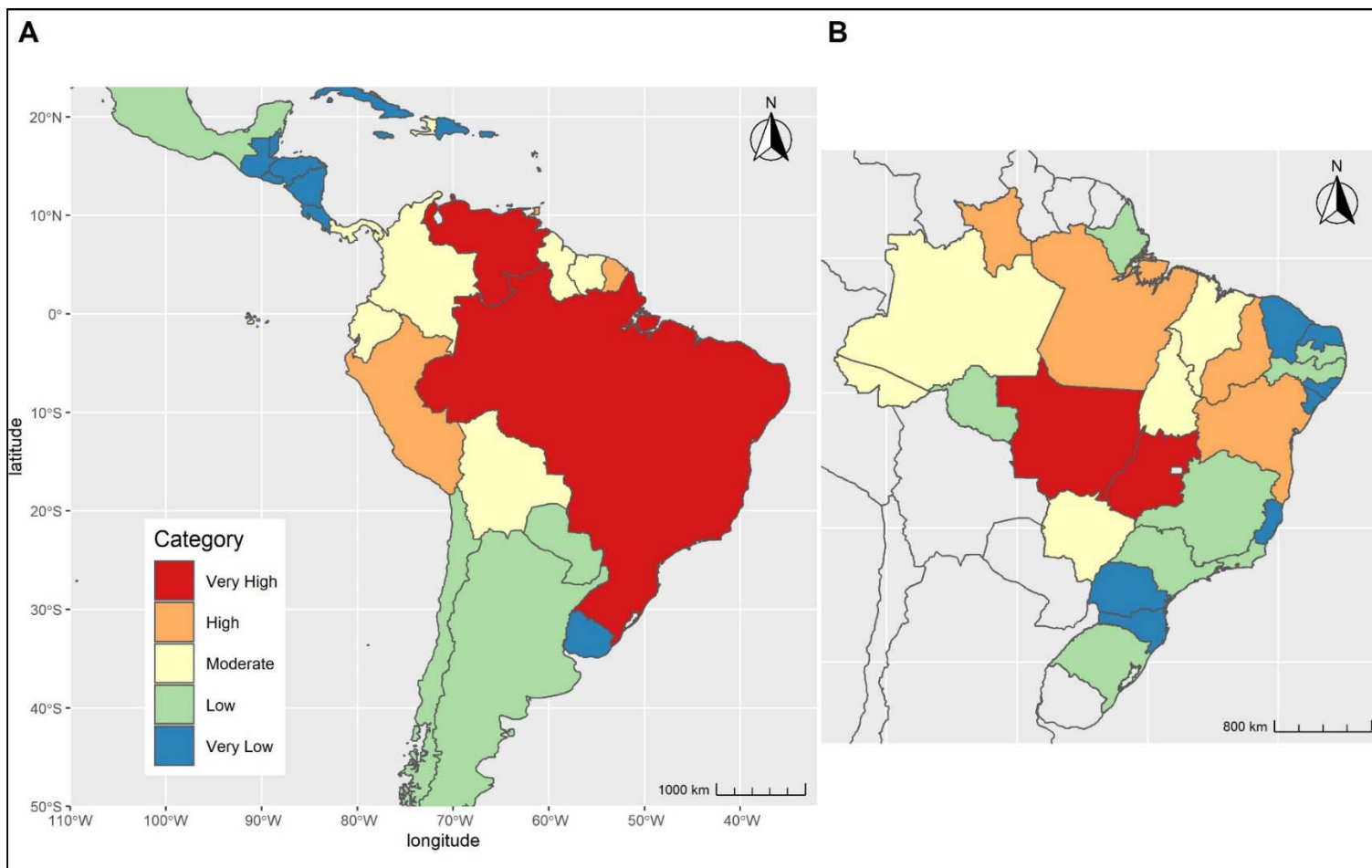
246 Our models showed MAYV suitability to be especially influenced by climatic variables
247 including nighttime LST (relative importance of 44.8%) and rainfall (relative importance of
248 23.6%) as well as EVI (relative importance of 8.1%). The partial dependence plots for nighttime
249 LST and rainfall are presented in Figure 5 and partial dependence plots for the remaining
250 variables are presented in the Supplementary Figure 3. The partial dependence plot for nighttime
251 LST reveals a steep increase in MAYV suitability around 12°C that peaks at ~22°C. The plot for
252 rainfall reveals a similarly steep increase starting at ~80mm that peaks at ~110mm and then
253 plateaus, with only a minor peak at ~375mm. After applying pairwise distance sampling to
254 remove spatial sorting bias, the model demonstrated good predictive power with an AUC of 0.78
255 ± 0.007 standard error. Other statistics from the 10-fold cross-validation procedure included the
256 following: PCC= 83%, sensitivity= 0.75, specificity= 0.91, and Kappa= 0.6.

257 We identified 0.488 as the threshold suitability value that encompassed 90% of the MAYV
258 occurrence locations. We applied this conservative value to classify pixels as either suitable or
259 unsuitable for MAYV transmission. Overall, we estimate that approximately 77 million people in
260 the Americas live in areas that are potentially suitable for MAYV transmission. Countries with
261 the greatest at-risk population include Brazil (43.4 million people), Colombia (6.9 million
262 people), and Mexico (4.9 million people). The majority of the Brazilian population living in
263 high-risk areas reside in the Amazon rainforest ecoregion.

264

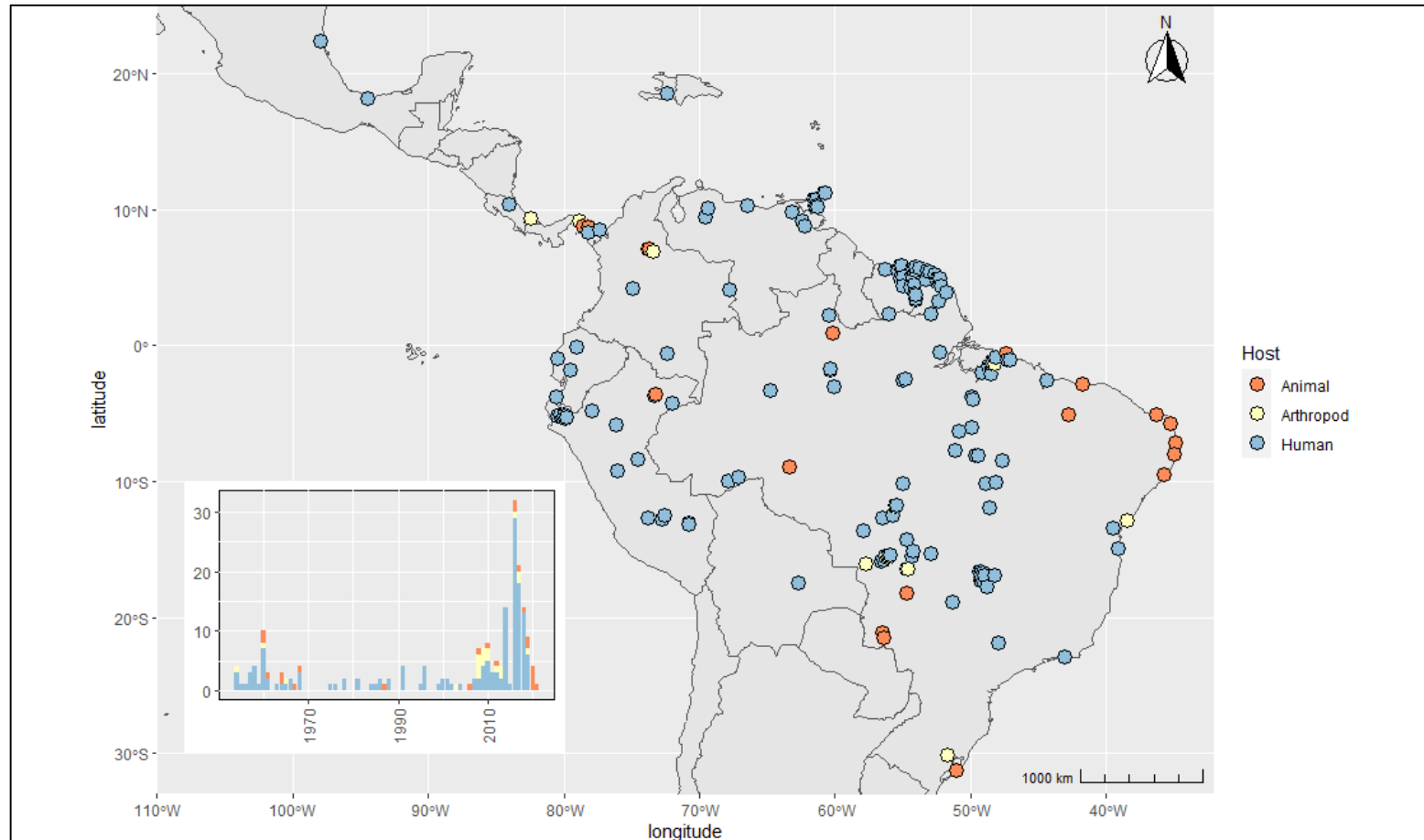
265 **Table 1. Total population living in areas potentially suitable for MAYV transmission**
266 **(millions).** Table 1 indicates the five countries in each region with the largest population in
267 suitable transmission regions. The North/Central American region includes Caribbean Island
268 nations.

Country/Region	Pop. at risk (millions)
South America	58.38
Brazil	43.37
Colombia	6.96
Venezuela	2.80
Peru	1.95
Trinidad and Tobago	0.90
North/Central America	18.33
Mexico	4.96
Haiti	3.09
Dominican Republic	2.52
Panama	2.30
Guatemala	1.21
Overall	76.71

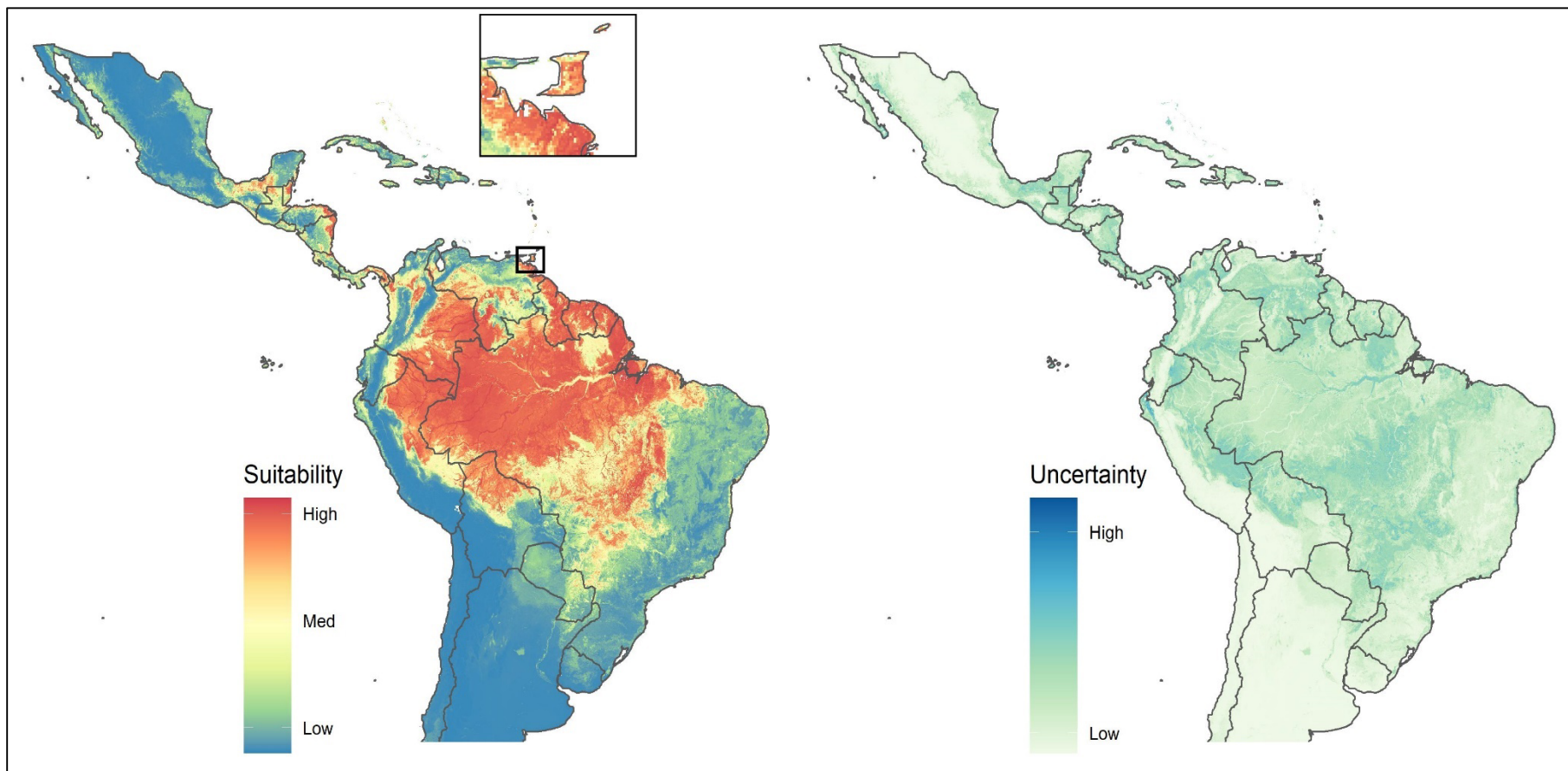


270

271 **Fig 2. Evidence consensus scores.** Evidence consensus is presented at the Admin0 level for all countries in the study (Fig. 2A) and at
 272 the Admin1 level (Fig. 2B) for Brazil. Scores are based on health organization status, date of most recent MAYV occurrence; validity
 273 of MAYV diagnostic test, recency of MAYV outbreaks or clinical cases, and recency of MAYV occurrence in animals or arthropods.
 274 Blue represents very low evidence consensus while red represents very high evidence consensus.

