Reliable estimates of Japanese encephalitis virus evolutionary rate supported by a formal test of temporal signal

Polina I. Ukraintceva [a], Artem N. Bondaryuk [b,*], Rodion G. Sakhabeev [a]

[a] Saint-Petersburg State Institute of Technology;

[b] Irkutsk Antiplague Research Institute of Siberia and Far East;

[*] Correspondence at: Artem N. Bondaryuk. E-mail: ui.artem.ui@gmail.com

Abstract

Japanese encephalitis virus (JEV) is a positive-strand RNA virus, that causing encephalitis in humans, which is highly spread in East and Southeast Asia, Western Pacific countries and Australia. The main vector of JEV is *Culex* spp. (mainly *C. tritaeniorhynchus*). The intraspecies structure of JEV comprises five genotypes (GI-GV). In the previous studies, the substitution rate and the age of the JEV most recent common ancestor (MRCA) based on complete open reading frame (ORF) / whole genome sequences were reported, however, the estimates are significantly different between some of the studies. Most importantly, the temporal signal in nucleotide data as the key aspect of the molecular clock application was evaluated by neither formal nor informal tests. In February 2023, we found 325 non-recombinant nucleotide sequences of ORF with known collection dates. The purpose of our study is the evaluation of temporal signal in the JEV ORF nucleotide data set using the Bayesian approach with the subsequent re-estimation of substitution rate and the age of the JEV common clade (GI-GV) and two JEV genotypes (GI and GIII) characterized by the number of sequences enough for reliable estimates. Results showed the presence of temporal signal in all analyzed data set. The substitution rate of the JEV GI-GV was 2.41×10^{-4} (95% highest posterior density (HPD) interval, 1.67×10^{-4}–3.14×10^{-4}) or approximately 2.5 nucleotide substitutions per ORF (10,245 nucleotides) per year that is one the lowest substitution rate among the species closely related to JEV and other well-studied members of mosquito-borne flaviviruses. The substitution rate of

GI and GIII evaluated as 4.13×10^{-4} (95% HPD interval, 3.45×10^{-4}–4.82×10^{-4}) and 6.17×10^{-5} (95% HPD interval 3.84×10^{-5}–8.62×10^{-5}), accordingly, with the ages of clades of 153 (95% HPD, 87–237) and 216 (95% HPD, 139–317), respectively. The root height of JEV is 1234 years (95% HPD, 407–2333). To the best of our knowledge, this study provides the first reliable substitution rate assessment of the global clade of JEV embracing all five genotypes.

Keywords: substitution rate, molecular clock, Orthoflavivirus, phylogeny, RNA virus, evolution

1. Introduction

Japanese encephalitis virus belongs to the species *Orthoflavivirus japonicum,* family *Flaviviridae* (Postler et al., 2023). The JEV genome consists of single-stranded positive-sense RNA with a length of about 10,965 nucleotides (nt). The genome encompasses a 10,245 nt single open reading frame (ORF), encoding a polyprotein, which includes three structural proteins: capsid (C), prosecutor membrane (prM) and envelope (E) proteins, and seven non-structural ones: NS1, NS2a, NS2b, NS3, N4a, NS4b, NS5 (Xu et al., 2023), (Lindenbach, 2007).

Mosquitoes are the vectors for JEV, in particular, *Culex tritaeniorhynchus* and *Culex annulirostris*, which were recognized as principal ones (Van den Eynde et al., 2022). The virus is zoonotic, and it is transmitted in most cases between pigs or birds by mosquitoes, however, Ricklin et al. (2016) described the direct way of virus transmission between two pigs via oronasal secretions. Humans, cattle and horses are dead-end hosts, due to insufficient viremia to infect mosquitoes. When these hosts are infected, encephalitis can develop with fever, convulsions, coma and even death (Van den Eynde et al., 2022), (Ladreyt et al., 2019).

Approximately 3 billion people live in JEV epidemic areas according to the World Health Organization data (https://www.who.int/news-room/fact-sheets/detail/japanese-encephalitis). JEV is the major causative agent of human virus encephalitis disease in Asia. Now the virus is endemic in Australia and Asia,

59 including 24 countries, namely China, Korea, Indonesia, Malaysia, Japan and even
60 Pakistan (Van den Eynde et al., 2022), (Mackenzie, 2005). Additionally, virus was
61 registered outside the endemic areas in Italy (Ravanini et al., 2012) and Angola
62 (Simon-Loriere et al., 2017).

