1 Increased interregional virus exchange and nucleotide diversity outline the expansion of the chikungunya virus ECSA lineage in Brazil.

2

34567890 111234567 111234567

Joilson Xavier^{1,2}^, Luiz Alcantara^{1,2}^*, Vagner Fonseca³^, Mauricio Lima^{1,4}, Emerson Castro^{1,4}, Hegger Fritsch^{1,2}, Carla Oliveira⁵, Natalia Jouson Xavier¹⁻², Luiz Alcantara^{1-2,6}, Vagner Fonseca^{2,6}, Mauricio Lima^{1,7}, Emerson Castro^{1,1}, Hegger Fritsch^{1,-2}, Carla Oliveira², Natalia Guimarães⁴, Talita Adelino⁴, Mariane Evaristo⁶, Evandra S. Rodrigues⁶, Elaine Vieira Santos⁶, Debora de La-Roque⁶, Laise de Moraes⁷, Stephane Tosta^{1,2}, Adelino Neto⁸, Alexander Rosewell³, Ana Flavia Mendonça⁹, Anderson Leite¹⁰, Andreza Vasconcelos¹¹, Arabela L. Silva de Mello¹², Bergson Vasconcelos¹³, Camila A. Montalbano¹⁴, Camila Zanluca¹⁵, Carla Freitas¹⁶, Carlos F. C. de Albuquerque³, Claudia Nunes Duarte dos Santos¹⁵, Cleiton S. Santos⁷, Cliomar Alves dos Santos¹⁷, Crhistinne C. Maymone Gonçalves¹⁸, Dalane Teixeira¹³, Daniel F. L. Neto¹⁶, Diego Cabral¹¹, Elaine C. de Oliveira¹⁹, Ethel L. Noia Maciel²⁰, Felicidade Mota Pereira¹², Felipe Iani⁴, Fernanda P. de Carvalho¹¹, Gabriela Andrade¹², Gabriela Bezerra¹⁷, Gislene G. de Castro Lichs²¹, Glauco Carvalho Pereira⁴, Haline Barroso¹³, Helena Cristina Ferreira Franz¹⁶, Hivylla Ferreira²², Pare²³ Lines M. Biodre²⁴ Lashe B. Badriguel¹¹, Izadore C. de Silva¹¹, Line Mardera Bioz²⁴, Lavallo L. B. Dadriguel¹¹, Leiner C. de Silva¹¹, Leira Mardera Bioz²⁴, Lavallo L. B. Dadriguel¹¹, Leiner C. de Silva¹¹, Leira Mardera Bioz²⁴, Lavallo L. B. Dadriguel¹¹, Leira C. de Silva¹¹, Leira Mardera Bioz²⁴, Lavallo L. B. Dadriguel¹¹, Leira Mardera Bioz²⁴, Lavallo L. B. Badriguel¹¹, Leira C. de Silva¹¹, Leira Mardera Bioz²⁴, Lavallo L. B. Badriguel¹¹, Leira Mardel¹¹, Leira Mardera Bioz²⁴, Lavallo L. B. Badriguel¹¹, Leira Mardera Bioz²⁴, Lavallo Lavallo Lavallo Handriguel¹², Lavallo L Gabriela Bezerra', Gislene G. de Castro Lichs'', Glauco Carvalho Peretra', Haline Barroso¹⁵, Helena Cristina Ferretra Franz'', Hivylla Ferretra²⁺, Iago Gomes²³, Irina N. Riediger²⁴, Isabela Rodrigues¹⁷, Isadora C. de Siqueira⁷, Jacilane Silva¹¹, Jairo Mendez Rico²⁵, Jaqueline Lima¹², Jayra Abrantes²³, Jean Phellipe M. do Nascimento¹⁰, Judith N. Wasserheit²⁶, Julia Pastor¹¹, Jurandy J. F. de Magalhães^{11,27}, Kleber Giovanni Luz²⁸, Lidio G. Lima Neto²², Livia C. V. Frutuoso²⁹, Luana Barbosa da Silva¹⁹, Ludmila Sena¹⁷, Luis Arthur F. de Sousa²², Luiz Augusto Pereira³, Luiz Demarchi²¹, Magaly C. B. Câmara²³, Marcela G. Astet¹², Maria Almiron²⁵, Maricelia Lima³⁰, Marina C. S. Umaki Zardin²¹, Mayra M. Presibella²⁴, Melissa B. Falcão³¹, Michael Gale Jr.³², Naishe Freire¹¹, Nelson Marques²⁴, Noely F. O. de Moura²⁹, Pedro E. Almeida Da Silva²⁰, Peter Rabinowitz³³, Rivaldo V. da Cunha⁴⁴, Karen S. Trinta³⁴, Rodrigo F. do Carmo Said³, Rodrigo Kato¹⁶, Rodrigo Stabeli³, Ronaldo de Jesus¹⁶, Roselene Hans Santos¹¹, Simone K. Haddad⁶, Svetoslav N. Slavov^{6,35}, Tamires Andrade¹³, Themis Rocha²³, Thiago Carneiro¹³, Vanessa Nardy¹², Vinicius da Silva⁹, Walterlene G. Carvalho⁸, Wesley C. Van Voorhis³⁶, Wildo N. Araujo²⁵, Ana M.B. de Filippis⁵*, Marta Giovanetti^{12,27}*

18

- 1-Instituto Rene Rachou, Fundação Oswaldo Cruz, Minas Gerais, Brazil
- 2-Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais, Brazil
- 3-Organização Pan-Americana da Saúde, Organização Mundial da Saúde, Brazil
- 4-Laboratório Central de Saúde Pública de Minas Gerais, Fundação Ezequiel Dias, Brazil
- 5-Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil
- 6-Fundação Hemocentro de Ribeirão Preto, Brazil
- 7-Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Bahia, Brazil
- 8-Laboratório Central de Saúde Pública do Piaui, Brazil
- 9-Laboratório Central de Saúde Pública de Goias, Brazil
- 10-Laboratório Central de Saúde Pública de Alagoas, Brazil
- 11-Laboratório Central de Saúde Pública de Pernambuco, Brazil
- 12-Laboratório Central de Saúde Pública da Bahia, Brazil
- 13-Laboratório Central de Saúde Pública da Paraíba, Brazil
- 14-Universidade Federal de Mato Grosso do Sul
- 15-Instituto Carlos Chagas, Fundação Oswaldo Cruz, Paraná, Brazil
- 16-Coordenação Geral dos Laboratórios de Saúde Pública, Ministério da Saúde, Brazil
- 17-Laboratório Central de Saúde Pública de Sergipe, Brazil
- 18-Secretaria de Saúde do Estado do Mato Grosso do Sul, Brazil
- 19-Laboratório Central de Saúde Pública do Mato Grosso, Brazil
- 20-Secretaria de Vigilância em Saúde e Ambiente, Ministério da Saúde, Brazil
- 21-Laboratório Central de Saúde Pública do Mato Grosso do Sul, Brazil
- 22-Laboratório Central de Saúde Pública do Maranhão, Brazil
- 23-Laboratório Central de Saúde Pública do Rio Grande do Norte, Brazil
- 24-Laboratório Central de Saúde Pública do Paraná, Brazil
- 25-Pan American Health Organization, Washington, USA
- 26-Department of Global Health and Medicine, University of Washington, USA
- 27-Universidade de Pernambuco Campus Serra Talhada
- 28-Universidade Federal do Rio Grande do Norte 29-Coordenação Geral das Arboviroses, Ministério da Saúde, Brazil
- 30-Universidade Estadual de Feira de Santana, Brazil
- 31-Secretaria de Saúde de Feira de Santana, Feira de Santana, Bahia, Brazil 32-Department of Immunology, University of Washington, USA
- 33-Department of Environmental and Occupational Health Sciences, University of Washington, USA
- 34-Fundação Oswaldo Cruz, Instituto de Tecnologia em Imunobiológicos, Brazil
- 35-Center for Research Development, CDC, Butantan Institute, Brazil
- 36- Department of Medicine, University of Washington, USA
- 37-Sciences and Technologies for Sustainable Development and One Health, University of Campus Bio-Medico, Italy
- 57 ^These authors contributed equally
- 58 *Correspondence: giovanetti.marta@gmail.com, ana.bispo@ioc.fiocruz.br & luiz.alcantara@ioc.fiocruz.br

59 Abstract

60 The emergence and reemergence of mosquito-borne diseases in Brazil such as Yellow Fever, Zika, Chikungunya, and Dengue have had serious 61 62 63 65 66 67 68 impacts on public health. Concerns have been raised due to the rapid dissemination of the chikungunya virus (CHIKV) across the country since its first detection in 2014 in Northeast Brazil. Faced with this scenario, on-site training activities in genomic surveillance carried out in partnership with the National Network of Public Health Laboratories have led to the generation of 422 CHIKV genomes from 12 Brazilian states over the past two years (2021-2022), a period that has seen more than 312 thousand chikungunya fever cases reported in the country. These new genomes increased the amount of available data and allowed a more comprehensive characterization of the dispersion dynamics of the CHIKV East-Central-South-African (ECSA) lineage in Brazil. Tree branching patterns revealed the emergence and expansion of two distinct subclades. Phylogeographic analysis indicated that the northeast region has been the leading hub of virus spread towards other regions. Increased frequency of C>T transitions among the new genomes suggested that host restriction factors from the immune system such as ADAR and AID/APOBEC deaminases might be driving 69 CHIKV ECSA lineage genetic diversity in Brazil.

70 Keywords: chikungunya virus, East-Central-South-African lineage, phylogenetics, nanopore sequencing, genomic surveillance, genetic diversity.

71

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

72 Introduction

73 Mosquito-borne viral diseases have impacted the lives of millions of people across several populations in the tropics and subtropics^{1,2}. This scenario prompted the World Health 74 Organization (WHO) to issue a guideline in 2017 with strategies for global vector control 75 response aiming to reduce the burden and threat of vector-borne diseases, such as Dengue, 76 Zika and Chikungunya fever, by 2030^3 . Five years later, considering the public health 77 78 implications brought by the coronavirus disease (COVID-19) pandemic, the WHO outlined ten proposals to strengthen health emergency preparedness, response, and resilience⁴. One of 79 the proposals calls for the development and establishment of a collaborative surveillance 80 81 system with improved laboratory capacity for pathogen and genomic surveillance that would 82 guide the public health response.