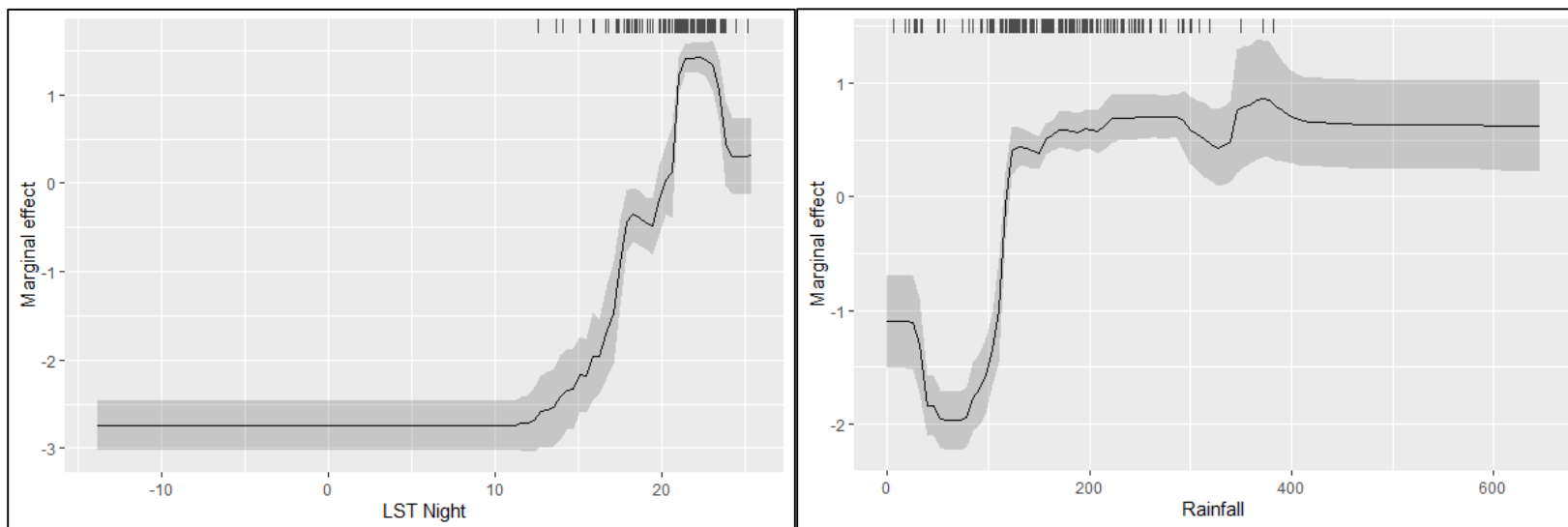


277 **Figure 3. Geographic distribution and temporal trend of MAYV occurrence.** The map shows the distribution of the 203
 278 occurrence locations (before the spatial thinning procedure) that were used to construct the boosted regression tree (BRT) model. The
 279 color corresponds to the host type of each point (human, animal, or arthropod). The inset chart displays total occurrences that were
 280 reported in each year since the initial human case was detected in 1954.



281

282 **Fig 4. Map of environmental suitability and prediction uncertainty for MAYV occurrence.** (A) Suitability ranged from blue (0 -
283 no suitability) to red (1 - very high suitability). The inset map shows greater detail in Trinidad and Tobago. (B) The per-pixel standard
284 deviation across the 100 sub-models is presented as a measure of the model's uncertainty.



285

286 **Fig 5. Partial dependence plots of the two most important variables.** The solid black line represents average response over 100
287 sub-models and the gray region represents one standard deviation. Tick marks represent values of each variable at occurrence location.

288 Discussion

289 In this paper we present an ensemble BRT model of MAYV environmental suitability in the
290 Americas and an evidence consensus framework that integrates multiple information sources. In
291 the absence of comprehensive epidemiological data and reliable and affordable methods for large
292 scale arboviral diagnosis, distribution models can serve an important role in guiding arboviral
293 surveillance (in humans and other hosts) and targeting vector control efforts. This is especially
294 true in the case of MAYV given its nonspecific febrile presentation and the uncertainty
295 surrounding its true distribution and non-human animal reservoirs. Our model provides important
296 information regarding regions of Central and South America that are at highest risk of MAYV
297 transmission allowing us to estimate the total population living in areas suitable for MAYV
298 transmission. Furthermore, we are able to identify areas with high predicted MAYV suitability
299 despite low or moderate evidence consensus.

300 Our model predicted the distribution of MAYV with relatively high accuracy, identifying several
301 regions of high environmental suitability. This included large areas of North and Central Brazil
302 (e.g., the states of Mato Grosso, Pará, and Goiás), French Guiana, Trinidad & Tobago, and
303 Northern Peru, all areas with well-documented evidence of MAYV transmission.²⁵ In addition,
304 several regions with minimal or no published evidence of MAYV transmission were also found
305 to be highly suitable for transmission, including the majority of Guyana as well as coastal
306 regions of Belize, Nicaragua, and Honduras. This finding highlights the utility of distribution
307 models in identifying areas that are particularly receptive to MAYV transmission that could be
308 targeted for increased surveillance or vector control.

309 The wide predicted geographic distribution of MAYV underscores the need for increased
310 surveillance and diagnostic capacity throughout the Americas. Our findings suggest that MAYV
311 is likely vastly underreported and that co-occurring arboviral epidemics (e.g., DENV or CHIKV)
312 may obfuscate the true MAYV disease burden. This is especially true in Brazil, where ~43
313 million people reside in areas that are potentially suitable for MAYV transmission. An additional
314 concern is the detection of MAYV in Haiti,²⁶ where ~3 million people are predicted to be at
315 potential risk according to our model. The discovery of MAYV in Haiti has prompted additional
316 questions about its potential vectors and the possibility of urban transmission, due to the lack of
317 *Haemagogus* mosquitoes and non-human primates on the island. These questions require further
318 entomological investigations in order to elucidate the vectorial capacity of urban mosquitoes and
319 potential animal reservoirs other than non-human primates.

320 Our model has several important limitations that must be considered when interpreting the
321 findings presented here. Georeferenced reports of disease occurrence are subject to sampling bias
322 related to the accessibility of certain locations, availability of laboratory infrastructure, or the
323 presence of a robust surveillance system that is able to detect arbovirus occurrence. Therefore,
324 the presence locations used in our model are likely subject to sampling bias and do not reflect the
325 true distribution of MAYV across the Americas. Following previously published modeling
326 studies,²⁷ we attempted to correct for the sampling bias in our dataset through the use of
327 pseudoabsence points with a similar spatial bias as the presence points. However, it is likely that
328 the model is still affected by some level of bias in the occurrence locations. Ideally, future
329 MAYV distribution models will be updated when new occurrence locations are reported (as in
330 Pigott et al.²⁸ updates to the Ebola niche map), leading to a more accurate prediction of the true
331 MAYV distribution.

332 Another limitation relates to the covariate set used to construct our model. Many aspects of the
333 MAYV epidemiology remain poorly understood and current knowledge of MAYV ecology is
334 lacking. We attempted to include covariates that likely have a strong influence on MAYV
335 transmission, including climate, landcover, and vegetation indices, which impact vector ecology
336 and therefore transmission risk. However, disease risk is influenced by other unknown or
337 understudied variables including socioeconomic factors and the presence of non-human reservoir
338 hosts. Although several non-human primate species appear to be important MAYV reservoirs,⁹
339 we opted not to include primate distribution in our model due to the lack of certainty regarding
340 their precise role in the transmission cycle. For systems like YFV, where the role of particular
341 primate species is better understood, the distributions of particular primate reservoir species may
342 be incorporated directly into predictive models.²⁹ As more research is conducted on MAYV
343 ecology and additional animal reservoirs are identified, future MAYV distribution models should
344 be updated to include primate distribution as a contributor to overall spillover risk. Our results
345 here may assist in targeting and designing animal reservoir studies.

346 Our model contributes several crucial additions to the limited literature on MAYV distribution.¹⁶
347 By including a comprehensive compendium of MAYV occurrence points along with high-
348 resolution, remotely sensed covariates at each point, the current ensemble model generates
349 actionable predictions through its demonstrated high predictive accuracy²⁰ across a wider
350 geographic range than any previously published model. Furthermore, the evidence consensus
351 scores provide an alternate method for estimating MAYV transmission, which complements the
352 modeled predictions. These estimates thus inform where and when both human and non-human
353 MAYV surveillance should be targeted. While it remains difficult to accurately estimate the
354 costs of epidemic response³⁰ our estimates of the total at-risk human population across multiple
355 countries also make clear that the potential risk to human health is disproportionately high
356 compared to investments in understanding the basic ecology and epidemiology of MAYV
357 needed to mitigate the risk of future spillover transmission.

358 **Disclaimer**

359 Material has been reviewed by the Walter Reed Army Institute of Research. There is no
360 objection to its presentation and/or publication. The contents, views or opinions expressed in this
361 publication are those of the authors and do not necessarily reflect official policy or position of
362 Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Uniformed
363 Services University of the Health Sciences, Walter Reed Army Institute of Research, the
364 Department of Defense (DoD), or Departments of the Army, Navy, or Air Force. Mention of
365 trade names, commercial products, or organizations does not imply endorsement by the U.S.
366 Government.

367 SP was supported by the National Institute of Allergy and Infectious Diseases, National Institutes
368 of Health, <https://www.niaid.nih.gov/>, under Inter-Agency Agreement Y1-AI-5072, and the
369 Defense Health Program, U.S. DoD, under award HU0001190002. AP and DP were financially
370 supported by the Armed Forces Health Surveillance Division – Global Emerging Infections
371 Surveillance (AFHSD-GEIS) award P0065_22_WR. The activities undertaken at the Walter
372 Reed Biosystematics Unit were performed in part under a Memorandum of Understanding
373 between the Walter Reed Army Institute of Research (WRAIR) and the Smithsonian Institution,
374 with institutional support provided by both organizations. The funders had no role in study
375 design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