63 Today, between 30,000 (Van den Eynde et al., 2022) and 69,000 (Auerswald
64 et al., 2021) cases registered annually according to different sources, but it is
65 considered as underestimated values due to insufficient monitoring and diagnosis.
66 The annual loss is reckoned as 709,000 disability-adjusted life years (Auerswald et
67 al., 2021). According to (McGuinness et al., 2023) JEV causes 100, 000 cases and
68 25, 000 deaths annually. The recent systematic review (Cheng et al., 2022) reported
69 that the case fatality rate of JEV in 2000–2018 was 14% (95% credible interval, 11–
70 17) and demonstrated the trend to decline in comparison with the previous periods.

71 Despite the fact that effective and safe vaccines have been developed (Vannice
72 et al., 2021), JEV is still one of the significant agents causing viral encephalitis in
73 the territory of Asia. Thus, to monitor and control this disease, it is important to
74 scrutinize evolutionary driving forces, which determine virus spread and
75 pathogenicity.

76 Currently, the JEV substitution rate estimates and the time to the most recent
77 common ancestor (MRCA) of the virus differ between the studies. Moreover, it is
78 important to underline that none of the observed works have performed the
79 evaluation of temporal signal in datasets prior to molecular clock application, and,
80 therefore, the reliability of the reported estimates is questionable.

81 The aim of our study is the reliable assessment of substitution rates and ages
82 of JEV and its genotypes using the Bayesian phylogenetic approach. For this
83 purpose, we used 325 complete ORF nucleotide sequences of all five JEV genotypes
84 with known collection dates available at the start of a study (January 2023).
85 Evolutionary rate and divergence times obtained were compared with the previous
86 estimates of JEV evolutionary rate and the rate of virus species, closely related to
87 JEV, and other well-studied mosquito-borne flaviviruses. Bayesian evaluation of

88 temporal signal makes reported results the first reliable assessment of JEV
89 substitution rate.

90 **2. Materials and methods**

91 **2.1 Nucleotide dataset**

92 Initially, 440 nucleotide sequences of complete or near complete genomes of
93 JEV were found in GenBank using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi).
94 Clones, vaccine and chimeric strains were removed from the dataset and then
95 collection dates were identified for 367 strains. Recombinant sequences detected by
96 RDP5 (Martin et al., 2021) were removed from the alignment. Thus, our research
97 was conducted using 325 nucleotide sequences of a polyprotein gene.

98 Next, sequence alignment was done with MAFFT v.7.490 (Katoh and
99 Standley, 2013). After that, the alignment was visualized in AliView v.1.24 (Larsson,
100 2014) to exclude ORFs with stop-codons. Non-coding regions were cut off using the
101 NCBI reference sequences (NC_001437.1).

102 JEV genotypes were determined in preliminary phylogenetic analysis in IQ-
103 TREE v.1.6.12 using the tree topology reported by (Takhampunya et al., 2011). It
104 was revealed that only GI and GIII have a sufficient number of sequences to obtain
105 reliable statistics in the further independent Bayesian phylogenetic analyses.
106 Therefore, we subsequently analyzed three data sets — the common clade of JEV
107 (JEV GI-GV), GI and GIII.

108 **2.2 Temporal signal assessment**

109 Temporal structure in each data sets (JEV GI-GV, GI and GIII) was assessed
110 with the Bayesian evaluation of temporal signal (BETS) (Duchene et al., 2020) in
111 BEAST v.2.6.7 (Bouckaert et al., 2019).

112 BETS involves the comparison of time-stamped (heterochronous) and
113 isochronous models (M_{het} and M_{iso}) by the difference between log marginal
114 likelihoods. The log marginal likelihoods for each model were calculated using the
115 path sampling estimator implemented in the MODEL_SELECTION package from
116 the BEAST2 Package Manager. The number of steps was set as 200 and the lengths

117 of MCMC for each step were determined so that the product of the number of steps
118 and MCMC length equaled the length of the MCMC during the main phylogenetic
119 analysis in BEAST with an effective sample size (ESS) at least of 200. Priors on
120 model parameters (molecular clock, substitution model, population model) were
121 chosen similar to those used in the main analysis with the exception of the tree prior
122 — in this case, we replaced the Bayesian skyline with the coalescent constant
123 population with a proper prior (i.e., integrated to 1) on effective population size
124 (Table 1):
125

126 Table 1. Priors used in the Bayesian evaluation of temporal signal (BETS) in all
127 three data sets (JEV GI-GV, GI, GIII)

Model	Parameter		Prior
Strict clock	Rate		Uniform(0.0, 1.0)
Optimized relaxed clock	Rates		LogNormal(M = 1.0, S = 0.2)
	Sigma		Γ(α = 0.5396, β = 0.3819)
	Mean		Γ(α = 0.1, β = 100)
Rate heterogeneity	Shape of Γ distribution		Exp(M = 1.0)
	Proportion of invariant sites		Uniform(0.0, 1.0)
Substitution model	GTR rates	Transition (A-G)	Γ(α = 0.05, β = 20)
		Transversions	Γ(α = 0.05, β = 10)
Tree prior (coalescent constant size)	Effective population size		Exp(M = 100)