83 Several genomic surveillance initiatives have been carried out in the last few years to build 84 knowledge regarding the genetic diversity and transmission dynamics of arboviruses in Brazil⁵⁻¹¹. Genomic sequencing has revealed that the first case of chikungunya fever 85 86 (CHIKF) reported in Brazil was an infection by the Asian lineage introduced in a northern 87 state in 2014, while another case reported seven days later in a north-eastern state represented 88 the first known introduction of the East-Central-South-African (ECSA) lineage in the 89 country¹². The establishment of the chikungunya virus (CHIKV) in the Brazilian territory was followed by several outbreaks reported across the country, accounting for more than 200 90 thousand confirmed cases in only the last two years¹³. Viral genomic data from Brazilian 91 cases has revealed that the ECSA lineage is widespread throughout the country and has been 92 93 linked to fatal cases observed in both risk and non-risk groups (young adults and no commodities)^{14,15}. 94

95 Viral genomic surveillance activities have been driven by the rapid development of DNA 96 sequencing technology and bioinformatics tools for genomic data analysis¹⁶. Such tools have 97 allowed the characterization of the genome and dispersal patterns of emerging and 98 reemerging pathogens^{6,9,17}. The use of such tools during the COVID-19 pandemic allowed, 99 for example, the rapid identification of emerging mutations likely associated with increased 100 transmissibility and immune escape¹⁸.

Despite technological advances and the high number of CHIKF cases reported in recent years
 in Brazil, the amount of genomic data available in public databases has consisted of genomes
 from localized outbreaks that could be limited in terms of representativeness across different

104 states and outbreak events. CHIKF can cause long last effects such as debilitating arthritis and arthralgia, and there are no effective treatments available¹⁹. Vaccine candidates in 105 106 development have reached phase 2 and 3 clinical trials using attenuated virus derived from 107 the Indian Ocean lineage (IOL) or virus-like particle containing recombinant structural proteins derived from a Senegalese viral strain^{20,21}. Since available vaccine candidates are 108 based on non-Brazilian variants, increasing the availability of genomic data to characterize 109 110 the genetic pool of the viral population circulating in Brazil might facilitate the future 111 development of an efficient and more representative CHIKV vaccine.

- 112 Recurring outbreaks demonstrate that CHIKV is currently endemic in Brazil. The existence 113 of abundant vectors, together with adequate climatic conditions for vector survival in areas of 114 high population density, create conditions that can modify the adaptive landscape, allowing the continued expansion and evolutionary adaptation of CHIKV²²⁻²⁴. Faced with a scenario of 115 limited availability of genetic information on potential strains causing a rapid increase in the 116 number of CHIKF cases over the past two years in Brazil, we carried out on-site training 117 118 activities in genomic surveillance in 12 Brazilian states covering four geographic regions to 119 increase the number of available viral genomic sequences. This has allowed comprehensive 120 monitoring of the expansion of the ECSA lineage and its variants circulating in different 121 states, in addition to the characterization of the most up-to-date structured phylogeny of the 122 CHIKV ECSA lineage in the country.
- 123

124 **Results**

East, Central, and South African (ECSA) lineage monitoring through countrywidegenomic surveillance

127 Nanopore sequencing was performed on selected CHIKV-positive samples provided by state 128 public health laboratories from 12 states across 4 geographic regions (Northeast, Midwest, 129 Southeast, and South) of Brazil (Figure 1A) during the years 2021-2022, which saw a 130 significant increase in the number of CHIKF cases reported across different Brazilian 131 regions, with a peak incidence rate of more than 20 cases per 100,000 population, and a total 132 of more than 312,000 reported cases nationally in this two-year period (Figure 1B). Due to 133 the portability and easy setup of the nanopore sequencing protocol, which allows data 134 generation in less than 24h, the collaborative work with the public health laboratories was 135 able to not only generate genomic data but also promote on-site genomic surveillance training activities for the local laboratory staff. This approach combining wet lab and basic data
analysis training allowed local teams to understand how genomic data can be linked to
demographic information in order to produce comprehensive and relevant inferences
regarding the epidemiology and evolution of CHIKV circulating in Brazil.

140 A total of 425 CHIKV-positive samples were subjected to nanopore whole-genome 141 sequencing, with 84.94% (n=361) of these samples originating in the Northeast region, 142 consisting largely of samples collected from the state of Bahia (n=102) (Figure 1A and 143 **Table 1**). These samples presented a mean RT-qPCR cycle threshold value of 24.04 (ranging 144 from 11 to 35.90) (Table 2). Patients' mean age upon sample collection was similar for both 145 females and males (39 years of age), with 57.88% (n=246) of the participants identified as 146 female (**Table 2**). The clinical status of patients at the time of sample collection, and travel 147 history data were not available for these samples.

148 Multiplex PCR-tiling amplicon sequencing on MinION allowed the recovery of 425 genomic 149 sequences from CHIKV with a mean genome coverage of 90.98% (range 31.80 to 96.19%) 150 (Table S1). Of the 425 sequences, 14 were recovered from old CHIKV-PCR samples 151 collected in Bahia state during July-August 2015 and stored since. These isolates had a mean 152 genome coverage of 70% (range 31.80 to 92.9 %). The remaining genomes have an 153 associated collection date ranging from April 2019 to June 2022. To better capture the 154 phylogenetic signal, only the sequences with genome coverage over 60% (n=422) were 155 considered for further analysis (discarded sequences are listed in the methods).

156 All the newly recovered genomes were assembled using Genome Detective software which 157 also classified all of them as belonging to the East, Central, and South African (ECSA) 158 lineage. To investigate the phylogenetic relationship of the new sequences with other 159 Brazilian and non-Brazilian sequences available in public databases, we built a global dataset 160 (n=1,987) composed of 1,565 CHIKV genomes retrieved from GenBank NCBI in addition to 161 422 sequences from this study. It is noteworthy that the two years of genomic surveillance 162 activities of this study contributed to a more than doubling of the number of CHIKV genomes 163 from Brazil available in the NCBI (by then there were 332 complete sequences) since the 164 virus emerged in the country in 2014 (Figure 2C).

165

166 Updated CHIKV phylogeny reveals two distinct emerging subclades

167 A preliminary Maximum Likelihood (ML) tree was reconstructed using the global dataset 168 that showed all Brazilian sequences grouped in the ECSA clade (Figure S1). It can also be 169 observed from the ML tree that most of the new sequences formed two well-distinct derived 170 clades, where it can be noticed that sequences collected in different geographical regions 171 were closely related (see ML tree in Supplemental files). To investigate in more detail the 172 phylogenetic features of these clades within a time-aware evolutionary framework, which can 173 benefit from the increased amount of genomic data obtained in this study, we performed a 174 Bayesian phylogenetic analysis using a down-sampled dataset (n=713) mostly composed of 175 Brazilian sequences.

176 Root-to-tip genetic distance regression indicated that the down-sampled dataset presented sufficient temporal signal ($R^2=0.70$ and correlation coefficient = 0.83) to infer a time-177 178 measured phylogeny (Figure 2B). Consistent with the ML tree, the inferred Maximum Clade 179 Credibility (MCC) tree also revealed two distinct more derived clades (henceforth clade I and 180 II) formed mainly by 2021-2022 sequences (Figure 2A). The Bayesian evolutionary analysis 181 estimated the time of the most recent common ancestor (tMRCA) of clade I to be late January 182 2018 (95% highest posterior density (HPD): December 2017 and March 2018), while clade II 183 presented a slightly late tMRCA estimated to be early February 2018 (95% HPD: January 184 2018 and March 2018).

185 Some composition differences can be noticed between these clades. Clade I comprises 186 sequences from 14 distinct states mostly collected from 2021 to 2022 (n=304) and from 187 northeastern Brazil (62.38%, n=204), with sequences from Sergipe (13.5%, Northeast), 188 Minas Gerais (1.8%), São Paulo (16.2%, Southeast), Goiás (11.6%), Mato Grosso do Sul 189 (0.6%, Midwest), and Paraná (4.3%, South) states uniquely present in this clade (Figure 2A). 190 Meanwhile, clade II is mostly composed of sequences collected in 2022 from northeastern 191 states (87.5%, n=147), with a total of 8 states sampled and Piaui state being the most 192 represented (46.42%, n=78) (Figure 2A).

A closer look at sequence distribution inside these clades reveals recurrent virus movement between states and regions, with midwestern isolates closely related to isolates from the Northeast and from the state of Minas Gerais (Southeast) in clade I. It can also be noticed in clade I that several distinct CHIKV introductions occurred into the state of Goiás (Midwest) and into northeastern states (Bahia, Rio Grande do Norte, Sergipe, Paraiba, Pernambuco, and Piaui) (**Figure 2A**). Contrarily, the clade I sequence distribution reveals that apparently, only one viral introduction event has happened in the southern state of Paraná, sharing a most
recent common ancestor (dated from Jun. 2019 to Jun. 2020, 95% HPD, posterior probability
= 1.0) with isolates from the state of São Paulo (Southeast cluster). Similarly, clade II
displays viral exchange between states, especially from the Northeast, as indicated by a single
well-supported subclade (posterior = 1.0), dated from Jan. 2020 to Dec. 2020 (95% HPD) and
dominated by sequences from northeastern states (Piauí, Bahia, Rio Grande do Norte,
Paraíba, Pernambuco, Maranhão, and Alagoas) (Figure 2A).