- 1 Anderson, C. R., Downs, W. G., Wattley, G. H., Ahin, N. W. & Reese, A. A. Mayaro virus: a new human disease agent. II. Isolation from blood of patients in Trinidad, B.W.I. *Am J Trop Med Hyg* **6**, 1012-1016, doi:10.4269/ajtmh.1957.6.1012 (1957).
- 2 LeDuc, J. W., Pinheiro, F. P. & Travassos da Rosa, A. P. An outbreak of Mayaro virus disease in Belterra, Brazil. II. Epidemiology. *Am J Trop Med Hyg* **30**, 682-688, doi:10.4269/ajtmh.1981.30.682 (1981).
- 3 Forshey, B. M. *et al.* Arboviral etiologies of acute febrile illnesses in Western South America, 2000-2007. *PLoS Negl Trop Dis* **4**, e787, doi:10.1371/journal.pntd.0000787 (2010).
- 4 Mota, M. T. D., Avilla, C. M. S. & Nogueira, M. L. Mayaro virus: a neglected threat could cause the next worldwide viral epidemic. *Future Virology* **14**, 375-378, doi:10.2217/fvl-2019-0051 (2019).
- 5 Hotez, P. J. & Murray, K. O. Dengue, West Nile virus, chikungunya, Zika-and now Mayaro? *PLoS Negl Trop Dis* **11**, e0005462, doi:10.1371/journal.pntd.0005462 (2017).
- 6 Halsey, E. S. *et al.* Long-term clinical, immunologic and virologic follow-up in a cohort infected with mayaro virus. *American Journal of Tropical Medicine and Hygiene* **87**, 126 (2012).
- 7 Suhrbier, A., Jaffar-Bandjee, M. C. & Gasque, P. Arthritogenic alphaviruses--an overview. *Nat Rev Rheumatol* **8**, 420-429, doi:10.1038/nrrheum.2012.64 (2012).
- 8 Hassing, R. J., Leparc-Goffart, I., Tolou, H., van Doornum, G. & van Genderen, P. J. Cross-reactivity of antibodies to viruses belonging to the Semliki forest serocomplex. *Euro Surveill.* **15** (2010).
- 9 Hoch, A. L., Peterson, N. E., LeDuc, J. W. & Pinheiro, F. P. An outbreak of Mayaro virus disease in Belterra, Brazil. III. Entomological and ecological studies. *Am J Trop Med Hyg* **30**, 689-698, doi:10.4269/ajtmh.1981.30.689 (1981).
- 10 de Thoisy, B., Gardon, J., Salas, R. A., Morvan, J. & Kazanji, M. Mayaro virus in wild mammals, French Guiana. *Emerg Infect Dis* **9**, 1326-1329, doi:10.3201/eid0910.030161 (2003).
- 11 Talarmin, A. *et al.* Mayaro virus fever in French Guiana: isolation, identification, and seroprevalence. *Am J Trop Med Hyg* **59**, 452-456, doi:10.4269/ajtmh.1998.59.452 (1998).
- 12 Izurieta, R. O. *et al.* Hunting in the Rainforest and Mayaro Virus Infection: An emerging Alphavirus in Ecuador. *J Glob Infect Dis* **3**, 317-323, doi:10.4103/0974-777x.91049 (2011).
- 13 Mourao, M. P. *et al.* Mayaro fever in the city of Manaus, Brazil, 2007-2008. *Vector Borne Zoonotic Dis* **12**, 42-46, doi:10.1089/vbz.2011.0669 (2012).
- 14 Wiggins, K., Eastmond, B. & Alto, B. W. Transmission potential of Mayaro virus in Florida *Aedes aegypti* and *Aedes albopictus* mosquitoes. *Med Vet Entomol* **32**, 436-442, doi:10.1111/mve.12322 (2018).
- 15 Pan American Health Organization / World Health Organization. Epidemiological Alert: Mayaro Fever. (2019).
- 16 Lorenz, C., Freitas Ribeiro, A. & Chiaravalloti-Neto, F. Mayaro virus distribution in South America. *Acta Trop* **198**, 105093, doi:10.1016/j.actatropica.2019.105093 (2019).
- 17 Bhatt, S. *et al.* The global distribution and burden of dengue. *Nature* **496**, 504-507, doi:10.1038/nature12060 (2013).
- 18 Messina, J. P. *et al.* Mapping global environmental suitability for Zika virus. *Elife* **5**, doi:10.7554/eLife.15272 (2016).
- 19 Nsoesie, E. O. *et al.* Global distribution and environmental suitability for chikungunya virus, 1952 to 2015. *Euro Surveill* **21**, doi:10.2807/1560-7917.Es.2016.21.20.30234 (2016).
- 20 Elith, J., Leathwick, J.R., Hastie, T. A working guide to boosted regression trees. *Journal of Animal Ecology* **77**, 802-813 (2008).
- 21 Brady, O. J. *et al.* Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl Trop Dis* **6**, e1760, doi:10.1371/journal.pntd.0001760 (2012).

- 22 Celone, M. *et al.* A systematic review and meta-analysis of the potential non-human animal reservoirs and arthropod vectors of the Mayaro virus. *PLoS Negl Trop Dis* **15**, e0010016, doi:10.1371/journal.pntd.0010016 (2021).
- 23 Celone, M. J. *et al.* A geopositioned, evidence-graded, pan-species compendium of Mayaro Virus occurrence. *medRxiv*, 2023.2003.2007.23286930, doi:10.1101/2023.03.07.23286930 (2023).
- 24 Hijmans RJ, P. S., Leathwick J, Elith J. *dismo: Species Distribution Modeling*. 2016. <http://CRAN.R-project.org/package=dismo>.
- 25 Ganjian, N. & Riviere-Cinnamond, A. Mayaro virus in Latin America and the Caribbean. *Revista panamericana de salud publica = Pan American journal of public health* **44**, e14, doi:10.26633/rpsp.2020.14 (2020).
- 26 Blohm, G. *et al.* Mayaro as a Caribbean traveler: Evidence for multiple introductions and transmission of the virus into Haiti. *Int J Infect Dis* **87**, 151-153, doi:10.1016/j.ijid.2019.07.031 (2019).
- 27 Pigott, D. M. *et al.* Mapping the zoonotic niche of Ebola virus disease in Africa. *Elife* **3**, e04395, doi:10.7554/eLife.04395 (2014).
- 28 Pigott, D. M. *et al.* Updates to the zoonotic niche map of Ebola virus disease in Africa. *Elife* **5**, doi:10.7554/eLife.16412 (2016).
- 29 Shearer, F. M. *et al.* Existing and potential infection risk zones of yellow fever worldwide: a modelling analysis. *Lancet Glob Health* **6**, e270-e278, doi:10.1016/s2214-109x(18)30024-x (2018).
- 30 Dobson, A. P. *et al.* Ecology and economics for pandemic prevention. *Science (New York, N.Y.)* **369**, 379-381, doi:10.1126/science.abc3189 (2020).

Appendices to “Understanding transmission risk and predicting environmental suitability for Mayaro Virus in the Americas”

Michael Celone¹, Sean Beeman¹, Barbara A. Han², Alexander M. Potter^{3,4,5}, David B. Pecor^{3,4,5}, Bernard Okech¹, Simon Pollett^{6,7}

¹ Uniformed Services University of the Health Sciences, F. Edward Hébert School of Medicine, Department of Preventive Medicine & Biostatistics, Bethesda, Maryland

² Cary Institute of Ecosystem Studies, Millbrook, NY, USA

³ One Health Branch, Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA

⁴ Walter Reed Biosystematics Unit, Smithsonian Museum Support Center, Suitland, MD 20746, USA

⁵ Department of Entomology, Smithsonian Institution—National Museum of Natural History (NMNH), Washington, DC 20560, USA

⁶ Infectious Disease Clinical Research Program, Department of Preventive Medicine and Biostatistics, Uniformed Services University of the Health Sciences, Bethesda, MD, USA

⁷ Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, MD, USA

*Corresponding author

Michael Celone

Michael.celone@usuhs.edu

Contents

Occurrence records.....	21
Evidence Consensus.....	21
Health Organization Reports (max 3 points).....	21
Peer-reviewed Evidence of Human Infection (max 6 points).....	22
Outbreaks and Clinical Cases (max 6 points).....	22
Animal and Arthropod Data (max 6 points).....	22
Supplementary Table 1. Evidence Categories and Possible Scores	22
Explanatory covariates.....	23
Covariate selection.....	24
Modelling approach.....	24
Model evaluation	25
Supplementary Figure 1. Map of environmental suitability MAYV occurrence using human-only data.	26
Supplementary Figure 2. Difference in predicted environmental suitability between the all-host model and human-only model.....	27
Supplementary Figure 3. Partial dependence plots of the two eight additional variables	28
Supplementary Table 1: Evidence consensus score by country.....	29
Supplementary Table 2. Evidence consensus by state (Brazil).....	32
References.....	35

Occurrence records

We previously developed a georeferenced compendium of MAYV occurrence based on methods that have been established for other pathogens including dengue and leishmaniasis.¹⁻⁴ MAYV occurrence among humans, non-human animals, and arthropods was compiled through a systematic review of the literature, including an evaluation of the quality of such evidence. These methods were described previously in greater detail.⁵ All occurrences were assigned to a point or polygon location, depending on the spatial resolution provided by the authors. Point data comprised precise locations with less than 5km of uncertainty (e.g., specific coordinates or a small town) while polygon data comprised larger areas or administrative units that exceeded 5km of uncertainty. The coordinates of point locations and polygon centroids were recorded in our database and used as the presence records in our current modeling study.

Presence points with ≤ 75 km of uncertainty were included in our current analysis, although the majority of points had substantially less uncertainty. Following a previously published modeling study that accepted up to 65km of uncertainty,⁶ we chose to accept a greater level of uncertainty in our occurrence data in order to include more occurrence locations during the model development process. Due to the limited size of our occurrence dataset, we deemed that the extra information gained from each occurrence record outweighed any issues associated with the greater uncertainty of some occurrence records.

The final MAYV database contained 265 unique georeferences in 15 countries, published between 1954 and 2022. Two-hundred and three of these occurrence points met the ≤ 75 km uncertainty threshold, and thus were eligible for inclusion in our model. We used the spThin package in the R statistical software to reduce clustering of presence records.⁷ A 5km distance threshold was applied to ensure that no more than one presence point occurred within each pixel of our covariate layers.

Evidence Consensus

Collating published reports of MAYV is an important first step in clarifying its distribution. However, heterogeneous surveillance capacity across countries and incomplete or unclear reporting of epidemiological data may impact our ability to definitively say that MAYV is present in a certain location. An evidence consensus approach takes several information sources into account in order to score the total available evidence supporting the presence or absence of a disease in a given country. This approach has been used previously to provide a refined description of the spatial limits of several pathogens including dengue virus,⁸ leishmaniasis,⁹ podoconiosis,¹⁰ and Lassa fever.¹¹ These studies considered multiple data sources to develop an evidence consensus score for disease presence or absence, including health organization status, peer reviewed evidence, case data, animal infection, economic status, and other supplementary evidence. We followed a similar procedure to generate a consensus score for each country in Latin America that quantifies the evidence supporting MAYV presence. This score ranged from 0 (“No evidence of MAYV presence”) to 21 (“Complete evidence of MAYV presence”) based on the categories described below.

Health Organization Reports (max 3 points)

International health organizations have been used previously to support evidence of pathogen presence or absence in specific countries.⁸ We followed a similar procedure, using health reports from two sources: Pan American Health Organization (PAHO)/World Health Organization (WHO) and the Global Infectious Diseases and Epidemiology Online Network (GIDEON). The PAHO issues weekly epidemiological alerts to update the public on the occurrence of significant health events. Similarly, the WHO issues Disease Outbreak News (DONs) related to public health issues of international importance. We searched these PAHO/WHO bulletins for relevant alerts related to MAYV in a given country. Countries were assigned a score of 1 if WHO/PAHO had issued an epidemiological alert for MAYV in that country.

GIDEON is a web application that compiles relevant news on infectious disease outbreaks and designates countries as endemic/potentially endemic for each pathogen. If a country was listed as endemic/potentially endemic for MAYV in the GIDEON database, it was assigned a score of 1. If a country fulfilled both criteria (i.e., listed as endemic in GIDEON and a relevant PAHO/WHO health alert) it was assigned a score of 3.

Peer-reviewed Evidence of Human Infection (max 6 points)

Based on methods proposed by Brady et al.,⁸ peer-reviewed evidence of human infection was scored based on the following two categories: contemporariness (3 for 2011-2020, 2 for 2000-2010, and 1 for 1999 and earlier) and diagnostic accuracy (3 for PCR, viral culture, or PRNT, 2 for serological methods only, 1 for presumed cases without diagnostic test or cases with unspecified diagnostic test). In the case of multiple MAYV reports in a given country, the highest scoring report was used. For example, if one study in Brazil reported serological evidence of MAYV transmission (score of 2) in 1990 (score of 1) and another study in Brazil reported MAYV viral culture (score of 3) in 2019 (score of 3), Brazil would receive a score of 6 for this category. We also considered returning traveler reports for this category if the case was definitively linked to the country of travel. These reports are useful for establishing evidence of pathogen presence because diagnosis is often pursued rigorously for travelers upon returning to their country of origin.⁸

Outbreaks and Clinical Cases (max 6 points)

Reported outbreaks of MAYV or clinical cases that were detected using PCR were scored according to total case numbers and contemporariness. Previous studies have used only the occurrence of reported outbreaks (with no consideration to clinical cases) to assign a score for this category.⁸ However, because only a limited number of MAYV outbreaks have occurred and because diagnostic techniques have been inconsistent across these outbreaks, we also considered clinical cases diagnosed by PCR or viral culture that were not necessarily considered to be an outbreak. The scoring system for this category was adapted based on methods used by Mylne et al.,¹¹ where higher scores were assigned to contemporary reports with 20 or more cases. In order to calculate a score for each country, we summed the cases detected across multiple studies within a single time period. For example, if two separate reports from Peru each detected 10 cases of MAYV using PCR between 2011 and 2020, we summed the reported cases (20 total) and assigned Peru 6 points for this category. However, cases from multiple reports were not summed across different time periods.