128

129 Besides, BETS allows for a simultaneous and coherent assessment of the
130 entire model, including the molecular clock. Therefore, we included a strict clock
131 (SC) and an optimized relaxed clock (ORC) models into the comparison (Douglas
132 et al., 2021).
133 The difference between log marginal likelihoods of competitive models (i.e.,
134 log Bayes factor) of at least 5.0 indicates very strong support, a value of 3 means
135 strong support and a value of 1 is considered to be positive evidence (Kass and
136 Raftery, 1995).
137 Also, for comparison purposes, we conducted widely used root-to-tip (RTT)
138 regression, which is an informal test with well-known limitations: data points are

not independent and strict clocklike behavior is assumed. RTT statistics were computed with Clockor2 (Featherstone et al., 2023). The trees for Clockor2 were reconstructed in IQ-TREE v.1.6.12 with the best-fit model determined for each data set using ModelFinder (Kalyaanamoorthy et al., 2017).

2.3 Dated phylogeny reconstruction and substitution rate evaluation

Bayesian phylogenetic analysis using JEV-common, GI and GIII data sets was performed in BEAST v.2.6.7. Selecting substitution model was based on the BIC previously assessed by ModelFinder in IQ-TREE. Molecular clock models were compared based on the log MLs calculated on the step of BETS. As a tree prior, we specified a flexible non-parametric Bayesian Skyline model. The lengths of MCMC in each run were chosen so that the ESS associated with estimated parameters reached at least 200 samples. The inspection of model parameters and MCMC convergence was carried out in Tracer v.1.7.1 (Rambaut et al., 2018). The burn-in thresholds of MCMC chains were defined individually depending on the chain behavior. Most clade credible (MCC) trees were annotated in TreeAnnotator (Bouckaert et al., 2019). Due to high computational cost, BEAST was run using a BEAGLE library (Ayres et al., 2012) on a GPU (NVIDIA GeForce RTX 3070).

3. Results and discussion

3.1 Degree of clocklike behavior of the common JEV dataset (GI-GV) and GI and GIII genotypes

In the first step, using three JEV data sets (GI-GV, GI, GIII) we conducted RTT regression to assess to which degree data behave in a clock-like fashion (Figure 1).