206

207 CHIKV dispersal in Brazil has been mainly seeded by the Northeast

208 In face of the recurrent virus movement observed across the country, as indicated by our 209 MCC phylogeny, we employed a Bayesian phylogeographic approach to reconstruct the 210 spatial dispersal dynamic of CHIKV in Brazil (closely related sequences from Haiti and 211 Paraguay were included) and to estimate the ancestral locations of clades I and II. The 212 resulting phylogeny kept the topology from the MCC tree shown in figure 2A and revealed 213 the Northeast as a leading source of CHIKV transmission in Brazil, seeding the network of 214 frequent virus exchange among states mainly from the Northeast, Southeast and Midwest, as 215 indicated by the location probability of the 5 early branching events inferred by the discrete 216 phylogeography (Figure 2D and Figure S2). Moreover, the ECSA lineage circulating in 217 Brazil extended its transmission network by reaching other countries in the region such as 218 Paraguay and Haiti likely via the Midwest and Northeast of Brazil, respectively. An 219 alternative approach, using a transmission network generated from transition states 220 summarized from the Bayesian phylogeography and centrality metrics, also indicated the 221 Northeast as a source (Source Hub Ratio of 0.66) in the network for the CHIKV spread in the 222 country, where intense interactions are displayed between the Northeast and Southeast 223 (Figure S3 and Table S2). Moreover, the discrete state ancestral reconstruction employed in 224 our Bayesian analysis estimated that both clades I and II might have emerged in the Southeast 225 region with a location probability of 1.0 and with estimated divergence times (clade I: 95% 226 HPD Nov. 2017 – Feb. 2018; clade II: Dec. 2017 – Mar. 2018) comparable to those obtained 227 in the MCC tree inferred using a comprehensive dataset displayed in Figure 2A (Figure S2).

These estimates place the divergence time of clades I and II in the period that marks the return of CHIKV-increased transmissions after two main epidemic seasons registered from 2016 to 2017 when a total of more than 565 thousand disease cases were notified in the 231 country (Figure 1B). From the time series graph of CHIKF cases, we can see an increase in 232 the incidence rate around early 2018 for the Midwest and North regions. In that same period, 233 clades I and II were estimated to emerge in the Southeast region, which also presented an 234 increased incidence rate. Since then, a seasonal epidemic pattern has been observed in the 235 CHIKV transmission dynamic with incidence peaks being displayed in the first semester of 236 the following years (Figure 1B). Consistent with the phylogeography and the transmission 237 network analyses, the CHIKF case time-series plot presents the Northeast region as a major 238 source of virus transmission in the country, as shown by the consecutive incidence peaks 239 registered for that region in the last three years.

240

241 Transition changes outline the evolutionary expansion of the Brazilian ECSA lineage

Despite having close divergence times, clades I and II have different sequence compositions that likely explain their apparent branching distance. While sequences from clades I and II were obtained from patients with similar median age, a significant difference (p<0.05) was observed between the median RT-qPCR cycle threshold value of the samples of each clade (**Figure 3A**). Such a difference, however, might have been affected by an imbalance present in the dataset used for comparison since clade I has almost twice as many sequences compared to clade II.

249 We also assessed these clades as to which selective regime they are likely to be subject to. 250 The results from the BUSTED analysis provided evidence that the envelope gene experienced 251 positive diversifying selection in both clades I (p=1.774e-8) and II (p=1.840e-11) (**Table S3**). 252 For comparison, we employed a second method, MEME, that identified 12 sites under 253 positive selection for clade I, whereas 13 sites were identified in clade II. Of these sites, 7 are 254 shared by both clades, while exclusive positive selected sites were reported for each clade 255 (Table S4). These results suggest the active status of these clades and consequently provide 256 evidence of the continuous evolutionary expansion of the Brazilian ECSA lineage.

Separate sequence alignments representative of each clade revealed a significant difference between the median frequency of 3 classes of single nucleotide variation (SNV) identified in the clades. Clade II presented a higher median frequency of SNV of type C>T, A>G, and T>C transitions (**Figure 3B**). By comparing the mutational profile of the two clades, we identified 27 SNVs exclusively present and shared by sequences from clade II, with three of them being non-synonymous substitutions leading putatively to amino acid change such as a

263 T288I substitution in the E1 protein gene (Table 3 and Figure 3C). The other two non-264 synonymous substitutions were identified in two nonstructural proteins (nsP2-P352A and 265 nsP4-A43V). Sequences from Clade I, on the other hand, presented 13 exclusive SNVs, of 266 which six are non-synonymous mutations present in two nonstructural proteins (nsP2 and 267 nsP4) and three structural proteins genes (capsid, 6k and E2) (**Table 3 and Figure 3C**). The 268 capsid gene sequence from this clade presented two contiguous transition substitutions in the 269 same codon (7787-7788) resulting in a Q74R change. In an alignment analysis of all 422 new 270 sequences, we found that among the total of CHIKV genome mutated positions the majority 271 consisted of type C>T transitions (29.5%, n= 368) followed by T>C substitutions (22.9%, n= 272 285) (Figure 3D). These transitions along with A>G and G>A transitions displayed a 273 comparable frequency distribution across the sequence dataset, where mutations at several 274 positions were identified in all new sequences (Figure 3E). Due to the possibility of these 275 substitutions resulting from sequencing errors, we performed the same comparative analysis 276 on a different dataset containing sequences that form a subclade (from the South region) in 277 clade I and that were generated by a different sequencing technology (Illumina) and a 278 different research group (data in Supplemental files). We observed the same pattern of 279 increased frequency (26.8%) of C>T transitions among those southern sequences, which 280 suggested the observed increased transition changes in clades I and II were likely derived 281 from mutation rather than sequencing artifacts.

282

283 Discussion

284 Advances in sequencing technology and bioinformatics tools have contributed to an increased 285 understanding of the genetic diversity and transmission dynamics of emerging viruses causing epidemics²⁵. In this study, we increased the number of CHIKV genomes deposited in 286 287 the NCBI by generating 422 genomes that cover 12 Brazilian states over the years 2021-288 2022. By combining genomic and demographic data this effort has provided a comprehensive 289 overview of CHIKV ECSA phylogeny in Brazil which summarizes and updates previous phylogenies from other studies of localized outbreaks ^{10,12,14,26,27}. The updated and metadata-290 291 enriched dataset allowed us not only to describe the genetic diversity of circulating CHIKV 292 variants but also to describe branching patterns observed in the updated phylogeny and infer 293 the divergence time and likely location of emerging distinct subclades from the Brazilian 294 lineage.

295 Following the arrival of ECSA lineage in the Northeast region, a number of other studies 296 have predicted and reported the establishment and spread of this lineage to the rest of the country ^{10–12,14,17,26–29}. The Bayesian phylogeographic approach employed in this study has 297 298 revealed the main routes of dispersion of CHIKV in the country. After being introduced in 299 the Northeast region, the ECSA lineage dispersed towards several states from all five regions 300 in Brazil. Both the Bayesian phylogeographic and transmission network approaches indicated 301 the Northeast region has been acting as the leading hub of virus spread towards other regions 302 in the country, including by forming an intense virus exchange network with the Southeast 303 region. From the Southeast, the virus dispersed to the southern state of Paraná (posterior 304 probability of 0.96) with the time of the most recent common ancestor shared with samples 305 from São Paulo estimated to be January 2020. In addition to spread in Brazil, CHIKV has 306 also extended its circulation to other countries such as Paraguay and Haiti. CHIKV 307 circulation in Paraguay and Haiti has been previously associated with transmission events 308 originating from Brazil through international viral exchange likely mediated by human movement ³⁰⁻³². These results evidence the lineage's potential expansion across Latin 309 310 America, despite limited information available about this lineage in the region as evidenced 311 by the lack of CHIKV ECSA sequences from other Latin American countries.

312 Our time-measured phylogeny corroborated previous studies and revealed emerging 313 branching patterns, mainly represented by two well-supported distinct subclades named clade 314 I and II. Both these clades were estimated to emerge in the Southeast region around the first 315 months of 2018. At the end of that year, the Southeast region accounted for more than 65 316 thousand reported cases. Moreover, differences in sequence composition were observed 317 between these clades, with clade I being more diverse as it comprises sequences from four 318 different regions (Northeast, Midwest, Southeast and South) collected in the years 2021-319 2022, while clade II contains mostly sequences from northeastern states collected in 2022. 320 This difference in the clades sequence composition profile might change as the lineage 321 continue to expand into the country. The observed differences in geographic diversity 322 between clades might reflect distinct transmission networks underlying the divergence and 323 expansion of these subclades.

Different viral lineages might be under distinct evolutionary pressure that together with the emergence and selection of mutations can drive viral adaption to a particular environment ^{33,34}. It has been argued that two different CHIKV lineages, IOL and Asian, have undergone different evolutionary trajectories leading to different vector adaptative potentials ³⁵. Here, we used the ratio of non-synonymous (dN) to synonymous (dS) nucleotide substitutions in the CHIKV envelope gene to assess the selective regime to which clades I and II might be subjected. Our analysis revealed that both clades have experienced positive diversifying selection. Such evolutionary pressure might be driven by the host antiviral immune response, as the viral envelope protein is a target for neutralizing antibodies ^{36,37}. Mutations in the envelope proteins have been implicated in the increased adaptation and transmission of the CHIKV IOL lineage in *Aedes albopictus* mosquitoes ^{22,38}.