A lack of MAYV case reports in a given country is not necessarily indicative of a lack of virus transmission. It is likely that MAYV cases may go undetected due to insufficient surveillance or diagnostic capacity. We attempted to account for this uncertainty using healthcare expenditure (HE) as a proxy of a country's capacity for detecting MAYV occurrence. If no outbreaks or clinical cases were reported in a given country, we used the current HE per capita from the World Health Organization 2017 dataset. Total HE for each country was designated as low ($HE < \$100$), medium ($\$100 \leq HE < \500), or high ($HE \geq \$500$) according to methods that were previously described.⁸ We also considered a country's proximity to other countries that have reported outbreaks or clinical cases diagnosed by PCR.¹¹ Adjacency to countries with outbreaks or clinical cases was combined with HE to assign a score. The highest score was assigned to countries with MAYV outbreaks/clinical cases in two or more neighboring countries and a low HE.

Animal and Arthropod Data (max 6 points)

Detection of MAYV occurrence in non-human animal or arthropod species provides additional evidence of MAYV presence. This may be indicative of the potential for disease spillover into the human population. Previous studies have considered the presence of competent vector species when calculating the evidence consensus score.^{12,13} However, because of uncertainties regarding the role of various mosquitoes in the MAYV transmission cycle (e.g., the possible role of *Aedes aegypti* in urban transmission¹⁴), we considered reports of any wild-caught arthropods that were identified as MAYV positive. Similarly, we assigned a separate score based on reports of infection in potential animal reservoirs. The highest scores for both animal and arthropod MAYV positivity were assigned to more contemporary studies. In the case of multiple MAYV reports in a given country, the highest scoring report was used. All data on MAYV positivity in non-human animals and arthropods was compiled in a systematic review that was previously described.⁵

Supplementary Table 1. Evidence Categories and Possible Scores

Evidence Category	Score
Health organization status	

Both GIDEON and PAHO/WHO	3
Either GIDEON or PAHO/WHO	1
Peer reviewed evidence	
Date of MAYV human occurrence	
2011-2020	3
2000-2010	2
1999 and earlier	1
Diagnostic procedure	
PCR or viral culture or PRNT	3
Serological methods (not including PRNT)	2
Presumptive diagnosis or not specified	1
Outbreaks and clinical cases	
20+ cases from 2011-2020	6
20+ cases from 2000-2010	5
20+ cases 1999 and earlier	4
<20 cases from 2011-2020	3
<20 cases from 2000-2010	2
<20 cases 1999 and earlier	1
If no case data: health expenditure in 2017 and adjacency	
HE <100 USD and 2 or more neighbors	4
100 USD ≤ HE <500 USD and 2 or more neighbors	3
HE <100 USD and 1 neighbor	2
100 USD ≤ HE <500 USD and 1 neighbor	1
Animal data	
Infected animal from 2011-2020	3
Infected animal from 2000-2010	2
Infected animal 1999 and earlier	1
Arthropod data	
Positive arthropod from 2011-2020	3
Positive arthropod from 2000-2010	2
Positive arthropod 1999 and earlier	1
Possible Evidence Consensus Score Categories (Maximum Possible Score: 21)	
Very high evidence of MAYV presence	16-21
High evidence of MAYV presence	11-15
Moderate evidence of MAYV presence	6-10
Little to no evidence of MAYV presence	0-5

Explanatory covariates

We considered several variables for inclusion in our models based on both previous studies of and hypotheses about MAYV ecology. Rainfall data and remotely sensed variables from the MODIS platform were provided by the Malaria Atlas Project (<https://malariaatlas.org/>) after a gap-filling algorithm was used to account for cloud cover.¹⁵ The variables were transformed to ensure matching spatial resolution of 2.5 arc-minutes (~5km) and matching extent. The variables were transformed to ensure that spatial resolution, extent, and boundaries were identical before modeling. Several variables described here were derived from NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) remote sensing platform.¹⁶

Temperature and rainfall play an important role in vector abundance and activity.¹⁷ Entomological surveys have demonstrated an association between *Hg. janthinomys* abundance and temperature^{18,19} and Alencar et al. reported that the mosquito's presence was correlated with high temperatures ranging from 24°C–30°C.²⁰ Several studies have also demonstrated that large diurnal temperature range can impact larval development time, adult survival, and reproductive output in *Aedes* and *Anopheles* populations²¹⁻²³ and an ecological niche model demonstrated that mean diurnal range was one of the most important predictors of *Hg. janthinomys* distribution.²⁴ Humidity and rainfall have also been shown to impact the density of adult *Hg. janthinomys* mosquito populations^{18,25-28} and *Hg. janthinomys* biting activity was shown to peak during intense rainfall in January.²⁶ Due to the impact of temperature and rainfall on vector abundance and vectorial capacity, we included three climate variables in our model, namely night-time and daytime land surface temperature (LST) and cumulative rainfall. LST Night and LST Day are remotely sensed variables from the NASA MODIS MOD11A2 satellite.²⁹ Annual LST Day and LST Night raster layers spanning the years 2000–2020 were used to calculate a single layer representing the mean values over this time period. We also used rainfall data from the Climate Hazards Group InfraRed Precipitation with Station (CHIRPS),³⁰ a quasi-global data set that incorporates satellite imagery at 0.05° resolution and meteorological station data to construct gridded rainfall estimates.

In addition, *Hg. janthinomys* mosquitoes thrive in arboreal habitats (e.g., tropical rainforests) and oviposit in water-filled natural plant cavities (e.g., tree holes or broken bamboo).³¹ Adult mosquitoes have predominantly been found in the forest canopy at heights of 16m and 30m.¹⁸ Therefore, the density of vegetation canopy and moisture supply may influence *Hg. janthinomys* abundance in a given area. To account for these factors, we included three covariates related to vegetation and surface moisture: enhanced vegetation index (EVI), tasseled cap wetness (TCW) and tasseled cap brightness (TCB). The EVI is a measure of vegetation canopy greenness displayed at a 500m spatial resolution³² and is derived from the MODIS MCD43B4 product.³³ The EVI has been used as a covariate in previous niche models of arboviruses including Yellow Fever,³⁴ chikungunya virus,¹² and Zika virus.³⁵ TCW and TCB, measures of surface moisture that are used to assess land cover change, were also generated from the MODIS MCD43B4 product.³⁶

Previous outbreaks of MAYV have occurred in towns close to the rainforest or jungle outposts in close proximity to the forest edge.³⁷⁻³⁹ MAYV most likely circulates in a sylvatic cycle involving canopy-dwelling mosquitoes and non-human primates, with occasional spill-overs into humans living close to the forest.⁴⁰ Entomological surveys have demonstrated that *Hg. janthinomys* are predominantly found in forest canopies at heights of 16m and 30m.¹⁸ Due to the strong impact of land cover on the probability of MAYV occurrence in a given area, we included two land cover covariates, namely evergreen broadleaf forest and urban/built-up, from the MODIS MOD13Q1 product.^{41,42} These covariates represent the proportion of each raster grid cell (ranging from 0-100) that is covered by the land cover class in question, whereby a value of zero represents the absence of land cover and 100 represents complete coverage. Lastly, we included slope and elevation covariates that we accessed from the US Geological Survey's Global Multi-resolution Terrain Elevation Data (GMTED).⁴³

Covariate selection

We implemented a data-driven variable selection process in the R package *SDMtune* to identify variables for inclusion in our models. This process involves removal of highly correlated variables based on an algorithm that first ranks the variables by permutation importance and evaluates the correlation between the most important variable and the remaining variables. A leave-one-out Jack-knife test is then used to remove the variable that has the smallest impact on model performance according to the AUC. However, no variables met the threshold for exclusion based on a Spearman correlation coefficient of 0.8. Therefore, all variables were included in the analysis.

Modelling approach

We used boosted regression trees (BRT) to model the environmental suitability for MAYV occurrence. BRT is a machine learning approach that has been used extensively to develop risk maps of vector-borne pathogens.^{12,13,35,44,45} This algorithm uses regression or classification trees to partition the dataset using recursive binary splits. It also incorporates boosting into the model-building process, a procedure that combines many simple models to improve overall model accuracy. The boosting algorithm is an iterative process that fits many small trees sequentially, building on previously fitted trees to improve model performance.⁴⁶ This process incorporates a level of

stochasticity by randomly selecting a subset of the data to fit each tree, thereby reducing the model variance.⁴⁶ BRT have several advantages including their ability to fit complex nonlinear relationships, handle missing data, and to accommodate many different types of covariates without any need for data transformation.⁴⁶

One of the most important aspects of modeling species distributions with presence-only data is the selection of pseudo-absence points that represent the range of environmental conditions where the species or pathogen was not detected. Random selection of pseudo-absence points may not be appropriate if the presence locations are spatially biased.⁴⁷ In most cases, the detection of disease presence locations may be subject to sampling bias if some locations are more likely to be surveyed than others (e.g., locations that are closer to roads).⁴⁸ Therefore, pseudo-absence points should be selected with a similar level of bias as the presence points to ensure that background and presence locations are biased in the same manner.⁴⁸ Following the methods of previous modeling studies,^{11,49-51} we selected 10,000 background points from the study region, biased towards more populous areas. Therefore, population density was used as a proxy for sampling bias. Pseudoabsence points were selected using the 2° method proposed by Barbet-Massin (2012),⁵² whereby each pseudo-absence point was at least 2° away from a presence location. In order to improve the model's discrimination capacity, the pseudoabsence points were down-weighted to ensure that the weighted sum of presence records equaled the sum of weighted background points.⁵²

We subsequently fitted 100 BRT 'sub-models' to separate bootstraps of the dataset. The bootstrapped datasets were chosen with replacement, subject to the condition that a minimum of 25 presence and 25 pseudo-absence points were selected. This bootstrapping procedure allowed us to quantify the uncertainty across models and to increase the model's robustness.⁵³ Each sub-model was fit in R using the `gbm.step` procedure in the `dismo` package.⁵⁴ This function uses cross-validation to identify the optimal number of trees for each sub-model to improve predictive capacity. The remaining BRT hyperparameters were held at their default values (tree complexity = 4, learning rate = 0.005, bag fraction = 0.75, cross-validation folds = 10, step size = 10). The final prediction map represents the mean MAYV suitability of each 5 x 5 km pixel across our ensemble of 100 models. We also generated a map of the model uncertainty, represented by per-pixel standard deviation. In order to avoid extrapolating the model to regions far outside of the MAYV niche, the model predictions were restricted to the Americas.

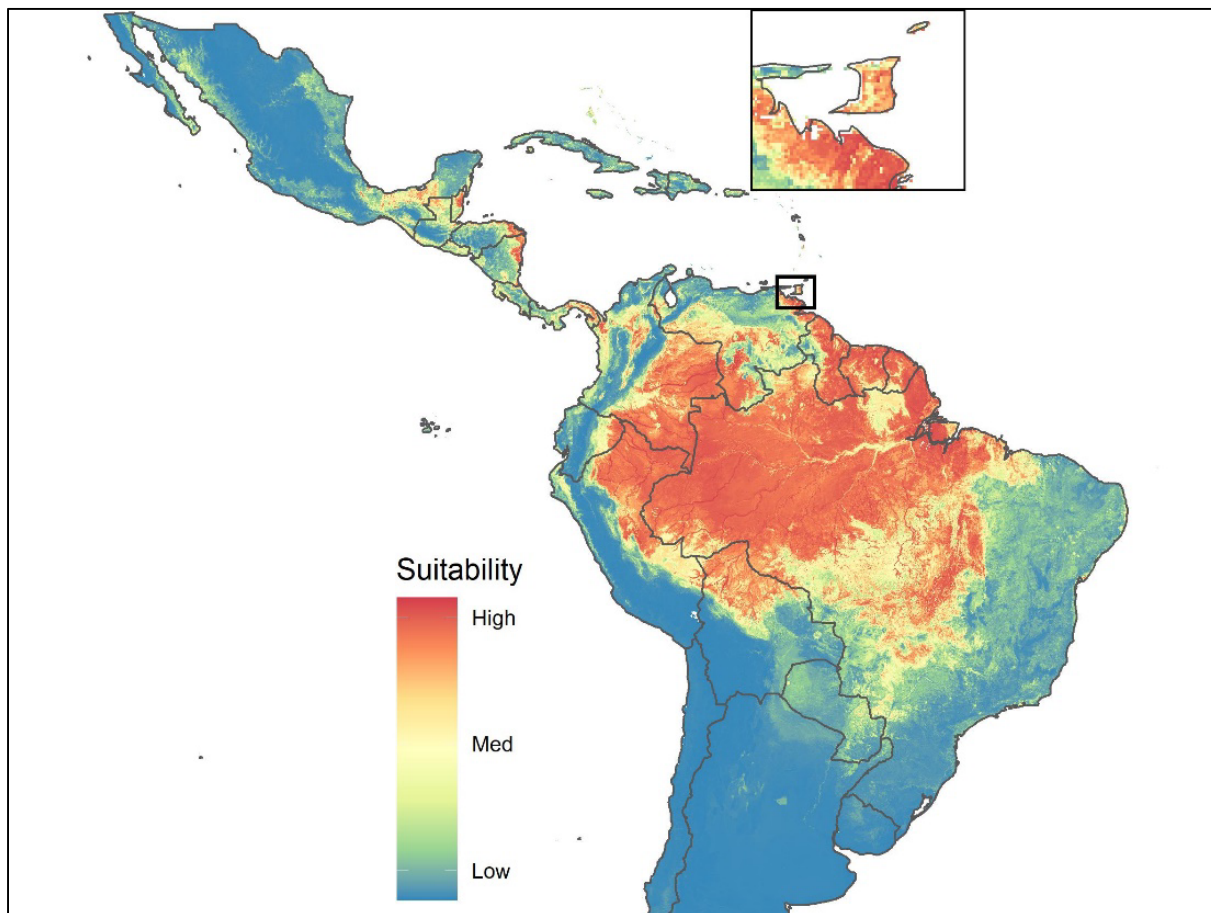
Two models were constructed. The model presented in the main text ("all-host") included all human, arthropod, and non-human animal occurrence data (203 occurrence points overall and 170 points after the spatial thinning procedure). An additional model presented in the Supplementary materials ("human-only") included only human occurrence data (160 occurrence points overall and 143 points after the spatial thinning procedure). Similar to a previous model of Marburg virus risk,⁵¹ we opted to use all occurrence points in the primary model to leverage all available data. We were also interested in exploring how the model predictions might change with the inclusion of multiple host types. We compared the output of the two models by subtracting the raster pixel values, allowing us to visualize the difference in predicted suitability between the human-only model and the all-host model. The all-host model is presented in the results below while the human-only model and comparison of model output is presented in Supplementary Materials.