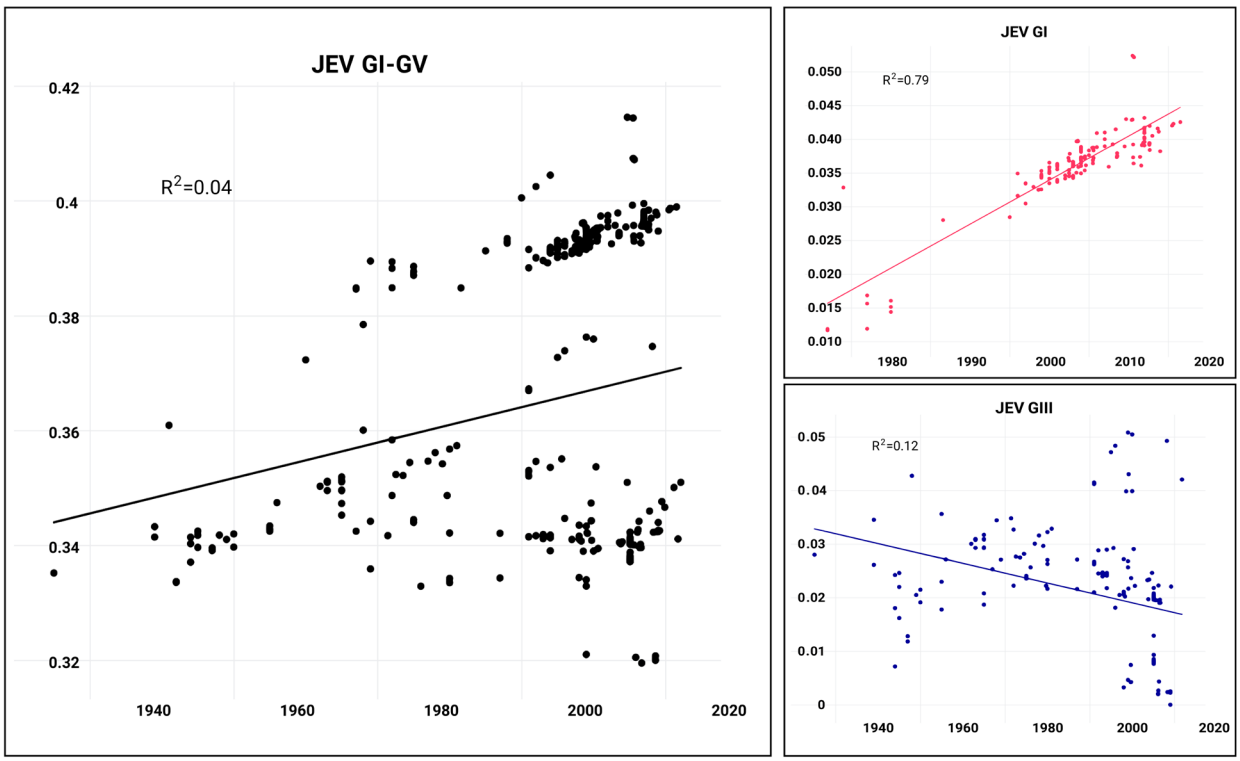


Figure 1 – Root-to-tips (RTT) regression analysis of three JEV data sets; X-axes represent calendar years, Y-axes — distances from a root to tips.

JEV GI-GV and GIII data sets showed R^2 values about zero (0.04 and 0.12 with a negative slope, respectively) indicating the weak clocklike behavior of the data which can be explained by the presence of a large amount of rate variation between lineages of the trees (95% HPD of CV, 1.04–1.62 for GI-GV and 1.29–2.4 for GIII). GI, on the contrary, demonstrated a stronger temporal structure with R^2 of 0.79 and the lowest CV values (95% HPD, 0.43–0.68) (Table 2).

Table 2. Comparison of the determination coefficients inferred in root-to-tip regression analysis and the coefficient of variation values.

Dataset	Number of samples	Sampling window, years	Determination coefficient (R^2) inferred in RTT analysis [1]	95% HPD of the coefficient of rate variation [2]
GI-GV	325	87	0.04	1.04–1.62
GI	134	45	0.79	0.43–0.68
GIII	171	87	0.12	1.29–2.4

[1] Root-to-tip (RTT) regression analysis was performed in Clockor2;
[2] 95% HPDs were inferred in BEAST analysis.

Importantly, the low R^2 values obtained should not be considered as the absence of a temporal signal. The point is that the RTT regression analysis appeared

178 to be uninformative when the data were sampled over a relatively narrow sampling
179 window (GI-GV data set) and there is large rate variation among lineages (GIII data
180 set) (Duchene et al., 2020). Therefore, we applied BETS — a formal test of temporal
181 signal — to assess the statistical support for including sampling times in a Bayesian
182 analysis.

183 **3.2 The presence of temporal signal in all JEV data sets**

184 BETS displayed very strong support (log Bayes factor > 5) for M_{het} over M_{iso}
185 in each pair of competitive models (Table 3) with the lowest log Bayes factor of 47.8
186 for the GIII data set (ORC_{het} relative to ORC_{iso}) which possessed the largest amount
187 of rate variation among branches (Table 2). Despite the last fact, in the log scale,
188 such BF values can be interpreted as overwhelming support for including sequence
189 sampling times into the model (Table 3).

190 Table 3. Results of the Bayesian test of temporal signal in Japanese
191 encephalitis virus data sets

Dataset	Molecular clock [1]	Log marginal likelihood	Log Bayes factor [2]
GI-GV	ORC_{iso}	−113559.93	140.38
	ORC_{het}	−113419.55	0
	SC_{iso}	−115098.41	1678.86
	SC_{het}	−114993.90	1574.34
GI	ORC_{iso}	−54633.51	129.02
	ORC_{het}	−54504.49	0
	SC_{iso}	−55182.96	678.48
	SC_{het}	−54734.24	229.75
GIII	ORC_{iso}	−46598.66	47.80
	ORC_{het}	−46550.86	0
	SC_{iso}	−47491.94	941.08
	SC_{het}	−47515.01	964.14

192 [1] ORC—optimized relaxed clock, SC—strict clock; iso—isochronous model, het—
193 heterochronous model;
194 [2] Best-performing model has a log Bayes factor value of 0.