A higher ratio of nonsynonymous substitutions is observed under a positive selection regime 335 promoted by virus-host interactions ³⁹. Since SNVs continue to arise in RNA virus 336 populations mainly driven by errors made by the virus replication complex that lead to 337 genetic diversity, we compared the mutational profiles of clades I and II⁴⁰. We identified 338 339 several SNVs across nonstructural and structural protein genes that were exclusive to each 340 clade. Clade II presented more SNVs (n=27) than clade I (n=13), of which three are non-341 synonymous substitutions (E1- T288I; nsP2-P352A, and nsP4-A43V). Literature research 342 revealed that E1-T288I change was previously identified in a 2017 sequence from Iran and also in CHIKV sequences collected in 2016 from infected cancer patients in Rio de Janeiro, 343 Brazil ^{41,42}. The nsP2-P352A substitution was also present in sequences collected between 344 2016 and 2017 in Rio de Janeiro^{14,43}. In turn, clade I sequences contain six non-synonymous 345 mutations across nonstructural (nsP2 and nsP4) and structural protein genes (capsid, 6k and 346 347 E2). The nsP4-V555I change was previously detected in sequences from Thailand in 2008-2009⁴⁴. The E2-L248F from clade I has been reported in Asian lineage sequences from 348 Colombia (2014-2015) and Philippines (2012)^{45,46}. Isolates from Thailand, Indonesia, Lao 349 PDR, Cameroon and India also presented the 6K-I54V mutation observed in clade I^{44,47-50}. 350 351 Despite the detection of these mutations in different countries (indicating homoplasy) by 352 other previous studies, there is no information about the functional impact of such 353 substitutions on CHIKV fitness, thus warranting further experimental studies to elucidate the 354 potential effects of SNVs on lineage-specific evolutionary adaptation. It has been argued that 355 not only non-synonymous mutations have the potential to promote adaptive changes but also 356 synonymous mutations can lead to changes in the viral RNA that can drive differential viral gene expression⁵¹. 357

Mutational analysis of the 422 sequences generated in this study revealed a higher amount of C>T and T>C transitions followed by A>G and G>A substitutions in the CHIKV genome and several genome positions presenting these transitions with higher frequency across all

361 new sequences. Moreover, Clade II has a significantly higher median frequency of C>T, 362 A>G, and T>C transitions compared to clade I. Although this study cannot experimentally 363 establish the significance of these transitions for CHIKV evolutionary adaptation, other 364 studies have associated this mutational pattern with the action of host antiviral immune repose mediated by AID/APOBEC and ADAR families of deaminases^{52,53}. These enzymes 365 are part of the interferon-stimulated innate immune response and promote viral genome 366 367 transitions mutations by catalyzing the deamination of adenosine to inosine to cause 368 A>G/T>C (by ADAR) substitutions or deamination of cytosine to uracil that leads to C>T/G>A (by AID/APOBEC) mutations⁵⁴. This RNA editing process has been 369 experimentally observed targeting specific viral sequences of SARS-CoV-2 to produce C>T 370 371 transitions and increasing viral replication in Caco-2 cells, thus promoting viral increased fitness and adaptative evolution⁵⁵. However, specific information about the effect of these 372 RNA editing processes on the CHIKV genome remains elusive, although the APOBEC3A 373 gene has been observed up-regulated in the expression profile of CHIKV-infected patients⁵⁶. 374 375 Despite arguments that these transitions happen mainly in phylogenetically uninformative 376 sites, the available evidence indicates that RNA editing processes might act as a significant 377 driver of viral sequence diversity and evolutionary adaptation through the introduction of nucleotide changes^{53,57}. 378

379 Recurring CHIKF epidemics, as indicated by the seasonal peak patterns displayed in the case 380 time-series plot, are evidence that the virus is endemic in Brazil. Human mobility, population 381 immunity, vector suitability, vegetation coverage, site socioeconomic status, and viral sequence variation are factors considered to mediate the dispersal of CHIKV in Brazil^{26,58,59}. 382 383 Although we did not find the E1:226V Aedes Albopictus-adaptive mutations in the Brazilian 384 sequences, the high abundance in the region of widely spread competent vectors, such as A. 385 aegypti and Ae. Albopictus, together with favorable climatic and social conditions in large 386 urban centers create conditions that modify the adaptive landscape of CHIKV, which in turn 387 can allow the continued expansion of the ECSA lineage in the country with a resultant increased impact on public health^{23,60,61}. Therefore, public health measures should be 388 389 undertaken to ensure continuous genomic surveillance of circulating CHIKV variants which 390 can help to identify viral transmission routes where focused vector control strategies could be 391 employed to reduce the risk of recurring CHIKF epidemics.

392

393 Limitations

394 Although our study presented the results of a Bayesian phylogeographic and mutational 395 profile analysis performed on 422 new sequences collected from 12 Brazilian states, not all 396 states were evenly represented in our dataset, which might limit our estimates relative to 397 divergence time and ancestral location reconstructions, prompting careful interpretation of 398 the results presented here. Ongoing sequencing efforts across the country could reduce this 399 disparity in the future. Moreover, although single nucleotide substitutions identified among 400 the new sequences offer insights into the evolutionary dynamics of CHIKV in Brazil, further 401 functional studies need to be undertaken to elucidate the actual adaptive effect of these 402 mutations.

403

404 Acknowledgements

405 This work was supported by The Pan American Health Organization (PAHO/WHO), the 406 Brazilian Ministry of Health grant SCON2021-00180 (Coordenação Geral de Laboratório de 407 Saúde Pública-CGLAB and Coordenação Geral de Vigilância de Arboviroses-CGARB), the 408 National Institutes of Health USA grant U01 AI151698 for the United World Arbovirus 409 Research Network (UWARN). We thank the State Health Secretariats and the 410 epidemiological surveillance services of Brazilian states and municipalities for all the support 411 provided. JX is supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-412 Brasil (CAPES)-Finance Code 001. MG is funded by PON "Ricerca e Innovazione" 2014-413 2020.

414

415 Author Contributions

416 Conception and Design: L.C.J.A., M.G., V.F., J.X.; Investigations: J.X., M.G., V.F.,

417 L.C.J.A., C.O., H.F., T.A., N.G., S.T., M.E., E.S., E.V.S., L.M., E.C., M.L., D.L.R., A.F.M.,

418 A.V., C.N.D.S., C.Z., C.S., D.C., F.C., G.B., H.B., I.G., I.R., I.C.S., J.B.S., G.A., J.L., J.A.,

419 J.P., K.G.L., L.B.S., L.A.F.S., M.G.A., M.L., M.B.F, M.C.S.U.Z., M.P., N.F., N.M., R.V.C.,

420 K.S.T., S.K.H., S.N.S., T.A., T.C., V.N., G.C.P., F.I., A.L.S.M., F.P., W.G.C., N.A.,

421 M.C.B.C., T.R., C.A.S., L.S., V.S., L.A.P., L.G.L.N., H.F., B.V., D.T., D.F.L.N, R.H.S.,

422 J.M., I.N.R., L.D., G.G.C.L., A.L., J.P.M.N., C.C.M.G., E.C.O., R.J., R.K., A.M.B.F.; Data

423 Curation: J.X., M.G., V.F., L.C.J.A; Formal Analysis: J.X., M.G., V.F., L.C.J.A.; Writing-

424	Original Draft Preparation: J.X., M.G., L.C.J.A.; Revision: J.X., M.G., V.F., L.C.J.A.;
425	Resources: J.M.R., M.A., W.N.A., A.R., R.F.C.S., C.F.C.A., W.C.V.V., P.R., M.G.J.,
426	J.N.W., R.S., E.L.N.M., P.EA.S., H.C.F.F., C.F., N.F.O.M., L.C.V.F., A.M.B.F., L.C.J.A.
427	
428	Declaration of interests
429	The authors declare no competing interests.
430	
431	Data and code availability
432	New sequences have been deposited in GenBank under accession numbers [available soon].
433	Input files used for the phylogenetic and mutational profile analyses will be available at the
434	GitHub repository https://github.com/genomicsurveillance.
435	
436	Figure legends

437 Figure 1. Spatiotemporal distribution of Chikungunya fever (CHIKF) in Brazil. A) Map of Brazil displays 438 states (colored) where samples were collected and sequenced. Bar plots indicate the number of isolates obtained 439 from each state. States abbreviations: AL=Alagoas, BA=Bahia, MA=Maranhão, PB=Paraíba, PE=Pernambuco, 440 PI=Piauí, RN=Rio Grande do Norte, SE=Sergipe, MG=Minas Gerais, PR=Paraná, GO=Goiás, MS=Mato 441 Grosso do Sul. B) Time series of monthly reported CHIKF cases normalized per 100 K individuals in five 442 Brazilian macroregions over 2014-2022 (until epidemiological week 28). Epidemic curves are colored according 443 to geographical macroregion: N = North, N = Northeast, M = Midwest, S = Southeast, S = Southeast, M = Northeast, M = Northeast, M = Northeast, M = Northeast, S = Northeast, M = Northeast, N = Northeast, 444 S = S outh. The shaded rectangle indicates the period in which samples were collected for this study.

445

446 Figure 2. Time-measured phylogeny of CHIKV ECSA lineage in Brazil. A) Maximum Clade Credibility 447 tree reconstructed using 706 sequences from Brazil (in addition to 5 sequences from Paraguay and 2 from Haiti) 448 and a molecular clock approach. Numbers in black show clade posterior probabilities of main nodes. Some 449 posterior probability values were omitted for clarity. Tip colors represent the sampling location. B) Root-to-tip 450 genetic distance regression in a maximum likelihood phylogeny of the CHIKV ECSA lineage (n=713). New sequences are colored in black. C) Number of genomes generated in this study (with >70% genome coverage) 451 compared to the number of Brazilian CHIKV-ESCA sequences available on the GenBank up to 27th Jan. 2023. 452 453 D). The spatial spread of CHIKV in Brazil estimated under a discrete diffusion model employed in the Bayesian 454 Phylogeographic approach using a dataset with 471 sequences. Size of colored circles was scaled by location 455 posterior support.

456

457 Figure 3. Genetic and demographic aspects of the newly generated CHIKV sequences. A) Boxplot of the 458 patient's age and samples' cycle threshold value distributions by clades I and II. Mann-Whitney U test with a 459 significance level alpha = 0.05. B) Boxplot of the frequency distributions for each class of single nucleotide 460 variation (SNV) identified in the genomes from clades I and II. Mann-Whitney U test with a significance level 461 alpha = 0.05. C) Genome positions and frequency of the non-synonymous substitutions uniquely identified in 462 sequences from clades I and II. D) Bar plot of the absolute number of CHIKV genome mutated positions for 463 each class of single nucleotide variation (SNV) identified in the alignment of the new CHIKV sequences 464 (n=422). E) Scatterplot of the frequency distributions for each class of single nucleotide variation (SNV) 465 identified in the alignment of the new CHIKV sequences (n=422). Each dot represents a distinct mutated 466 genome position.