Model evaluation

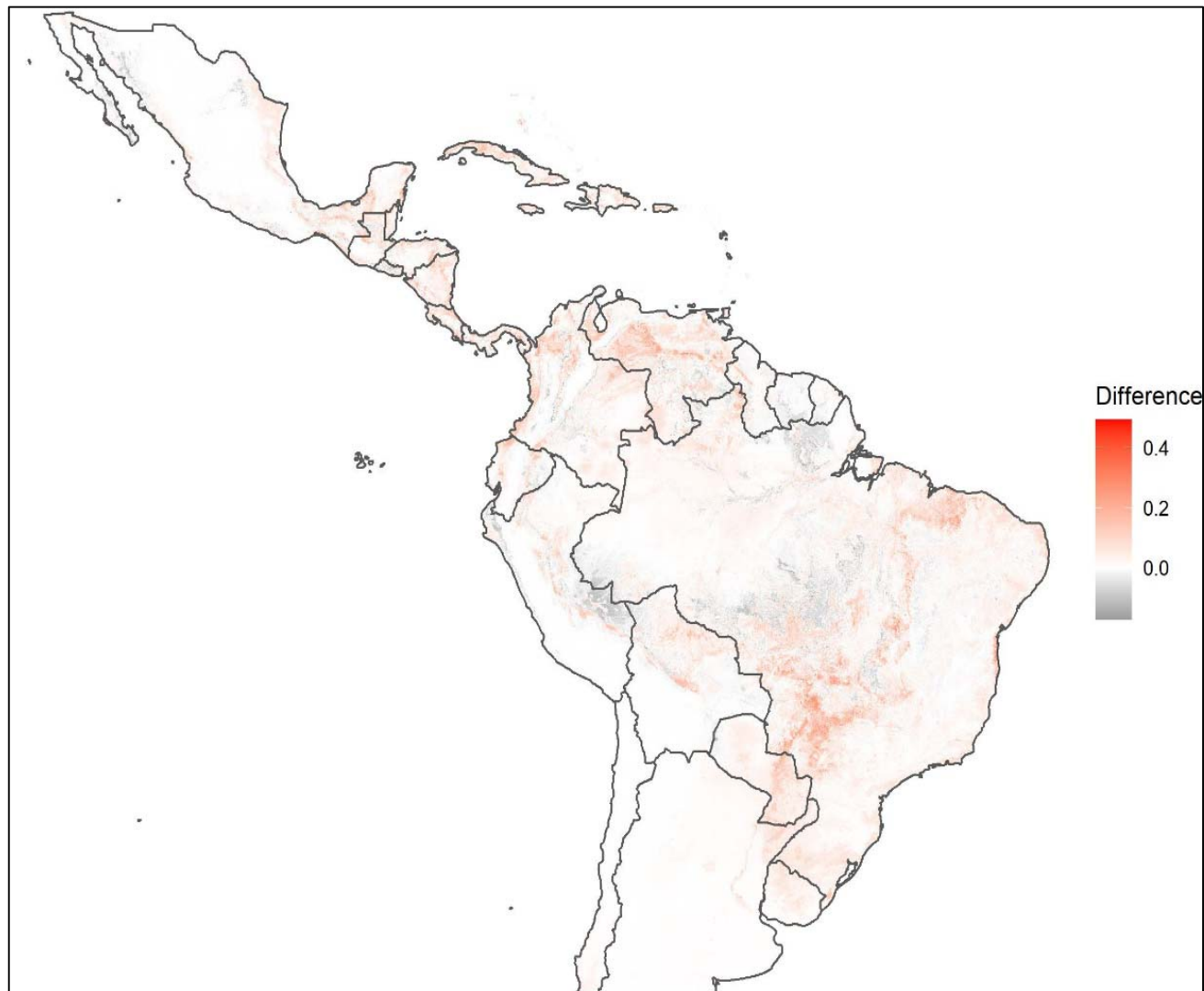
The predictive accuracy of the model was assessed using several performance metrics including the area under the receiver operator characteristic curve (AUC), sensitivity, specificity, Kappa statistic, and percent correctly classified (PCC). Metrics were calculated for each sub-model using 10-fold cross validation. The cross-validation procedure involved randomly splitting each bootstrapped dataset into 10 folds with approximately the same number of presence and absence records in each fold. The model was subsequently trained on nine of the folds and the withheld fold was used to evaluate the model performance. The performance metrics for each sub-model represent the mean values across the 10 folds. These values were then averaged across each of the 100 sub-models to generate an estimate of overall model performance.

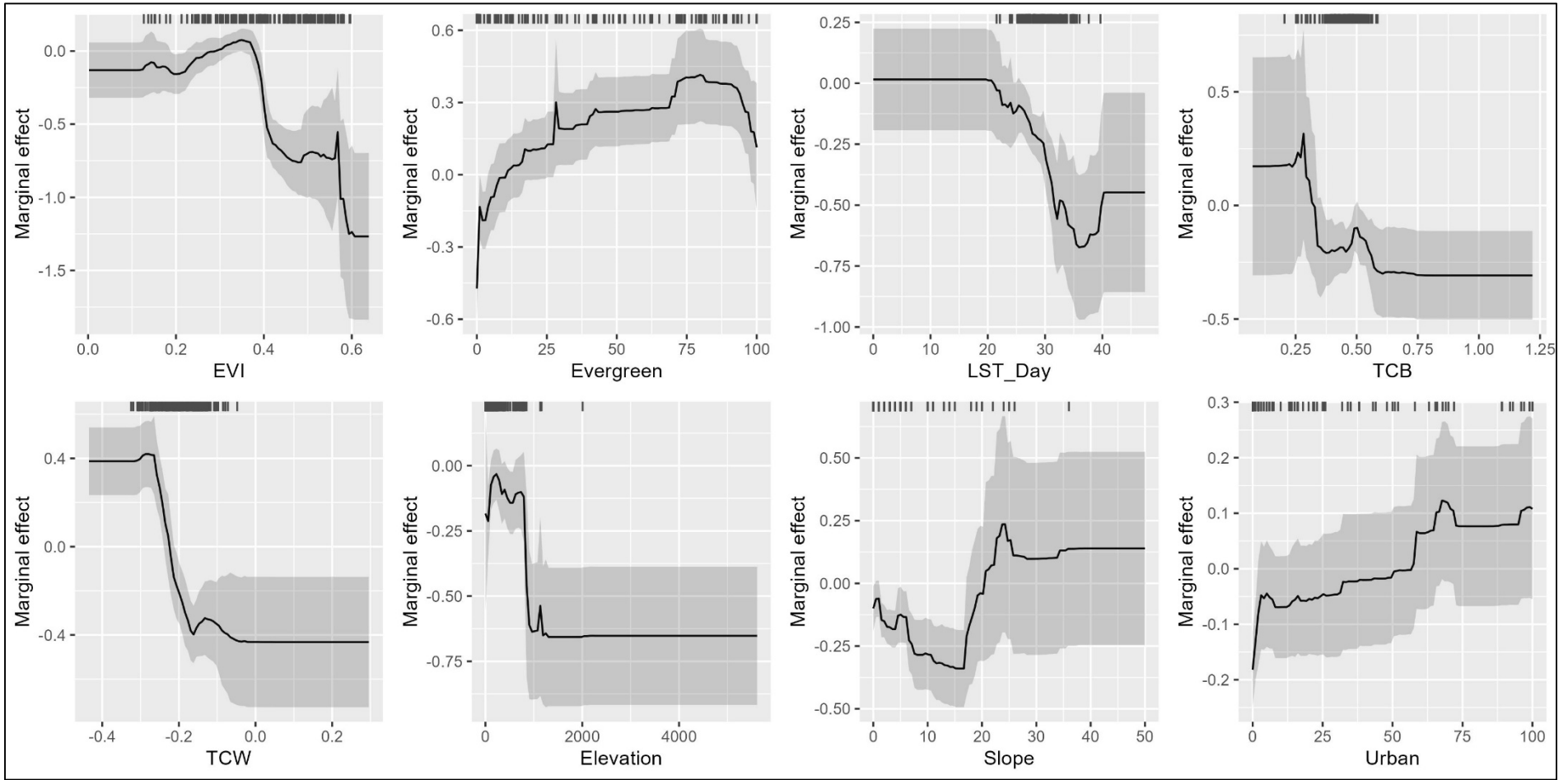
Supplementary Figure 1. Map of environmental suitability MAYV occurrence using human-only data.

Suitability ranged from blue (0- no suitability) to red (1- very high suitability). The inset map shows greater detail in Trinidad and Tobago.



Supplementary Figure 2. Difference in predicted environmental suitability between the all-host model and human-only model. This map was created by subtracting the pixel-wise probability values in the all-host suitability model from the values in the human-only model. Red represents regions where the all-host model predicted higher suitability while black represents areas where the human-only model predicted higher suitability. White pixels represent agreement between the two models.





EVI = Enhanced vegetation index; LST = Land surface temperature; TCB = Tasseled cap brightness; TCW = Tasseled cap wetness.

Supplementary Figure 3. Partial dependence plots of the two eight additional variables. The solid black line represents average response over 100 sub-models and the gray region represents one standard deviation. Tick marks represent values of each variable at occurrence location. Partial dependence plots for the two most important variables are presented in the main text (Figure 5).