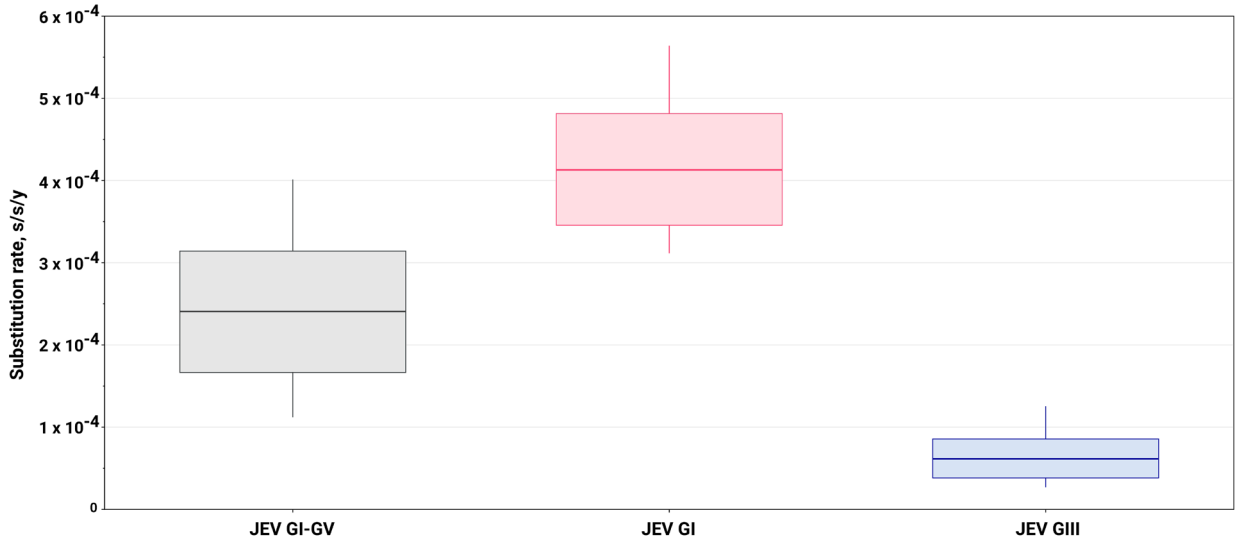
195 The comparison of the SC_{het} and ORC_{het} models also showed very strong
196 evidence in favor of ORC in each dataset. The lowest log Bayes factor of 229.75
197 was detected for, as expected, GI which had the highest R^2 in the RTT regression
198 analysis.

199 Thus, all three data sets passed the Bayesian test of temporal signal and were
200 analyzed with ORC in the next step of phylogenetic analysis.

201 **3.3 Substitution rates**

202 The substitution rate of the common JEV GI-GV clade was 2.41×10^{-4}
203 nucleotide substitutions per site per year (s/s/y) (95% HPD, 1.67×10^{-4}–3.14×10^{-4}
204 s/s/y) or approximately 2.5 nucleotide substitutions per ORF (10,245 nucleotides)
205 per year. GI has the highest rate of evolution: 4.13×10^{-4} s/s/y (95% HPD, 3.45×10^{-4}–
206 4.82×10^{-4} s/s/y) or approximately 4.2 nucleotide substitutions per ORF per year. The
207 lowest substitution rate was obtained for GIII: 6.17×10^{-5} s/s/y (95% HPD,
208 3.84×10^{-4}–8.63×10^{-4} s/s/y) or approximately 0.6 nucleotide substitutions per ORF
209 per year. Notably, the 95% HPD intervals of estimates do not overlap (Figure 2).

210



211
212 Figure 2. Substitution rates of common JEV clade (GI-GV), genotypes I (GI) and III (GIII).
213 Upper and lower ends of the boxes are 95 % highest posterior density interval borders.

214 **3.3.1 Comparing the substitution rates of JEV yielded in our analysis with**
215 **estimates reported previously**

216 We compared the JEV substitution rate values obtained in our study with the
217 previous estimates based on the analysis of complete ORF gene sequences (Figure
218 3).