467

468 Tables

469

Table 1: Number of Chikungunya virus-positive samplessequenced during the genomic surveillance in Brazil, 2021-2022, by geographical origin.

Region and State	# Samples (total n= 425)
Northeast	(n=361)
Alagoas (AL)	3 (0,71%)
Bahia (BA)	102 (24,00%)
Maranhão (MA)	38 (8,94%)
Paraíba (PB)	23 (5,41%)
Pernambuco (PE)	16 (3,77%)
Piauí (PI)	90 (21,18%)
Rio Grande do Norte (RN)	45 (10,59%)
Sergipe (SE)	44 (10,35%)
Midwest	(n=42)
Goiás (GO)	40 (9,41%)
Mato Grosso do Sul (MS)	2 (0,47%)
Southeast	
Minas Gerais (MG)	8 (1,88%)

medRxiv preprint doi: https://doi.org/10.1101/2023.03.28.23287733; this version posted April 3, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license. South Paraná (PR) 14 (3,29%) 470 Table 2: Demographic and laboratory characteristics of Chikungunya virus-infected patients. Patients (n= 425) Sex Female 246 (57.88 %) Male 179 (42.12 %) Mean age on sample collection (SD) Female 39.10 ± 18.83 39.37 ± 22.03 Male

 3.47 ± 8.52

 24.04 ± 4.15

Median days of symptoms prior to sampling (SD)

RT-qPCR mean Ct (cycle threshold) value (SD)

*SD = Standard Deviation

472

471

Table 3. List o	of single nucleot	ide variation	s (SNV) unique	y identified	for clades	I and II.		
Genomic Position ^a	Reference base	Altered base	Substitution type	aa ^b change	Clade	Genomic region	aa ^c position	Frequency ^d
3208	С	Т	S	D>D	1	nsP2	509	1.00
3217	Т	С	S	Y>Y	1	nsP2	512	1.00
3277	С	Т	S	S>S	1	nsP2	532	1.00
3295	G	А	S	P>P	1	nsP2	538	1.00
3314	G	Т	NS	A>S	1	nsP2	544	1.00
7328	G	Α	NS	V>I	1	nsP4	555	1.00
7787	Α	G	NS	Q>R	1	capsid	74	1.00
7788	G	Α	NS	Q>R	1	capsid	74	1.00
8641	С	Т	S	L>L	1	E2	34	0.68
9283	С	Т	NS	L>F	1	E2	248	0.22
9285	С	Т	S	L>L	1	E2	248	0.73
9936	Т	С	S	C>C	1	6K	42	1.00
9970	Α	G	NS	I>V	1	6K	54	1.00
1754	Т	С	S	L>L	2	NSP2	25	1.00
1759	Т	А	S	V>V	2	NSP2	26	1.00
1783	Т	G	S	R>R	2	NSP2	34	1.00
1789	А	G	S	Q>Q	2	NSP2	36	1.00
1795	Т	С	S	L>L	2	NSP2	38	1.00
2735	С	G	NS	P>A	2	NSP2	352	1.00
2767	А	G	S	G>G	2	NSP2	362	1.00
2887	Т	С	S	L>L	2	NSP2	402	1.00
3403	Т	С	S	F>F	2	NSP2	574	1.00

3901	А	G	S	V>V	2	NSP2	740	1.00	
3916	Т	С	S	F>F	2	NSP2	745	1.00	
4711	Т	С	S	D>D	2	NSP3	212	1.00	
4837	А	G	S	S>S	2	NSP3	254	1.00	
4843	С	Т	S	P>P	2	NSP3	256	1.00	
4939	Т	С	S	S>S	2	NSP3	288	1.00	
5783	С	Т	S	L>L	2	NSP4	40	1.00	
5793	С	Т	NS	A>V	2	NSP4	43	1.00	
5807	С	Т	S	L>L	2	NSP4	48	1.00	
5857	Т	С	S	Y>Y	2	NSP4	64	1.00	
5921	Т	С	S	L>L	2	NSP4	86	1.00	
5956	С	Т	S	T>T	2	NSP4	97	1.00	
6994	Т	С	S	N>N	2	NSP4	443	1.00	
7213	А	G	S	T>T	2	NSP4	516	1.00	
9153	Т	С	S	G>G	2	E2	204	0.89	
9642	Т	А	S	T>T	2	E2	367	0.80	
10856	С	Т	NS	T>I	2	E1	288	0.86	
11241	Т	С	S	G>G	2	E1	416	0.81	
a	Genomic pos	sition accord	ling to the re	ference sequ	ence NC	C 004162.2.			

473 ^aGenomic position accord
474 ^bAmino acid change.

475 ^cAmino acid position in the respective peptide.

476 ^dFrequency of the substitution in the dataset from the respective clade.

477 Non-synonymous substitutions are highlighted in bold.

478

479

480

481 Methods

482 Sample collection

483 Residual samples (serum or plasma) were obtained from the epidemiological surveillance 484 routine of the Brazilian Central Public Health Laboratories (LACEN) from different states 485 (Alagoas, Bahia, Goiás, Paraíba, Paraná, Pernambuco, Piauí, Maranhão, Minas Gerais, Mato 486 Grosso do Sul, Rio Grande do Norte, and Sergipe). These samples were submitted to nucleic 487 acid purification using the MagMax Viral RNA Mini kit (Thermo Fischer Scientific), 488 following the manufacturer's recommendations, and were previously screened by each 489 LACEN. CHIKV RT-qPCR positive samples were selected for sequencing based on the cycle 490 threshold value ≤ 30 and the availability of demographic metadata such as sex, age, and 491 municipality of residency. These demographic patient data were provided by LACENs and 492 were collected through a questionnaire filled out at local health care services.

493 Ethical statement

494 The project was approved by the Pan American World Health Organization (PAHO) and the 495 Brazilian Ministry of Health (MoH) as part of the arboviral genomic surveillance efforts within the terms of Resolution 510/2016 of CONEP (Comissão Nacional de Ética em 496 497 Pesquisa, Ministério da Saúde; National Ethical Committee for Research, Ministry of 498 Health). This authorizes the use of clinical samples collected in the Brazilian Central Public 499 Health Laboratories to accelerate knowledge building and contribute to surveillance and 500 outbreak response. The study protocol was reviewed and approved by Research Ethics 501 Committee of the Universidade Federal de Minas Gerais with approval No. 502 32912820.6.1001.5149. Personally identifying information were de-identified in the 503 datasets and tables in a way that minimizes the risk of unintended disclosure of 504 identity of individuals and information about them.

505

506 cDNA synthesis and whole genome sequencing

507 Extracted RNA from positive CHIKV samples were provided by collaborating LACENs and 508 submitted to cDNA synthesis and PCR, using a sequencing protocol based on multiplex PCR tiling amplicon approach design for MinION nanopore sequencing⁶². All reactions were 509 performed at biosafety level 2 facilities and using no template controls. PCR products were 510 511 purified using 1x AMpure beads Beckman Coulter, UK) and quantified using Qubit 3.0 512 instrument (Life Technologies) and the Qubit dsDNA High Sensitivity assay. DNA library 513 preparation was performed on all amplified samples using the Ligation Sequencing Kit 514 (Oxford Nanopore Technologies). Individual samples were barcoded using the Native 515 Barcoding Kit (NBD104, Oxford Nanopore Technologies, Oxford, UK). Sequencing library 516 was loaded onto a R9.4 flow cell and data were collected for up to 48 sequencing hours.

517

518 Generation of consensus sequences

519 Basecalling of raw FAST5 files and demultiplex of barcodes were performed using the 520 software Guppy (<u>https://github.com/nanoporetech</u>). Consensus sequences were generated by 521 a hybrid assembling approach implemented on Genome Detective 522 (<u>https://www.genomedetective.com/</u>)⁶³.

523

524 **Phylogenetic reconstruction**

We used MAFFT to align 422 new sequences (with coverage over >60% according to Thézé 525 et al. ⁶⁶, samples 736.22_RED, FS0116, and FS0132 were discarded) in addition to 1,565 526 527 CHIKV whole genome sequences publicly available in NCBI up to August 2022, forming a 528 global dataset (n=1,987) that includes all lineages. This global dataset was used to infer a Maximum Likelihood (ML) phylogeny using the IQ-TREE 2.1.1 software^{64,65}. Statistical 529 530 support for tree nodes was estimated using the ultrafast bootstrap (UFBoot) 531 feature implemented in IQ-TREE with 1,000 replicates. We then used the ML tree from the 532 global dataset to extract the Brazilian ECSA clade and use it to form a second dataset (total 533 n=713, 706 Brazilian sequences, 2 from Haiti, and 5 from Paraguay; sequences with genome coverage >60% according to Thézé et al. ⁶⁶) which was used to infer a time-scaled phylogeny 534 using BEAST v1.10.4. First, we investigated the temporal signal regressing root-to-tip 535 genetic distances from this ML tree against sample collection dates using TempEst v.1.5.1⁶⁷. 536 537 Secondly, we employed a stringent model selection analysis using both path-sampling (PS) 538 and stepping-stone (SS) procedures to estimate the most appropriate molecular clock model for the Bayesian phylogenetic analysis⁶⁸. For the Bayesian analysis, the uncorrelated relaxed 539 540 molecular clock was chosen as indicated by estimating marginal likelihoods, also employing 541 the HKY+G4 nucleotide substitution model, and the nonparametric Bayesian Skyline coalescent model. We combined two independent runs of 200 million states each⁶⁹. The 542 convergence of MCMC chains was checked using Tracer⁷⁰. Maximum clade credibility 543 544 (MCC) trees were summarized using TreeAnnotator after discarding 10% as burn-in.