Supplementary Table 2: Evidence consensus score by country

Country	Health Org. Status	Date of Human Occurrence	Diagnostic Procedure	Outbreaks and Clinical Cases	Health Expenditure and Adjacency	Animal Data	Arthropod Data	Total Score
Brazil	1	3	3	6	N/A	3	3	19- Very high
	GIDEON only	Most recent occurrence: 2021 ⁵⁵	RT-PCR ⁵⁶	26 clinical cases in 2017-18 ⁵⁶		Detected in horses in 2016 ⁵⁷	Detected in mosquitoes 2018-19 ⁵⁸	
Venezuela	3	3	3	6	N/A	1	0	16- Very high
	GIDEON and PAHO alert ⁵⁹	Most recent occurrence: 2016 ⁶⁰	Isolation ⁶⁰	77 clinical cases during outbreak in 2010 ⁶¹		Detected in sentinel hamster in 1999 ⁶²	None detected	
Peru	1	3	3	6	N/A	2	0	15- High
	GIDEON only	Most recent occurrence: 2019 ⁶³	RT-PCR and isolation ⁶⁴	86 clinical cases in 2016 ⁶⁵		Detected in various animals in 2008 ⁶⁶	None detected	
French Guiana	3	3	3	3	N/A	1	0	13- High
	GIDEON and PAHO alert ⁶⁷	Most recent occurrence: 2020 ⁶⁷	RT-PCR ⁶⁷	13 clinical cases in 2020 ⁶⁷		Detected in various animals in 1994-95 ⁶⁸	None detected	
Trinidad & Tobago	1	3	3	3	N/A	0	3	13- High
	GIDEON only	Most recent occurrence: 2014 ⁶⁹	RT-PCR ⁶⁹	9 clinical cases in 2014 ⁶⁹		None detected	Identified by metatranscriptomics in 2018 ⁷⁰	
Bolivia	1	2	3	5	N/A	0	0	11- High
	GIDEON only	Most recent occurrence: 2007 ⁷¹	RT-PCR and isolation ⁷¹	23 clinical cases from 2000-07 ⁷¹		None detected	None detected	
Colombia	1	1	3	0	4	1	1	11- High
	GIDEON only	Most recent occurrence: 1966 ⁷²	NT ⁷³	No clinical cases		Per capita CHE was \$513; adjacent to Brazil, Venezuela, Peru	Detected in primates in 1957 ⁷³	
Haiti	1	3	3	3	N/A	0	0	10- Moderate
	GIDEON only	Most recent occurrence: 2015 ⁷⁵	RT-PCR ⁷⁵	5 clinical cases 2014-15 ⁷⁵		None detected	None detected	
Ecuador	1	3	3	2	N/A	0	0	9- Moderate

	GIDEON only	Most recent occurrence: 2019 ⁷⁶	Isolation, RT-PCR, or IgM seroconversion (not specified) ⁷¹	1 clinical case 2000-07 ⁷¹		None detected	None detected	
Guyana	1	1	2	0	5	0	0	9- Moderate
	GIDEON only	Most recent occurrence: 1955 ⁷⁷	Serology (not specified) ⁷⁷	No clinical cases	Per capita CHE was \$296; adjacent to Brazil and Venezuela	None detected	None detected	
Panama	1	3	3	0	0	1	1	9- Moderate
	GIDEON only	Most recent occurrence: 2017 ⁷⁸	NT ⁷⁸	No clinical cases	No neighbors	Detected in various animals in 1974-76 ⁷⁹	Isolated from mosquito in 1972-79 ⁸⁰	
Suriname	1	1	3	0	3	0	0	8- Moderate
	GIDEON only	Most recent occurrence: 1964 ⁸¹	Isolation ⁸²	No clinical cases	Per capita CHE was \$474; adjacent to Brazil and French Guiana	None detected	None detected	
Mexico	1	2	2	0	0	0	0	5- Low
	GIDEON only	Most recent occurrence: 2001 ⁸³	IgM ELISA ⁸³	No clinical cases	No neighbors	None detected	None detected	
Paraguay	0	0	0	0	5	0	0	5- Low
	Neither	No cases	No cases	No clinical cases	Per capita CHE was \$400; adjacent to Brazil and Bolivia	None detected	None detected	
Argentina	0	0	0	0	4	0	0	4- Low
	Neither	No cases	No cases	No clinical cases	Per capita CHE was \$1128; adjacent to Brazil and Bolivia	None detected	None detected	
Chile	0	0	0	0	4	0	0	4- Low
	Neither	No cases	No cases	No clinical cases	Per capita CHE was \$1456; adjacent to Bolivia and Peru	None detected	None detected	
Costa Rica	0	1	2	0	0	0	0	3- Low
	Neither	Most recent occurrence: 1968 ⁸⁴	HI test ⁸⁴	No clinical cases	No neighbors	None detected	None detected	
Uruguay	0	0	0	0	1	0	0	1- Low

	Neither	No cases	No cases	No clinical cases	Per capita CHE was \$1590; adjacent to Brazil	None detected	None detected	
Dominican Republic	0	0	0	0	1	0	0	1- Low
	Neither	No cases	No cases	No clinical cases	Per capita CHE was \$462; adjacent to Haiti	None detected	None detected	
Belize	0	0	0	0	0	0	0	0- Low
	Neither	No cases	No cases	No clinical cases	No neighbors	None detected	None detected	
Guatemala	0	0	0	0	0	0	0	0- Low
	Neither	No cases	No cases	No clinical cases	No neighbors	None detected	None detected	
Honduras	0	0	0	0	0	0	0	0- Low
	Neither	No cases	No cases	No clinical cases	No neighbors	None detected	None detected	
Nicaragua	0	0	0	0	0	0	0	0- Low
	Neither	No cases	No cases	No clinical cases	No neighbors	None detected	None detected	
El Salvador	0	0	0	0	0	0	0	0- Low
	Neither	No cases	No cases	No clinical cases	No neighbors	None detected	None detected	

HI: hemagglutination inhibition; NT: neutralization test; RT-PCR: Reverse transcription polymerase chain reaction; CHE: Current health expenditure

Supplementary Table 3. Evidence consensus by state (Brazil)

Country	Health Org. Status ¹	Date of Human Occurrence	Diagnostic Procedure	Outbreaks and Clinical Cases	Health Expenditure ² and Adjacency	Animal Data	Arthropod Data	Total Score
Mato Grosso	1	3	3	6	N/A	0	3	16- Very high
	GIDEON only	Most recent occurrence: 2017 ⁸⁵	RT-PCR ⁸⁶	68 positives by RT-PCR from 2011-17 ⁸⁶⁻⁸⁸		No cases	Mosquito pools positive by RT-PCR and isolation in 2018 ⁸⁹	
Goiás	1	3	3	6	N/A	0	3	16- Very high
	GIDEON only	Most recent occurrence: 2017-18 ^{56,90}	RT-PCR ^{56,90}	104 positives by RT-PCR from 2016-17 ^{56,90}		No cases	Mosquitoes positive by RT-PCR in 2018-2019 ⁵⁸	
Pará	1	3	3	3	N/A	2	2	14- High
	GIDEON only	Most recent occurrence: 2016 ⁹¹	RT-PCR ⁹¹	4 positives by RT-PCR in 2016 ⁹¹		Infected animals in 2009 ⁹²	Isolated from mosquitoes in 2008 ⁹³	
Roraima	1	3	3	3	N/A	3	0	13- High
	GIDEON only	Most recent occurrence: 2012 ⁹⁴	RT-PCR ⁹⁴	7 positives by RT-PCR in 2012 ⁹⁴		Infected animals in 2016 ⁵⁷	No cases	
Bahia	1	1	2	0	3	3	2	12- High
	GIDEON only	Most recent occurrence: 1984 ⁹⁵	Serology ⁹⁵	No outbreak or clinical cases	Per capita CHE was \$172; adjacent to Piauí and Goiás	Infected animal from 2012-17 ⁹⁶	One pool positive by RT-PCR from 2009-14 ⁹⁷	
Piauí	1	3	3	3	N/A	2	0	12- High
	GIDEON only	Most recent occurrence: 2016-17 ⁹⁸	RT-PCR ⁹⁸	One positive by RT-PCR from 2016-17 ⁹⁸		Infected animal from 2008-10 ⁹⁹	No cases	
Mato Grosso do Sul	1	2	3	2	N/A	3	0	11- High
	GIDEON only	Most recent occurrence: 2000 ¹⁰⁰	Viral culture ¹⁰⁰	One isolate in 2000 ¹⁰⁰		Infected animals from 2012-14 ¹⁰¹	No cases	
Amazonas	1	3	3	3	N/A	0	0	10- Moderate
	GIDEON only	Most recent occurrence: 2016	RT-PCR ¹⁰²	13 positives by RT-PCR from 2014-16 ¹⁰²		No cases	No cases	

Maranhão	1	3	3	3	N/A	0	0	10- Moderate
	GIDEON only	Most recent occurrence: 2016-18 ¹⁰³	RT-PCR ¹⁰³	One positive by RT-PCR from 2016-18 ¹⁰³		No cases	No cases	
Tocantins	1	3	3	3	N/A	0	0	10- Moderate
	GIDEON only	Most recent occurrence: 2017 ¹⁰⁴	RT-PCR ¹⁰⁴	6 positives by RT-PCR in 2017 ¹⁰⁴		No cases	No cases	
Acre	1	2	2	0	3	0	0	8- Moderate
	GIDEON only	Most recent occurrence: 2004 ¹⁰⁵	Serology ¹⁰⁵	No outbreak or clinical cases	Per capita CHE was \$258; adjacent to Amazonas and Peru	No cases	No cases	
São Paulo	1	3	2	0	1	0	0	7- Moderate
	GIDEON only	Most recent occurrence: 2017 ¹⁰⁶	ELISA ¹⁰⁶	No clinical cases	Per capita CHE was \$252; adjacent to Mato Grosso do Sul	No cases	No cases	
Amapá	1	1	2	0	3	0	0	7- Moderate
	GIDEON only	Most recent occurrence: 1995 ¹⁰⁷	Serology ¹⁰⁷	No outbreak or clinical cases	Per capita CHE was \$220; adjacent to French Guiana and Para	No cases	No cases	
Paraíba	1	1	3	0	0	2	0	7- Moderate
	GIDEON only	Most recent occurrence: 1964 ¹⁰⁸	Plaque reduction NT ¹⁰⁸	No outbreak or clinical cases	No neighbors	Infected animal from 2008-10 ^{99,109}	No cases	
Rio Grande do Sul	1	0	0	0	0	2	3	6- Moderate
	GIDEON only	No cases	No cases	No clinical cases	No neighbors	Infected animal in 2002 ¹¹⁰	Isolated from mosquitoes 2011	
Rondônia	1	0	0	0	3	1	0	5- Low
	GIDEON only	No cases	No cases	No outbreak or clinical cases	Per capita CHE was \$225; adjacent to Bolivia, Amazonas, Mato Grosso	Infected animals in 1987-88 ¹¹¹	No cases	
Rio de Janeiro	1	3	1	0	0	0	0	5- Low
	GIDEON only	Most recent occurrence: 2019 ¹¹²	Not specified ¹¹²	No clinical cases	No neighbors	No cases	No cases	
Pernambuco	1	0	0	0	1	2	0	4- Low
	GIDEON only	No cases	No cases	No outbreak or	Per capita CHE was	Infected animal	No cases	

				clinical cases	\$188; adjacent to Piauí	from 2008-10 ⁹⁹		
Minas Gerais	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	3 Per capita CHE was \$216; adjacent to Goiás and Mato Grosso do Sul	0 No cases	0 No cases	4- Low
Alagoas	1 GIDEON only	0 No cases	0 No cases	0 No outbreak or clinical cases	0 No neighbors	2 Infected animal from 2008-10 ⁹⁹	0 No cases	3- Low
	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	2 Infected animal from 2008-10 ⁹⁹	0 No cases	
Rio Grande do Norte	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	2 Infected animal from 2008-10 ⁹⁹	0 No cases	3- Low
	1 GIDEON only	0 No cases	0 No cases	0 No outbreak or clinical cases	1 Per capita CHE was \$231; adjacent to Mato Grosso do Sul	0 No cases	0 No cases	
Ceará	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	1 Per capita CHE was \$184; adjacent to Piauí	0 No cases	0 No cases	2- Low
	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	0 No cases	0 No cases	
Espírito Santo	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	0 No cases	0 No cases	1- Low
	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	0 No cases	0 No cases	
Santa Catarina	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	0 No cases	0 No cases	1- Low
	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	0 No cases	0 No cases	
Sergipe	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	0 No cases	0 No cases	1- Low
	1 GIDEON only	0 No cases	0 No cases	0 No clinical cases	0 No neighbors	0 No cases	0 No cases	

1 Due to lack of data granularity, the overall Health Organization Status score for the country of Brazil was applied to each administrative unit

2 Mention Alternate data source

References

- 1 Ramshaw, R. E. *et al.* A database of geopositioned Middle East Respiratory Syndrome Coronavirus occurrences. *Sci Data* **6**, 318, doi:10.1038/s41597-019-0330-0 (2019).
- 2 Messina, J. P. *et al.* A global compendium of human dengue virus occurrence. *Sci Data* **1**, 140004, doi:10.1038/sdata.2014.4 (2014).
- 3 Messina, J. P. *et al.* A global compendium of human Crimean-Congo haemorrhagic fever virus occurrence. *Sci Data* **2**, 150016, doi:10.1038/sdata.2015.16 (2015).
- 4 Pigott, D. M. *et al.* Global database of leishmaniasis occurrence locations, 1960-2012. *Sci Data* **1**, 140036, doi:10.1038/sdata.2014.36 (2014).
- 5 Celone, M. *et al.* A systematic review and meta-analysis of the potential non-human animal reservoirs and arthropod vectors of the Mayaro virus. *PLoS Negl Trop Dis* **15**, e0010016, doi:10.1371/journal.pntd.0010016 (2021).
- 6 Carlson, C. J., Dougherty, E. R. & Getz, W. An Ecological Assessment of the Pandemic Threat of Zika Virus. *PLoS Negl Trop Dis* **10**, e0004968, doi:10.1371/journal.pntd.0004968 (2016).
- 7 Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B. & Anderson, R. P. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography* **38**, 541-545 (2015).
- 8 Brady, O. J. *et al.* Refining the global spatial limits of dengue virus transmission by evidence-based consensus. *PLoS Negl Trop Dis* **6**, e1760, doi:10.1371/journal.pntd.0001760 (2012).
- 9 Pigott, D. M. *et al.* Global distribution maps of the leishmaniasis. *Elife* **3**, doi:10.7554/eLife.02851 (2014).
- 10 Deribe, K. *et al.* Mapping the global distribution of podocniosis: Applying an evidence consensus approach. *PLoS Negl Trop Dis* **13**, e0007925, doi:10.1371/journal.pntd.0007925 (2019).
- 11 Mylne, A. Q. *et al.* Mapping the zoonotic niche of Lassa fever in Africa. *Trans R Soc Trop Med Hyg* **109**, 483-492, doi:10.1093/trstmh/trv047 (2015).
- 12 Nsoesie, E. O. *et al.* Global distribution and environmental suitability for chikungunya virus, 1952 to 2015. *Euro Surveill* **21**, doi:10.2807/1560-7917.Es.2016.21.20.30234 (2016).
- 13 Messina, J. P. *et al.* The global distribution of Crimean-Congo hemorrhagic fever. *Trans R Soc Trop Med Hyg* **109**, 503-513, doi:10.1093/trstmh/trv050 (2015).
- 14 Mourao, M. P. *et al.* Mayaro fever in the city of Manaus, Brazil, 2007-2008. *Vector Borne Zoonotic Dis* **12**, 42-46, doi:10.1089/vbz.2011.0669 (2012).
- 15 Weiss, D. J. *et al.* An effective approach for gap-filling continental scale remotely sensed time-series. *ISPRS J Photogramm Remote Sens* **98**, 106-118, doi:10.1016/j.isprsjprs.2014.10.001 (2014).