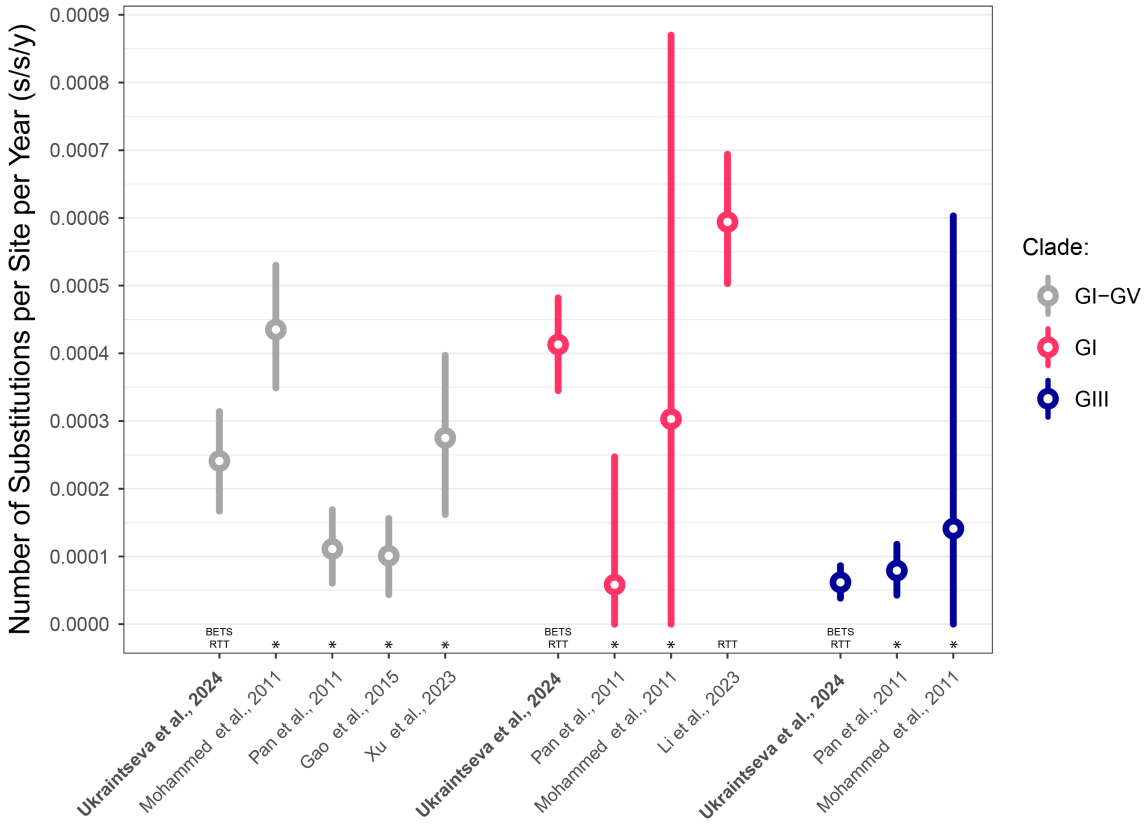


Figure 3. Comparison of the Japanese encephalitis virus (JEV) substitution rate evaluated in our study based on complete open reading frame nucleotide sequences and the JEV substitution rates according to the literature. The vertical bars represent 95% highest posterior density intervals, circles—median or mean values. The X-axis labels are references to appropriate studies, the methods on temporal signal evaluation applied are indicated above the X-axis: BETS—Bayesian evaluation of temporal signal, RTT—root-to-tips regression, *—temporal signal was not assessed.

First of all, we found only one study where at least the degree of clock-like behavior was assessed with informal RTT (Li et al., 2023) when analyzing GI. Therefore, the substitution rate values reported in the literature should be treated with caution, especially in the cases where the number of samples was relatively low. Hence, to our knowledge, we provide here, for the first time, the JEV evolutionary rate estimates confirmed by a formal statistical test.

The comparison showed that the evolutionary rate of the JEV common clade including all five genotypes GI-GV obtained in our study (95% HPD, 1.67×10^{-4}–3.14×10^{-4} s/s/y) is significantly lower than values reported by Mohammed et al. (2011) (95% HPD, 3.49×10^{-4}–5.30×10^{-4}) that can be explained by the smaller sample size (n=35). In contrast, our estimates are significantly higher than the JEV substitution rates provided by Pan et al. (2011) (95% HPD, 6.04×10^{-5}–1.69×10^{-4}) and Gao et al. (2015) (95% HPD, 4.37×10^{-5}–1.56×10^{-4}) who analyzed 98 (GI-GIV)

239 and 100 (GI-GV) sequences, respectively. These disparities may be accounted for
240 by the lack of temporal signal in the data sets of about 100 sequences. In the case of
241 low informative data, the Bayesian analysis turns out to be very sensitive to a prior
242 choice. Therefore, an inferred substitution rate is largely rested on, for example, a
243 continuous-time Markov chain (CTMC) rate reference prior — gamma probability
244 distribution with a shape parameter α=0.5 and rate parameter β=T, where T is tree
245 length (Ferreira and Suchard, 2010) — which is, in turn, ultimately conditioned by
246 a coalescent population size prior (i.e., the more population size, the longer tree).
247 In one of the last studies, Xu et al. (2023) reported a JEV GI-GV substitution
248 rate (95% HPD, 1.62×10^{-4}–3.97×10^{-4}) comparable with our estimates. Notably, the
249 sample size (115), clock model (UCLD) and tree prior (Bayesian Skyline) applied
250 in the work of Xu et al. (2023) were similar to two previously discussed studies with
251 significantly lower JEV substitution rates (Pan et al., 2011; Gao et al., 2015). The
252 possible explanation is that Xu et al. (2023) provide new JEV nucleotide sequences,
253 however without testing at least the measure of clock-like evolution by RTT, the
254 question about the true factor of distorted inference of substitution rate persists
255 unresolved.
256 Our estimates of JEV GI evolutionary rate (95% HPD, 3.45×10^{-4}–4.82×10^{-4})
257 are significantly higher than those calculated for the GI-GV clade (95% HPD,
258 1.67×10^{-4}–3.14×10^{-4}) (Figure 3). GI rate values inferred in our study are also
259 statistically different from the literature data, except for the estimates with the widest
260 95% HPD interval (3.42×10^{-9}–2.47×10^{-4}) provided by Mohammed et al. (2011).
261 This wide width of the 95% HPD interval is likely due to a small sample size (n=10)
262 and, as a consequence, the absence of a temporal signal which was not evaluated.
263 The only study (Li et al., 2023) in which the extent of time-dependency in data was
264 assessed (RTT) reported statistically higher JEV GI evolutionary rates (95% HPD,
265 5.03×10^{-4}–6.94×10^{-4}) than our estimates. We speculated on two possible
266 explanations of the discrepancy — the first one is Li et al., 2023 removed RTT
267 regression outliers without explaining the reason (e.g., sequencing errors, long
268 passage history, lab contamination, etc.) that is forcing the data to fit a strict clock