545 CHIKV ECSA lineage movements across Brazil were investigated using the Bayesian 546 phylogeographic approach with a discrete trait phylogenetic model. A trait file was used to 547 discretize sequences sapling location by five Brazilian regions (North, Northeast, Southeast, 548 South and Midwest). For this analysis, we downsampled our Brazilian ECSA clade to a 549 dataset containing 471 sequences to maximize the temporal signal in the dataset. MCMC 550 analyses were performed in BEAST v1.10.4, running in duplicate for 200 million interactions 551 and sampling every 20,000 steps in the chain. Convergence for each run was assessed in 552 Tracer. MCC trees for each run were summarized using TreeAnnotator after discarding the initial 10% as burn-in. Finally, we used SPREAD 4 tool to map spatiotemporal information 553 embedded in the MCC trees 71 . 554

555

556 Transmission network analysis

18 pg.

A transmission network was reconstructed, using the StrainHub tool, from transition states summarized from the Bayesian phylogeography⁷². Centrality metrics on the tree nodes were also estimated for the network.

560

561 **Comparative mutational analysis**

For comparative mutational analysis, we assembled separate alignment datasets for each subclade (clade I = 327; clade II = 168) and for all new sequences (n=422). The sequence datasets were compared against the NCBI reference strain NC_004162.2 using MALVIRUS⁷³. We filtered and selected substitutions with an occurrence frequency of 100% across the whole dataset and substitutions with a frequency above 60% across the envelope genes E1 and E2.

568

569 Selection pressure analysis

570 Since the ratio of non-synonymous (dN) to synonymous (dS) nucleotide substitutions can be 571 used to study selection pressure on genomic sequences, we used the HYPHY software 572 package that employs statistical methods that estimate the dN/dS to detect diversifying selection⁷⁴. For that we performed an alignment of the envelope gene sequences of 720 573 574 Brazilian isolates and used BUSTED (restricting the analysis to each subclade I or II), an alignment-wide method implemented in HYPHY, that aims to detect evidence of episodic 575 diversifying selection⁷⁵. For comparison, we also used a different method, called MEME, a 576 site-level approach that aims to detect evidence of both pervasive and episodic diversifying 577 578 selection at individual sites (also restricting the analysis to each subclade I or II) 76 .

579

580 Epidemic curves from Chikungunya cases reported in Brazil

Data of weekly notified and laboratory-confirmed cases of infection by CHIKV in Brazil
from 2014 to 2022 were supplied by the Brazilian Ministry of Health (BrMoH)(Ministério da

583 Saúde, 2022). These data were used to calculate incidence and to plot time series charts.

584

585

586 **References**

19 pg.

medRxiv preprint doi: https://doi.org/10.1101/2023.03.28.23287733; this version posted April 3, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International licer 587 588 589 1. Mourad, O., Makhani, L., and Chen, L.H. (2022). Chikungunya: an emerging public health concern. Curr. Infect. Dis. Rep. 24, 217-228. 10.1007/s11908-022-00789-y. 590 2. 591 Chala, B., and Hamde, F. (2021). Emerging and Re-emerging Vector-Borne Infectious Diseases and the Challenges for Control: A Review. Front. Public Health 9, 592 593 715759. 10.3389/fpubh.2021.715759. 594 3. World Health Organization (2017). Global vector control response 2017–2030 (World 595 Health Organization). 596 4. World Health Organization (2022). 10 proposals to build a safer world together – 597 Strengthening the Global Architecture for Health Emergency Preparedness, Response 598 and Resilience (WHO Health Emergencies Programme). Adelino, T.É.R., Giovanetti, M., Fonseca, V., Xavier, J., de Abreu, Á.S., do 599 5. 600 Nascimento, V.A., Demarchi, L.H.F., Oliveira, M.A.A., da Silva, V.L., de Mello, A.L.E.S., et al. (2021). Field and classroom initiatives for portable sequence-based 601 602 monitoring of dengue virus in Brazil. Nat. Commun. 12, 2296. 10.1038/s41467-021-603 22607-0. 604 6. Faria, N.R., Quick, J., Claro, I.M., Thézé, J., de Jesus, J.G., Giovanetti, M., Kraemer, 605 M.U.G., Hill, S.C., Black, A., da Costa, A.C., et al. (2017). Establishment and cryptic 606 transmission of Zika virus in Brazil and the Americas. Nature 546, 406–410. 10.1038/nature22401. 607 Costa, É.A., Giovanetti, M., Silva Catenacci, L., Fonseca, V., Aburjaile, F.F., 608 7. 609 Chalhoub, F.L.L., Xavier, J., Campos de Melo Iani, F., da Cunha E Silva Vieira, M.A., Freitas Henriques, D., et al. (2021). West nile virus in brazil. Pathogens 10. 610 611 10.3390/pathogens10070896. 612 8. Faria, N.R., Kraemer, M.U.G., Hill, S.C., Goes de Jesus, J., Aguiar, R.S., Iani, 613 F.C.M., Xavier, J., Quick, J., du Plessis, L., Dellicour, S., et al. (2018). Genomic and 614 epidemiological monitoring of yellow fever virus transmission potential. Science 361, 894-899. 10.1126/science.aat7115. 615 Giovanetti, M., Pereira, L.A., Adelino, T.É.R., Fonseca, V., Xavier, J., de Araújo 616 9. 617 Fabri, A., Slavov, S.N., da Silva Lemos, P., de Almeida Marques, W., Kashima, S., et al. (2022). A retrospective overview of zika virus evolution in the midwest of brazil. 618 619 Microbiol. Spectr. 10, e0015522. 10.1128/spectrum.00155-22. Naveca, F.G., Claro, I., Giovanetti, M., de Jesus, J.G., Xavier, J., Iani, F.C. de M., do 620 10. Nascimento, V.A., de Souza, V.C., Silveira, P.P., Lourenço, J., et al. (2019). 621 Genomic, epidemiological and digital surveillance of Chikungunya virus in the 622 Brazilian Amazon. PLoS Negl. Trop. Dis. 13, e0007065. 623 624 10.1371/journal.pntd.0007065.

625 626 627 628 629	11.	Xavier, J., Giovanetti, M., Fonseca, V., Thézé, J., Gräf, T., Fabri, A., Goes de Jesus, J., Lima de Mendonça, M.C., Damasceno Dos Santos Rodrigues, C., Mares-Guia, M.A., et al. (2019). Circulation of chikungunya virus East/Central/South African lineage in Rio de Janeiro, Brazil. PLoS ONE <i>14</i> , e0217871. 10.1371/journal.pone.0217871.
630 631 632 633	12.	Nunes, M.R.T., Faria, N.R., de Vasconcelos, J.M., Golding, N., Kraemer, M.U.G., de Oliveira, L.F., Azevedo, R. do S. da S., da Silva, D.E.A., da Silva, E.V.P., da Silva, S.P., et al. (2015). Emergence and potential for spread of Chikungunya virus in Brazil. BMC Med. <i>13</i> , 102. 10.1186/s12916-015-0348-x.
634 635 636	13.	Pan American Health Organization (2023). Cases of chikungunya virus disease. PAHO/WHO Data - Weekly Report. https://www3.paho.org/data/index.php/en/mnu- topics/chikv-en/550-chikv-weekly-en.html.
637 638 639 640 641	14.	Souza, T.M.L., Vieira, Y.R., Delatorre, E., Barbosa-Lima, G., Luiz, R.L.F., Vizzoni, A., Jain, K., Miranda, M.M., Bhuva, N., Gogarten, J.F., et al. (2019). Emergence of the East-Central-South-African genotype of Chikungunya virus in Brazil and the city of Rio de Janeiro may have occurred years before surveillance detection. Sci. Rep. <i>9</i> , 2760. 10.1038/s41598-019-39406-9.
642 643 644 645	15.	de Lima, S.T.S., de Souza, W.M., Cavalcante, J.W., da Silva Candido, D., Fumagalli, M.J., Carrera, JP., Simões Mello, L.M., De Carvalho Araújo, F.M., Cavalcante Ramalho, I.L., de Almeida Barreto, F.K., et al. (2021). Fatal outcome of chikungunya virus infection in brazil. Clin. Infect. Dis. <i>73</i> , e2436–e2443. 10.1093/cid/ciaa1038.
646 647 648	16.	Robishaw, J.D., Alter, S.M., Solano, J.J., Shih, R.D., DeMets, D.L., Maki, D.G., and Hennekens, C.H. (2021). Genomic surveillance to combat COVID-19: challenges and opportunities. Lancet Microbe 2, e481–e484. 10.1016/S2666-5247(21)00121-X.
649 650 651 652	17.	Xavier, J., Fonseca, V., Bezerra, J.F., do Monte Alves, M., Mares-Guia, M.A., Claro, I.M., de Jesus, R., Adelino, T., Araújo, E., Cavalcante, K.R.L.J., et al. (2021). Chikungunya virus ECSA lineage reintroduction in the northeasternmost region of Brazil. Int. J. Infect. Dis. <i>105</i> , 120–123. 10.1016/j.ijid.2021.01.026.
653 654 655 656	18.	Garcia-Beltran, W.F., Lam, E.C., St Denis, K., Nitido, A.D., Garcia, Z.H., Hauser, B.M., Feldman, J., Pavlovic, M.N., Gregory, D.J., Poznansky, M.C., et al. (2021). Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. Cell <i>184</i> , 2372-2383.e9. 10.1016/j.cell.2021.03.013.
657 658	19.	Powers, A.M. (2018). Vaccine and therapeutic options to control chikungunya virus. Clin. Microbiol. Rev. <i>31</i> . 10.1128/CMR.00104-16.
659 660 661 662 663	20.	ClinicalTrials.gov (2022). Pivotal Study to Evaluate Safety and Immunogenicity of a Live-Attenuated Chikungunya Virus Vaccine Candidate in Adults - Full Text View - ClinicalTrials.gov. ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT04546724?term=NCT04546724&draw=2&ran k=1.