- 16 Tatem, A. J., Goetz, S. J. & Hay, S. I. Terra and Aqua: new data for epidemiology and public health. *Int J Appl Earth Obs Geoinf* **6**, 33-46, doi:10.1016/j.jag.2004.07.001 (2004).
- 17 Huber, J. H., Childs, M. L., Caldwell, J. M. & Mordecai, E. A. Seasonal temperature variation influences climate suitability for dengue, chikungunya, and Zika transmission. *PLoS Negl Trop Dis* **12**, e0006451, doi:10.1371/journal.pntd.0006451 (2018).
- 18 Pinto, C. S., Confalonieri, U. E. & Mascarenhas, B. M. Ecology of *Haemagogus* sp. and *Sabethes* sp. (Diptera: Culicidae) in relation to the microclimates of the Caxiuanã National Forest, Pará, Brazil. *Memórias do Instituto Oswaldo Cruz* **104**, 592-598 (2009).
- 19 Hendy, A. *et al.* The vertical stratification of potential bridge vectors of mosquito-borne viruses in a central Amazonian forest bordering Manaus, Brazil. *Sci Rep* **10**, 18254, doi:10.1038/s41598-020-75178-3 (2020).
- 20 Alencar, J. *et al.* Circadian and seasonal preferences for hematophagy among *Haemagogus capricornii*, *Hg. janthinomys*, and *Hg. leucocelaenus* (Diptera: Culicidae) in different regions of Brazil. *Journal of Vector Ecology* **33**, 389-392 (2008).
- 21 Beck-Johnson, L. M. *et al.* The importance of temperature fluctuations in understanding mosquito population dynamics and malaria risk. *R Soc Open Sci* **4**, 160969, doi:10.1098/rsos.160969 (2017).
- 22 Carrington, L. B., Seifert, S. N., Willits, N. H., Lambrechts, L. & Scott, T. W. Large diurnal temperature fluctuations negatively influence *Aedes aegypti* (Diptera: Culicidae) life-history traits. *J Med Entomol* **50**, 43-51, doi:10.1603/me11242 (2013).
- 23 Lambrechts, L. *et al.* Impact of daily temperature fluctuations on dengue virus transmission by *Aedes aegypti*. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 7460-7465, doi:10.1073/pnas.1101377108 (2011).
- 24 Celone, M. *et al.* An ecological niche model to predict the geographic distribution of *Haemagogus janthinomys*, Dyar, 1921 a yellow fever and Mayaro virus vector, in South America. *PLoS Negl Trop Dis* **16**, e0010564, doi:10.1371/journal.pntd.0010564 (2022).
- 25 Chadee, D. D., Tikasingh, E. S. & Ganesh, R. Seasonality, biting cycle and parity of the yellow fever vector mosquito *Haemagogus janthinomys* in Trinidad. *Med Vet Entomol* **6**, 143-148, doi:10.1111/j.1365-2915.1992.tb00592.x (1992).
- 26 Tubaki, R. M., de Menezes, R. M. T., Vesgueiro, F. T. & Cardoso, R. P. Observations on *Haemagogus janthinomys* Dyar (Diptera: Culicidae) and other Mosquito Populations within Tree Holes in a Gallery Forest in the Northwestern Region of Sao Paulo State, Brazil. *Neotropical Entomology* **39**, 664-670, doi:10.1590/s1519-566x2010000400030 (2010).
- 27 Alencar, J. *et al.* INFLUENCE OF CLIMATIC FACTORS ON THE POPULATION DYNAMICS OF HAEMAGOGUS JANTHINOMYS (DIPTERA: CULICIDAE), A VECTOR OF SYLVATIC YELLOW FEVER. *Entomological News* **121**, 45-52, doi:10.3157/021.121.0109 (2010).
- 28 Silva, S. O. F., Ferreira de Mello, C., Figueiró, R., de Aguiar Maia, D. & Alencar, J. Distribution of the Mosquito Communities (Diptera: Culicidae) in Oviposition Traps Introduced into the

- Atlantic Forest in the State of Rio de Janeiro, Brazil. *Vector borne and zoonotic diseases* (Larchmont, N.Y.) **18**, 214-221, doi:10.1089/vbz.2017.2222 (2018).
- 29 Wan, Z. & Li, Z.-L. A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. *IEEE Transactions on Geoscience and Remote Sensing* **35**, 980-996 (1997).
- 30 Funk, C. *et al.* The climate hazards infrared precipitation with stations--a new environmental record for monitoring extremes. *Sci Data* **2**, 150066, doi:10.1038/sdata.2015.66 (2015).
- 31 Ali, R. *et al.* Changing patterns in the distribution of the Mayaro virus vector *Haemagogus* species in Trinidad, West Indies. *Acta Trop* **199**, 105108, doi:10.1016/j.actatropica.2019.105108 (2019).
- 32 Didan, K. (ed NASA EOSDIS Land Processes DAAC) (2015).
- 33 Lin, Q. in *2012 5th International Congress on Image and Signal Processing*. 1043-1046.
- 34 Shearer, F. M. *et al.* Existing and potential infection risk zones of yellow fever worldwide: a modelling analysis. *Lancet Glob Health* **6**, e270-e278, doi:10.1016/s2214-109x(18)30024-x (2018).
- 35 Messina, J. P. *et al.* Mapping global environmental suitability for Zika virus. *Elife* **5**, doi:10.7554/eLife.15272 (2016).
- 36 Lobser, S. & Cohen, W. MODIS tasselled cap: land cover characteristics expressed through transformed MODIS data. *International Journal of Remote Sensing* **28**, 5079-5101 (2007).
- 37 Izurieta, R. O. *et al.* Hunting in the Rainforest and Mayaro Virus Infection: An emerging Alphavirus in Ecuador. *J Glob Infect Dis* **3**, 317-323, doi:10.4103/0974-777x.91049 (2011).
- 38 Schaeffer, M., Gajdusek, D. C., Lema, A. B. & Eichenwald, H. Epidemic jungle fevers among Okinawan colonists in the Bolivian rain forest. I. Epidemiology. *Am J Trop Med Hyg* **8**, 372-396, doi:10.4269/ajtmh.1959.8.372 (1959).
- 39 LeDuc, J. W., Pinheiro, F. P. & Travassos da Rosa, A. P. An outbreak of Mayaro virus disease in Belterra, Brazil. II. Epidemiology. *Am J Trop Med Hyg* **30**, 682-688, doi:10.4269/ajtmh.1981.30.682 (1981).
- 40 Hoch, A. L., Peterson, N. E., LeDuc, J. W. & Pinheiro, F. P. An outbreak of Mayaro virus disease in Belterra, Brazil. III. Entomological and ecological studies. *Am J Trop Med Hyg* **30**, 689-698, doi:10.4269/ajtmh.1981.30.689 (1981).
- 41 Friedl, M. & Sulla-Menashe, D. (ed NASA EOSDIS Land Processes DAAC.) (2015).
- 42 Friedl, M. A. *et al.* MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote sensing of Environment* **114**, 168-182 (2010).
- 43 Danielson, J. J. & Gesch, D. B. Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010). 26 (2011).

- 44 Hardcastle, A. N. *et al.* Informing Rift Valley Fever preparedness by mapping seasonally varying environmental suitability. *Int J Infect Dis* **99**, 362-372, doi:10.1016/j.ijid.2020.07.043 (2020).
- 45 Messina, J. P. *et al.* The current and future global distribution and population at risk of dengue. *Nat Microbiol* **4**, 1508-1515, doi:10.1038/s41564-019-0476-8 (2019).
- 46 Elith, J., Leathwick, J.R., Hastie, T. A working guide to boosted regression trees. *Journal of Animal Ecology* **77**, 802-813 (2008).
- 47 Phillips, S. J. D. M. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography* **31**, 161-175 (2008).
- 48 Phillips, S. J. *et al.* Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecol Appl* **19**, 181-197, doi:10.1890/07-2153.1 (2009).
- 49 Pigott, D. M. *et al.* Mapping the zoonotic niche of Ebola virus disease in Africa. *Elife* **3**, e04395, doi:10.7554/eLife.04395 (2014).
- 50 Pigott, D. M. *et al.* Mapping the zoonotic niche of Marburg virus disease in Africa. *Trans R Soc Trop Med Hyg* **109**, 366-378, doi:10.1093/trstmh/trv024 (2015).
- 51 Shearer, F. M. *et al.* Estimating Geographical Variation in the Risk of Zoonotic Plasmodium knowlesi Infection in Countries Eliminating Malaria. *PLoS Negl Trop Dis* **10**, e0004915, doi:10.1371/journal.pntd.0004915 (2016).
- 52 Barbet-Massin, M., Jiguet, F., Albert, C. H. & Thuiller, W. Selecting pseudo-absences for species distribution models: how, where and how many? *Methods in ecology and evolution* **3**, 327-338 (2012).
- 53 Araújo, M. B. & New, M. Ensemble forecasting of species distributions. *Trends in ecology & evolution* **22**, 42-47 (2007).
- 54 Hijmans RJ, P. S., Leathwick J, Elith J. dismo: Species Distribution Modeling. 2016. <http://CRAN.R-project.org/package=dismo>.
- 55 Outbreak News. *Brazil: 8 Mayaro virus cases reported in Pará in 2021*, <<https://outbreaknewstoday.com/brazil-8-mayaro-virus-cases-reported-in-para-in-2021/>> (2022).
- 56 de Paula Silveira-Lacerda, E. *et al.* Molecular epidemiological investigation of Mayaro virus in febrile patients from Goiania City, 2017-2018. *Infect Genet Evol*, 104981, doi:10.1016/j.meegid.2021.104981 (2021).
- 57 Gomes, F. A. *et al.* Serological evidence of arboviruses and coccidia infecting horses in the Amazonian region of Brazil. *PloS One* **14**, e0225895, doi:10.1371/journal.pone.0225895 (2019).
- 58 de Curcio, J. S. *et al.* Detection of Mayaro virus in Aedes aegypti mosquitoes circulating in Goiânia-Goiás-Brazil. *Microbes Infect* **24**, 104948, doi:10.1016/j.micinf.2022.104948 (2022).
- 59 Pan American Health Organization. *Epidemiological Alert, Outbreak of Mayaro Fever in the Americas (7 June 2010)*, <<https://iris.paho.org/handle/10665.2/50840>> (2010).