269 model, which does not assume among-lineage rate variation, instead of using a
270 proper model (e.g., relaxed clock) to describe the data; the second one is the new
271 JEV nucleotide sequences generated in the study of Li et al. (2023) and not included
272 in our work.

273 Among all three data sets analyzed (Table 1), JEV GIII was characterized by
274 the lowest evolutionary rate (95% HPD, 3.84×10^{-5}–8.63×10^{-5}) (Figure 3). Thus,
275 JEV GIII evolves on the other order of magnitude than JEV GI and JEV GI-GV. The
276 very similar estimates (4.28×10^{-5}–1.18×10^{-4}) were reported by Pan et al. (2011) with
277 a sample size of 53 and a model combination similar to the one we applied in our
278 work with the exception of a relaxed molecular clock (UCLD). In the study of
279 Mohammed et al. (2011), the 95% HPD interval was the widest one (7.05×10^{-8}–
280 6.03×10^{-4}) like in the case of JEV GI and the sample size was probably one of the
281 main reasons for that (n=22). To date, JEV GIII substitution rate estimates derived
282 in our study are relatively the most accurate.

283 To summarize, GI is the fastest-evolving genotype of JEV and, conversely,
284 GIII is the slowest JEV genotype. Concerning 95% HPD interval width, our JEV
285 rate estimates are the most or one of the most robust.

286 **3.3.2 Comparing the substitution rate of JEV (GI-GV clade) and other**
287 **mosquito-borne flaviviruses**

288 To provide a better understanding of the JEV evolutionary rate, we compared
289 the rate of JEV GI-GV clade obtained with rates of JEV most closely related species
290 and the other well-studied mosquito-borne flaviviruses (Figure 4).

291 For the comparison of the JEV substitution rate, we considered the data only
292 if 95% HPD intervals were provided. Also, the estimates reported with mistakes or
293 obvious typos were excluded from the comparison. We mainly focused on the
294 substitution rates derived from the analysis of as many virus genotypes as possible
295 (i.e., the genetic diversity of viruses was meant to be represented to its greatest
296 extent).