664 665 666 667 668	21.	Bennett, S.R., McCarty, J.M., Ramanathan, R., Mendy, J., Richardson, J.S., Smith, J., Alexander, J., Ledgerwood, J.E., de Lame, PA., Royalty Tredo, S., et al. (2022). Safety and immunogenicity of PXVX0317, an aluminium hydroxide-adjuvanted chikungunya virus-like particle vaccine: a randomised, double-blind, parallel-group, phase 2 trial. Lancet Infect. Dis. <i>22</i> , 1343–1355. 10.1016/S1473-3099(22)00226-2.
669 670 671 672 673	22.	Tsetsarkin, K.A., Chen, R., Yun, R., Rossi, S.L., Plante, K.S., Guerbois, M., Forrester, N., Perng, G.C., Sreekumar, E., Leal, G., et al. (2014). Multi-peaked adaptive landscape for chikungunya virus evolution predicts continued fitness optimization in Aedes albopictus mosquitoes. Nat. Commun. <i>5</i> , 4084. 10.1038/ncomms5084.
674 675 676 677	23.	Pancetti, F.G.M., Honório, N.A., Urbinatti, P.R., and Lima-Camara, T.N. (2015). Twenty-eight years of Aedes albopictus in Brazil: a rationale to maintain active entomological and epidemiological surveillance. Rev. Soc. Bras. Med. Trop. <i>48</i> , 87– 89. 10.1590/0037-8682-0155-2014.
678 679 680 681	24.	Carvalho, R.G., Lourenço-de-Oliveira, R., and Braga, I.A. (2014). Updating the geographical distribution and frequency of Aedes albopictus in Brazil with remarks regarding its range in the Americas. Mem. Inst. Oswaldo Cruz <i>109</i> , 787–796. 10.1590/0074-0276140304.
682 683 684	25.	Grubaugh, N.D., Ladner, J.T., Lemey, P., Pybus, O.G., Rambaut, A., Holmes, E.C., and Andersen, K.G. (2019). Tracking virus outbreaks in the twenty-first century. Nat. Microbiol. <i>4</i> , 10–19. 10.1038/s41564-018-0296-2.
685 686 687 688 689	26.	Charlys da Costa, A., Thézé, J., Komninakis, S.C.V., Sanz-Duro, R.L., Felinto, M.R.L., Moura, L.C.C., Barroso, I.M. de O., Santos, L.E.C., Nunes, M.A. de L., Moura, A.A., et al. (2017). Spread of chikungunya virus east/central/south african genotype in northeast brazil. Emerging Infect. Dis. <i>23</i> , 1742–1744. 10.3201/eid2310.170307.
690 691 692 693 694	27.	Maeda, A.Y., Nogueira, J.S., Campos, K.R., Camargo, C.H., da Silva Vasami, F.G., Arvigo, A.P.B., Santos, M.B.N., Abbud, A., and Sacchi, C.T. (2022). Circulation of Chikungunya virus East-Central-South African genotype during the 2020–21 outbreak in São Paulo State, Brazil. Journal of Clinical Virology Plus 2, 100070. 10.1016/j.jcvp.2022.100070.
695 696 697 698 699	28.	de Castro Moreira, D., Junior, F.F.D., Júnior, J.R.P., Jorge, F.A., Dos Santos Rando, F., Thomazella, M.V., Presibella, M.M., Riediger, I.N., Fernandez, M.A., and Bertolini, D.A. (2022). Genetic characterization of Chikungunya virus circulating in individuals from Paraná, Brazil. Braz. J. Microbiol. <i>53</i> , 641–645. 10.1007/s42770-022-00680-x.
700 701 702 703	29.	Teixeira, M.G., Andrade, A.M.S., Costa, M. da C.N., Castro, JN.S.M., Oliveira, F.L.S., Goes, C.S.B., Maia, M., Santana, E.B., Nunes, B.T.D., and Vasconcelos, P.F.C. (2015). East/Central/South African genotype chikungunya virus, Brazil, 2014. Emerging Infect. Dis. <i>21</i> , 906–907. 10.3201/eid2105.141727.

704 705 706 707	30.	White, S.K., Mavian, C., Salemi, M., Morris, J.G., Elbadry, M.A., Okech, B.A., Lednicky, J.A., and Dunford, J.C. (2018). A new "American" subgroup of African- lineage Chikungunya virus detected in and isolated from mosquitoes collected in Haiti, 2016. PLoS ONE <i>13</i> , e0196857. 10.1371/journal.pone.0196857.
708 709 710 711	31.	Gräf, T., Vazquez, C., Giovanetti, M., de Bruycker-Nogueira, F., Fonseca, V., Claro, I.M., de Jesus, J.G., Gómez, A., Xavier, J., de Mendonça, M.C.L., et al. (2021). Epidemiologic history and genetic diversity origins of chikungunya and dengue viruses, paraguay. Emerging Infect. Dis. 27, 1393–1404. 10.3201/eid2705.204244.
712 713 714 715 716	32.	de Oliveira, E.C., Fonseca, V., Xavier, J., Adelino, T., Morales Claro, I., Fabri, A., Marques Macario, E., Viniski, A.E., Campos Souza, C.L., Gomes da Costa, E.S., et al. (2021). Short report: Introduction of chikungunya virus ECSA genotype into the Brazilian Midwest and its dispersion through the Americas. PLoS Negl. Trop. Dis. <i>15</i> , e0009290. 10.1371/journal.pntd.0009290.
717 718 719 720	33.	Spielman, S.J., Weaver, S., Shank, S.D., Magalis, B.R., Li, M., and Kosakovsky Pond, S.L. (2019). Evolution of viral genomes: interplay between selection, recombination, and other forces. Methods Mol. Biol. <i>1910</i> , 427–468. 10.1007/978-1-4939-9074-0_14.
721 722 723	34.	Yu, X., and Cheng, G. (2022). Adaptive Evolution as a Driving Force of the Emergence and Re-Emergence of Mosquito-Borne Viral Diseases. Viruses <i>14</i> . 10.3390/v14020435.
724 725 726 727	35.	Chen, R., Plante, J.A., Plante, K.S., Yun, R., Shinde, D., Liu, J., Haller, S., Mukhopadhyay, S., and Weaver, S.C. (2021). Lineage Divergence and Vector- Specific Adaptation Have Driven Chikungunya Virus onto Multiple Adaptive Landscapes. MBio <i>12</i> , e0273821. 10.1128/mBio.02738-21.
728 729 730	36.	Hakim, M.S., Annisa, L., Gazali, F.M., and Aman, A.T. (2022). The origin and continuing adaptive evolution of chikungunya virus. Arch. Virol. <i>167</i> , 2443–2455. 10.1007/s00705-022-05570-z.
731 732 733 734 735	37.	Quiroz, J.A., Malonis, R.J., Thackray, L.B., Cohen, C.A., Pallesen, J., Jangra, R.K., Brown, R.S., Hofmann, D., Holtsberg, F.W., Shulenin, S., et al. (2019). Human monoclonal antibodies against chikungunya virus target multiple distinct epitopes in the E1 and E2 glycoproteins. PLoS Pathog. <i>15</i> , e1008061. 10.1371/journal.ppat.1008061.
736 737 738	38.	Tsetsarkin, K.A., Vanlandingham, D.L., McGee, C.E., and Higgs, S. (2007). A single mutation in chikungunya virus affects vector specificity and epidemic potential. PLoS Pathog. <i>3</i> , e201. 10.1371/journal.ppat.0030201.
739 740 741	39.	Duggal, N.K., and Emerman, M. (2012). Evolutionary conflicts between viruses and restriction factors shape immunity. Nat. Rev. Immunol. <i>12</i> , 687–695. 10.1038/nri3295.

742 743 744	40.	Collins, N.D., Beck, A.S., Widen, S.G., Wood, T.G., Higgs, S., and Barrett, A.D.T. (2018). Structural and nonstructural genes contribute to the genetic diversity of RNA viruses. MBio <i>9</i> . 10.1128/mBio.01871-18.
745 746 747 748	41.	Pouriayevali, M.H., Rezaei, F., Jalali, T., Baniasadi, V., Fazlalipour, M., Mostafavi, E., Khakifirouz, S., Mohammadi, T., Fereydooni, Z., Tavakoli, M., et al. (2019). Imported cases of Chikungunya virus in Iran. BMC Infect. Dis. <i>19</i> , 1004. 10.1186/s12879-019-4637-4.
749 750 751 752	42.	Familiar-Macedo, D., Gama, B.E., Emmel, V.E., Vera-Lozada, G., Abdelhay, E., Martins, I.S., and Hassan, R. (2022). Molecular aspects of Chikungunya virus infections in cancer patients. Mem. Inst. Oswaldo Cruz <i>117</i> , e210383. 10.1590/0074-02760210383.
753 754 755 756	43.	Cunha, M.S., Cruz, N.V.G., Schnellrath, L.C., Medaglia, M.L.G., Casotto, M.E., Albano, R.M., Costa, L.J., and Damaso, C.R. (2017). Autochthonous transmission of east/central/south african genotype chikungunya virus, brazil. Emerging Infect. Dis. 23, 1737–1739. 10.3201/eid2310.161855.
757 758 759	44.	Pongsiri, P., Auksornkitti, V., Theamboonlers, A., Luplertlop, N., Rianthavorn, P., and Poovorawan, Y. (2010). Entire genome characterization of Chikungunya virus from the 2008-2009 outbreaks in Thailand. Trop. Biomed. <i>27</i> , 167–176.
760 761 762 763	45.	Rodas, J.D., Kautz, T., Camacho, E., Paternina, L., Guzmán, H., Díaz, F.J., Blanco, P., Tesh, R., and Weaver, S.C. (2016). Genetic Characterization of Northwestern Colombian Chikungunya Virus Strains from the 2014-2015 Epidemic. Am. J. Trop. Med. Hyg. <i>95</i> , 639–646. 10.4269/ajtmh.16-0091.
764 765 766	46.	Tan, KK., Sy, A.K.D., Tandoc, A.O., Khoo, JJ., Sulaiman, S., Chang, LY., and AbuBakar, S. (2015). Independent emergence of the cosmopolitan asian chikungunya virus, philippines 2012. Sci. Rep. <i>5</i> , 12279. 10.1038/srep12279.
767 768 769	47.	Hakim, M.S., Annisa, L., and Aman, A.T. (2023). The evolution of chikungunya virus circulating in Indonesia: Sequence analysis of the orf2 gene encoding the viral structural proteins. Int. Microbiol. 10.1007/s10123-023-00337-1.
770 771 772 773	48.	Khan, N., Bhat, R., Jain, V., Raghavendhar B, S., Patel, A.K., Nayak, K., Chandele, A., Murali-Krishna, K., and Ray, P. (2021). Epidemiology and molecular characterization of chikungunya virus from human cases in North India, 2016. Microbiol. Immunol. <i>65</i> , 290–301. 10.1111/1348-0421.12869.
774 775 776 777	49.	Calvez, E., Bounmany, P., Somlor, S., Xaybounsou, T., Viengphouthong, S., Keosenhom, S., Brey, P.T., Lacoste, V., and Grandadam, M. (2022). Multiple chikungunya virus introductions in Lao PDR from 2014 to 2020. PLoS ONE <i>17</i> , e0271439. 10.1371/journal.pone.0271439.
778 779 780 781	50.	Agbodzi, B., Yousseu, F.B.S., Simo, F.B.N., Kumordjie, S., Yeboah, C., Mosore, M T., Bentil, R.E., Prieto, K., Colston, S.M., Attram, N., et al. (2021). Chikungunya viruses containing the A226V mutation detected retrospectively in Cameroon form a new geographical subclade. Int. J. Infect. Dis. <i>113</i> , 65–73. 10.1016/j.ijid.2021.09.058.