- 60 Blohm, G. M. *et al.* Isolation of Mayaro Virus from a Venezuelan Patient with Febrile Illness, Arthralgias, and Rash: Further Evidence of Regional Strain Circulation and Possible Long-Term Endemicity. *Am J Trop Med Hyg* **101**, 1219-1225, doi:10.4269/ajtmh.19-0357 (2019).
- 61 Auguste, A. J. *et al.* Evolutionary and Ecological Characterization of Mayaro Virus Strains Isolated during an Outbreak, Venezuela, 2010. *Emerging Infectious Diseases* **21**, 1742-1750, doi:10.3201/eid2110.141660 (2015).
- 62 Medina, G. *et al.* Genetic diversity of Venezuelan alphaviruses and circulation of a Venezuelan equine encephalitis virus subtype IAB strain during an interepizootic period. *Am J Trop Med Hyg* **93**, 7-10, doi:10.4269/ajtmh.14-0543 (2015).
- 63 Gobierno Regional de Cusco - Direccion Regional de Salud Cusco. Boletín de Vigilancia en Salud Pública. *Direccion Ejecutiva de Inteligencia Sanitaria*, Vol. **XIX - No. 8** (2019).
- 64 de la Cruz, C. H., Martínez, S. L. A., Failoc-Rojas, V. E. & Aguilar-Gamboa, F. R. Momento de considerar otras arbovirosis luego del virus mayaro. *Rev Cubana Med Gen Integr* **35** (2019).
- 65 Aguilar-Luis, M. A. *et al.* A silent public health threat: emergence of Mayaro virus and co-infection with Dengue in Peru. *BMC Res Notes* **14**, 29, doi:10.1186/s13104-021-05444-8 (2021).
- 66 Perez, J. G. *et al.* Serologic Evidence of Zoonotic Alphaviruses in Humans from an Indigenous Community in the Peruvian Amazon. *Am J Trop Med Hyg*, doi:10.4269/ajtmh.18-0850 (2019).
- 67 World Health Organization. *Mayaro virus disease - French Guiana, France*, <<https://www.who.int/csr/don/25-october-2020-mayaro-fever-french-guiana-france/en/>> (2020).
- 68 de Thoisy, B., Gardon, J., Salas, R. A., Morvan, J. & Kazanji, M. Mayaro virus in wild mammals, French Guiana. *Emerg Infect Dis* **9**, 1326-1329, doi:10.3201/eid0910.030161 (2003).
- 69 Gonzalez-Escobar, G., Churaman, C., Rampersad, C., Singh, R. & Nathaniel, S. Mayaro virus detection in patients from rural and urban areas in Trinidad and Tobago during the Chikungunya and Zika virus outbreaks. *Pathog Glob Health*, 1-9, doi:10.1080/20477724.2021.1878445 (2021).
- 70 Ali, R. *et al.* Characterization of the virome associated with Haemagogus mosquitoes in Trinidad, West Indies. *Sci Rep* **11**, 16584, doi:10.1038/s41598-021-95842-6 (2021).
- 71 Forshey, B. M. *et al.* Arboviral etiologies of acute febrile illnesses in Western South America, 2000-2007. *PLoS Negl Trop Dis* **4**, e787, doi:10.1371/journal.pntd.0000787 (2010).
- 72 Prías-Landínez, D. E., Bernal-Cúbides, C., de Torres, S. V. & Romero-León, M. Encuesta serologica de virus transmitidos por artrópodos. *Bol Oficina Sanit Panam* (1970).
- 73 Groot, H. Estudios sobre virus transmitidos por artrópodos en Colombia. *Rev Acad Colomb Cienc* **12**, 191-217, doi:10.18257/raccefyn.565 (1964).
- 74 Groot, H., Morales, A. & Vidales, H. Virus isolations from forest mosquitoes in San Vicente de Chucuri, Colombia. *Am J Trop Med Hyg* **10**, 397-402, doi:10.4269/ajtmh.1961.10.397 (1961).

- 75 Blohm, G. *et al.* Mayaro as a Caribbean traveler: Evidence for multiple introductions and transmission of the virus into Haiti. *Int J Infect Dis* **87**, 151-153, doi:10.1016/j.ijid.2019.07.031 (2019).
- 76 Pan American Health Organization / World Health Organization. Epidemiological Alert: Mayaro Fever. (2019).
- 77 Downs, W. G. & Anderson, C. R. Distribution of Immunity to Mayaro Virus Infection in the West Indies. *West Indian Med J* **7**, 190-194 (1958).
- 78 Carrera, J.-P. *et al.* Endemic and epidemic human alphavirus infections in Eastern Panama; An Analysis of Population-based Cross-Sectional Surveys. *bioRxiv*, 2020.2001.2010.901462, doi:10.1101/2020.01.10.901462 (2020).
- 79 Seymour, C., Peralta, P. H. & Montgomery, G. G. Serologic evidence of natural togavirus infections in Panamanian sloths and other vertebrates. *Am J Trop Med Hyg* **32**, 854-861, doi:10.4269/ajtmh.1983.32.854 (1983).
- 80 Galindo, P., Adames, A., Peralta, P., Johnson, C. & Read, R. Impacto de la hidroeléctrica de Bayano en la transmisión de arbovirus. *Rev Med Pan* **8**, 89-134 (1983).
- 81 Karbaat, J. Arbovirus infections in Dutch military personnel stationed in Surinam. *Trop Geogr Med* **17**, 187-189 (1964).
- 82 Metselaar, D. Isolation of arboviruses of group A and group C in Surinam. *Trop Geogr Med* **18**, 137-142 (1966).
- 83 Navarrete-Espinosa, J. & Gomez-Dantes, H. Arbovirus causales de fiebre hemorrágica en pacientes del Instituto Mexicano del Seguro Social. *Rev Med Inst Mex Seguro Soc* **44**, 347-353 (2006).
- 84 Fuentes, L. G. & Mora, J. A. Encuesta serologica sobre arbovirus en Costa Rica. *Rev. Latinoam. Microbiol.* **13**, 25-28 (1971).
- 85 Fumagalli, M. J. *et al.* Development of an Enzyme-Linked Immunosorbent Assay To Detect Antibodies Targeting Recombinant Envelope Protein 2 of Mayaro Virus. *Journal of clinical microbiology* **57**, doi:10.1128/jcm.01892-18 (2019).
- 86 de Souza Costa, M. C. *et al.* Arbovirus investigation in patients from Mato Grosso during Zika and Chikungunya virus introduction in Brazil, 2015-2016. *Acta Trop* **190**, 395-402, doi:10.1016/j.actatropica.2018.12.019 (2019).
- 87 Pilatti, M. *et al.* Perfil clínico-epidemiológico dos pacientes infectados com o vírus Mayaro (MAYV) em Mato Grosso. *TCC-Biomedicina* (2018).
- 88 Zuchi, N., Heinen, L. B., Santos, M. A., Pereira, F. C. & Shlessarenko, R. D. Molecular detection of Mayaro virus during a dengue outbreak in the state of Mato Grosso, Central-West Brazil. *Mem Inst Oswaldo Cruz* **109**, 820-823, doi:10.1590/0074-0276140108 (2014).
- 89 da Silva Neves, N. A. *et al.* Chikungunya, Zika, Mayaro, and Equine Encephalitis virus detection in adult Culicinae from South Central Mato Grosso, Brazil, during the rainy season of 2018.

- Brazilian journal of microbiology* : [publication of the Brazilian Society for Microbiology] **53**, 63-70, doi:10.1007/s42770-021-00646-5 (2022).
- 90 Correa, J. F. *Avaliação sorológica e molecular de pacientes com suspeita de arboviroses em uma unidade municipal de saúde de Goiânia-Goiás*. 2020. Mestrado em Ciências Biológicas thesis, Universidade Federal de Goiás, (2020).
- 91 Saatkamp, C. J. *et al.* Mayaro virus detection in the western region of Pará state, Brazil. *Rev Soc Bras Med Trop* **54**, e0055-2020, doi:10.1590/0037-8682-0055-2020 (2021).
- 92 Casseb, A. d. R. *et al.* Prevalence of antibodies to equine alphaviruses in the State of Pará, Brazil. *Arq. Inst. Biol.* **83**, doi:10.1590/1808-1657000202014 (2016).
- 93 Azevedo, R. S. *et al.* Mayaro fever virus, Brazilian amazon. *Emerg Infect Dis* **15**, 1830, doi:10.3201/eid1511.090461 (2009).
- 94 Meneses, C. A. R. *Identificação de arbovírus, mayaro e oropouche em amostras com dengue negativo NS1 no estado de Roraima, no ano de 2012* Mestrado em Ciências da Saúde thesis, Universidade Federal de Roraima, (2013).
- 95 Tavares-Neto, J. *et al.* Pesquisa de anticorpos para arbovírus no soro de residentes no povoado de Corte de Pedra, Valença, Bahia. *Memórias do Instituto Oswaldo Cruz* **81**, 351-358, doi:10.1590/S0074-02761986000400001 (1986).
- 96 Moreira-Soto, A. *et al.* Limited Evidence for Infection of Urban and Peri-urban Nonhuman Primates with Zika and Chikungunya Viruses in Brazil. *mSphere* **3**, doi:10.1128/mSphere.00523-17 (2018).
- 97 Catenacci, L. S. *Abordagem one health para vigilância de arbovirus na Mata Atlântica do sul da Bahia, Brasil*. Doutorado em Virologia thesis, Instituto Evandro Chagas, (2017).
- 98 Barros, E. L. T. *Caracterização molecular de Chikungunya virus e investigação dos arbovírus Dengue virus e Mayaro virus no estado do Piauí* Mestre em Ciências Biomédicas thesis, Universidade Federal do Piauí, (2018).
- 99 Laroque, P. O. *et al.* Levantamento soropidemiológico para arbovírus em macaco-prego-galego (*Cebus flavius*) de vida livre no estado da Paraíba e em macaco-prego (*Cebus libidinosus*) de cativeiro do nordeste do Brasil. *Pesq Vet Bras* **34**, 462-468 (2014).
- 100 Coimbra, T. L. *et al.* Mayaro virus: imported cases of human infection in Sao Paulo State, Brazil. *Revista do Instituto de Medicina Tropical de Sao Paulo* **49**, 221-224, doi:10.1590/s0036-46652007000400005 (2007).
- 101 Paulo, M. B. *et al.* Serosurvey of Arbovirus in Free-Living Non-Human Primates (*Sapajus* Spp.) in Brazil. *Journal of Environmental Analytical Chemistry* **2** (2015).
- 102 da Silva, M. S. *Estudo de casos suspeitos de dengue negativos no teste sorológico para detecção do antígeno NS1: falha no diagnóstico ou emergência de outras arboviroses?* Mestrado em Imunologia Básica e Aplicada thesis, Universidade Federal do Amazonas, (2017).

- 103 Nunes, J. P. P. One-Step reverse transcriptase PCR for detection of arboviruses in serum samples of patients assisted in Basic health Units in the State of Maranhão, Brazil. *Braz J Dev* **5** (2019).
- 104 Dos Santos Souza Marinho, R. *et al.* Re-emergence of mayaro virus and coinfection with chikungunya during an outbreak in the state of Tocantins/Brazil. *BMC Res Notes* **15**, 271, doi:10.1186/s13104-022-06153-6 (2022).
- 105 da Silva-Nunes, M. The Acre Project: the epidemiology of malaria and arthropod-borne virus infections in a rural Amazonian population. *Cad. Saúde Pública* (2006).
- 106 Romeiro, M. F., Fumagalli, M. J., Dos Anjos, A. B. & Figueiredo, L. T. M. Serological evidence of Mayaro virus infection in blood donors from São Carlos, São Paulo, Brazil. *Trans R Soc Trop Med Hyg* **114**, 686-689, doi:10.1093/trstmh/traa016 (2020).
- 107 Souto, R. N. P. & Souto, R. C. C. Sorologia para arbovírus em população humana da região da Reserva Extrativista do Cajari, Amapá, Brasil. *Biota Amazônia* **2**, 8-14, doi:DOI: <http://dx.doi.org/10.18561/2179-5746/biotaamazonia.v2n2p8-14> (2012).
- 108 Niederman, J. C., Henderson, J. R., Opton, E. M., Black, F. L. & Skvrnova, K. A nationwide serum survey of Brazilian military recruits, 1964. II. Antibody patterns with arboviruses, polioviruses, measles and mumps. *Am J Epidemiol* **86**, 319-329 (1967).
- 109 Araujo, F. A. A. *et al.* Anticorpos antialfavírus detectados em equinos durante diferentes epizootias de encefalite equina, Paraíba, 2009. *Rev Bras Ciênc Vet* **19**, 80-85, doi:10.4322/rbcv.2014.086 (2012).
- 110 Araujo, F. A. A. *et al.* Vol. 1 (ed Ministério da Saúde Secretaria de Vigilância em Saúde) (Boletim Eletrônico Epidemiológico, , 2003).
- 111 Degallier, N. *et al.* Modifications of arbovirus transmission in relation to construction of dams in Brazilian Amazonia *Ciência e Cultura* **44**, 124-135 (1992).
- 112 Ziegler, M. F. *Mechanism that triggers the inflammatory process by Mayaro virus is discovered*, <https://agencia.fapesp.br/mechanism-that-triggers-the-inflammatory-process-during-infection-by-mayaro-virus-is-discovered/32066/> (2019).