782 783 784	51.	Faure, G., Ogurtsov, A.Y., Shabalina, S.A., and Koonin, E.V. (2017). Adaptation of mRNA structure to control protein folding. RNA Biol. <i>14</i> , 1649–1654. 10.1080/15476286.2017.1349047.
785 786 787	52.	Chaw, SM., Tai, JH., Chen, SL., Hsieh, CH., Chang, SY., Yeh, SH., Yang, WS., Chen, PJ., and Wang, HY. (2020). The origin and underlying driving forces of the SARS-CoV-2 outbreak. J. Biomed. Sci. 27, 73. 10.1186/s12929-020-00665-8.
788 789 790 791	53.	Simmonds, P., and Ansari, M.A. (2021). Extensive C->U transition biases in the genomes of a wide range of mammalian RNA viruses; potential associations with transcriptional mutations, damage- or host-mediated editing of viral RNA. PLoS Pathog. <i>17</i> , e1009596. 10.1371/journal.ppat.1009596.
792 793 794	54.	Pecori, R., Di Giorgio, S., Paulo Lorenzo, J., and Nina Papavasiliou, F. (2022). Functions and consequences of AID/APOBEC-mediated DNA and RNA deamination. Nat. Rev. Genet. <i>23</i> , 505–518. 10.1038/s41576-022-00459-8.
795 796 797	55.	Kim, K., Calabrese, P., Wang, S., Qin, C., Rao, Y., Feng, P., and Chen, X.S. (2022). The roles of APOBEC-mediated RNA editing in SARS-CoV-2 mutations, replication and fitness. Sci. Rep. <i>12</i> , 14972. 10.1038/s41598-022-19067-x.
798 799 800 801	56.	Soares-Schanoski, A., Baptista Cruz, N., de Castro-Jorge, L.A., de Carvalho, R.V.H., Santos, C.A.D., Rós, N. da, Oliveira, Ú., Costa, D.D., Santos, C.L.S.D., Cunha, M.D.P., et al. (2019). Systems analysis of subjects acutely infected with the Chikungunya virus. PLoS Pathog. <i>15</i> , e1007880. 10.1371/journal.ppat.1007880.
802 803 804	57.	Piontkivska, H., Wales-McGrath, B., Miyamoto, M., and Wayne, M.L. (2021). ADAR Editing in Viruses: An Evolutionary Force to Reckon with. Genome Biol. Evol. 13. 10.1093/gbe/evab240.
805 806 807 808	58.	Churakov, M., Villabona-Arenas, C.J., Kraemer, M.U.G., Salje, H., and Cauchemez, S. (2019). Spatio-temporal dynamics of dengue in Brazil: Seasonal travelling waves and determinants of regional synchrony. PLoS Negl. Trop. Dis. <i>13</i> , e0007012. 10.1371/journal.pntd.0007012.
809 810 811 812	59.	Freitas, L.P., Schmidt, A.M., Cossich, W., Cruz, O.G., and Carvalho, M.S. (2021). Spatio-temporal modelling of the first Chikungunya epidemic in an intra-urban setting: The role of socioeconomic status, environment and temperature. PLoS Negl. Trop. Dis. <i>15</i> , e0009537. 10.1371/journal.pntd.0009537.
813 814 815 816 817	60.	de Melo Ximenes, M. de F.F., de Araújo Galvão, J.M., Inacio, C.L.S., Macêdo E Silva, V.P., Pereira, R.L. do N., Pinheiro, M.P.G., de Medeiros Silva, M.M., and Gomes, C.E.S. (2020). Arbovirus expansion: New species of culicids infected by the Chikungunya virus in an urban park of Brazil. Acta Trop. <i>209</i> , 105538. 10.1016/j.actatropica.2020.105538.
818 819 820 821	61.	Costa-da-Silva, A.L., Ioshino, R.S., Petersen, V., Lima, A.F., Cunha, M.D.P., Wiley, M.R., Ladner, J.T., Prieto, K., Palacios, G., Costa, D.D., et al. (2017). First report of naturally infected Aedes aegypti with chikungunya virus genotype ECSA in the Americas. PLoS Negl. Trop. Dis. <i>11</i> , e0005630. 10.1371/journal.pntd.0005630.

822 823 824 825	62.	Quick, J., Grubaugh, N.D., Pullan, S.T., Claro, I.M., Smith, A.D., Gangavarapu, K., Oliveira, G., Robles-Sikisaka, R., Rogers, T.F., Beutler, N.A., et al. (2017). Multiplex PCR method for MinION and Illumina sequencing of Zika and other virus genomes directly from clinical samples. Nat. Protoc. <i>12</i> , 1261–1276. 10.1038/nprot.2017.066.
826 827 828 829	63.	Vilsker, M., Moosa, Y., Nooij, S., Fonseca, V., Ghysens, Y., Dumon, K., Pauwels, R., Alcantara, L.C., Vanden Eynden, E., Vandamme, AM., et al. (2019). Genome Detective: an automated system for virus identification from high-throughput sequencing data. Bioinformatics <i>35</i> , 871–873. 10.1093/bioinformatics/bty695.
830 831 832	64.	Nguyen, LT., Schmidt, H.A., von Haeseler, A., and Minh, B.Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. <i>32</i> , 268–274. 10.1093/molbev/msu300.
833 834 835	65.	Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. <i>33</i> , 511–518. 10.1093/nar/gki198.
836 837 838 839	66.	Thézé, J., Li, T., du Plessis, L., Bouquet, J., Kraemer, M.U.G., Somasekar, S., Yu, G., de Cesare, M., Balmaseda, A., Kuan, G., et al. (2018). Genomic epidemiology reconstructs the introduction and spread of zika virus in central america and mexico. Cell Host Microbe <i>23</i> , 855-864.e7. 10.1016/j.chom.2018.04.017.
840 841 842	67.	Rambaut, A., Lam, T.T., Max Carvalho, L., and Pybus, O.G. (2016). Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2, vew007. 10.1093/ve/vew007.
843 844 845 846	68.	Baele, G., Lemey, P., Bedford, T., Rambaut, A., Suchard, M.A., and Alekseyenko, A.V. (2012). Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. Mol. Biol. Evol. <i>29</i> , 2157–2167. 10.1093/molbev/mss084.
847 848 849	69.	Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., and Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol. <i>4</i> , vey016. 10.1093/ve/vey016.
850 851 852	70.	Rambaut, A., Drummond, A.J., Xie, D., Baele, G., and Suchard, M.A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67, 901–904. 10.1093/sysbio/syy032.
853 854 855	71.	Nahata, K.D., Bielejec, F., Monetta, J., Dellicour, S., Rambaut, A., Suchard, M.A., Baele, G., and Lemey, P. (2022). SPREAD 4: online visualisation of pathogen phylogeographic reconstructions. Virus Evol. 8, veac088. 10.1093/ve/veac088.
856 857 858 859	72.	de Bernardi Schneider, A., Ford, C.T., Hostager, R., Williams, J., Cioce, M., Çatalyürek, Ü.V., Wertheim, J.O., and Janies, D. (2020). StrainHub: a phylogenetic tool to construct pathogen transmission networks. Bioinformatics <i>36</i> , 945–947. 10.1093/bioinformatics/btz646.

860 861 862	73.	Ciccolella, S., Denti, L., Bonizzoni, P., Della Vedova, G., Pirola, Y., and Previtali, M. (2022). MALVIRUS: an integrated application for viral variant analysis. BMC Bioinformatics <i>22</i> , 625. 10.1186/s12859-022-04668-0.
863 864 865 866	74.	Kosakovsky Pond, S.L., Poon, A.F.Y., Velazquez, R., Weaver, S., Hepler, N.L., Murrell, B., Shank, S.D., Magalis, B.R., Bouvier, D., Nekrutenko, A., et al. (2020). HyPhy 2.5-A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. Mol. Biol. Evol. <i>37</i> , 295–299. 10.1093/molbev/msz197.
867 868 869	75.	Murrell, B., Weaver, S., Smith, M.D., Wertheim, J.O., Murrell, S., Aylward, A., Eren, K., Pollner, T., Martin, D.P., Smith, D.M., et al. (2015). Gene-wide identification of episodic selection. Mol. Biol. Evol. <i>32</i> , 1365–1371. 10.1093/molbev/msv035.
870 871	76.	Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., and Kosakovsky Pond, S.L. (2012). Detecting individual sites subject to episodic diversifying selection.

PLoS Genet. 8, e1002764. 10.1371/journal.pgen.1002764.



25

って



Epidemiological week

Location: NE N S SE MW

CHIKV

CHIKV