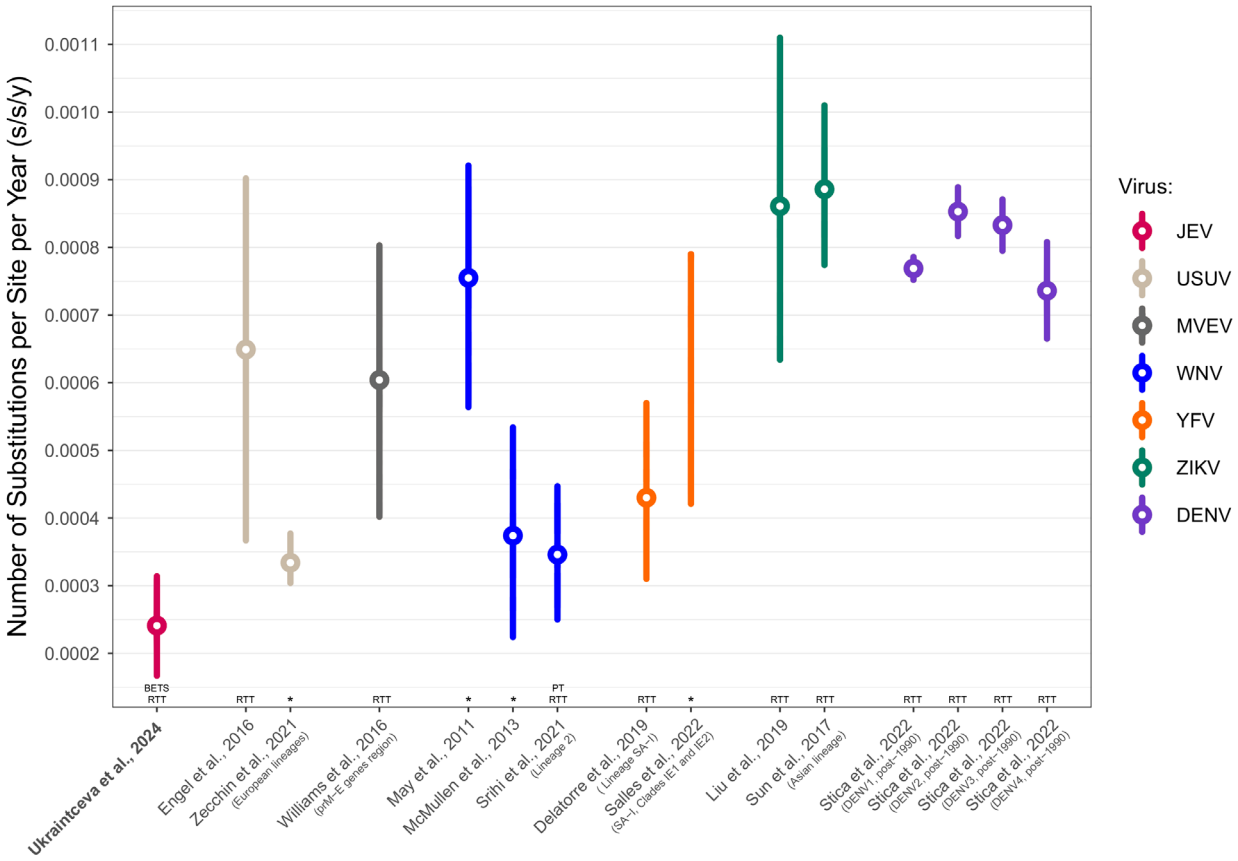


Figure 4. Comparison of the Japanese encephalitis virus (JEV) substitution rate inferred in our study using 325 complete open reading frame nucleotide sequences of the genotype I–V clade (GI-GV) and the substitution rates for the other mosquito-borne flaviviruses according to the literature: USUV—Usutu virus (Engel et al., 2016; Zecchin et al., 2021), MVEV—Murray Valley encephalitis virus (Williams et al., 2015), WNV—West Nile virus (May et al., 2011; McMullen et al., 2013; Srihi et al., 2021), YFV—yellow fever virus (Delatorre et al., 2019; Moreira Salles et al., 2022), ZIKV—Zika virus (Sun et al., 2017; Liu et al., 2019), DENV—dengue virus (Stica et al., 2022). The vertical bars represent 95% highest posterior density intervals, circles—median or mean values if available. The X-axis labels are references to appropriate studies. The methods on temporal signal evaluation applied are indicated above the X-axis: BETS—Bayesian evaluation of temporal signal, RTT—root-to-tips regression, PT—permutation test, *—temporal signal was not assessed.

Unlike the literature data on JEV, the most of MBFV rate estimates were supported by the RTT regressions, and, furthermore, a permutation or randomization test was applied in one study (Srihi et al., 2021). This makes the comparison much more reliable, but, for some virus species, reported rates are statistically different between studies.

First, we compared our JEV substitution rate estimates with the rates of JEV most closely related species (Moureau et al., 2015) — Usutu virus (USUV), Murray Valley encephalitis virus (MVEV), West Nile virus (WNV) — for which the sample size allows the evolutionary rate of the virus to be reliably evaluated. Second, we

included in the comparison the most widely understood members of MBFVs — yellow fever virus (YFV), Zika virus (ZIKV), and dengue virus (DENV).

As can be seen in Figure 4, JEV is one of the slowest changing MBFVs represented in our analysis (95% HPD, 1.67×10^{-4}–3.14×10^{-4}). The overlap of 95% HPD intervals is observed in the case of two assessments of WNV (2.24×10^{-4}–5.34×10^{-4} according to McMullen et al. (2013) and 2.50×10^{-4}–4.47×10^{-4} for WNV lineage 2 reported by Srihi et al. (2021)). There is the slight overlap with the rate values of the USUV European lineage (95% HPD, 3.04×10^{-4}–3.77×10^{-4}). More distant MBFV members evolve faster, especially ZIKV (Sun et al., 2017; Liu et al., 2019) and DENV (Stica et al., 2022), the substitution rates of which are near the threshold of 10^{-3} and 10^{-2} orders of magnitude or even above it.

3.4 Divergence times

The evaluation of JEV divergence times was performed using 325 complete ORF nucleotide sequences (Figure 5). The results of the analysis in BEAST suggest that JEV emerged 1234 years ago (95% HPD, 407–2333 years).

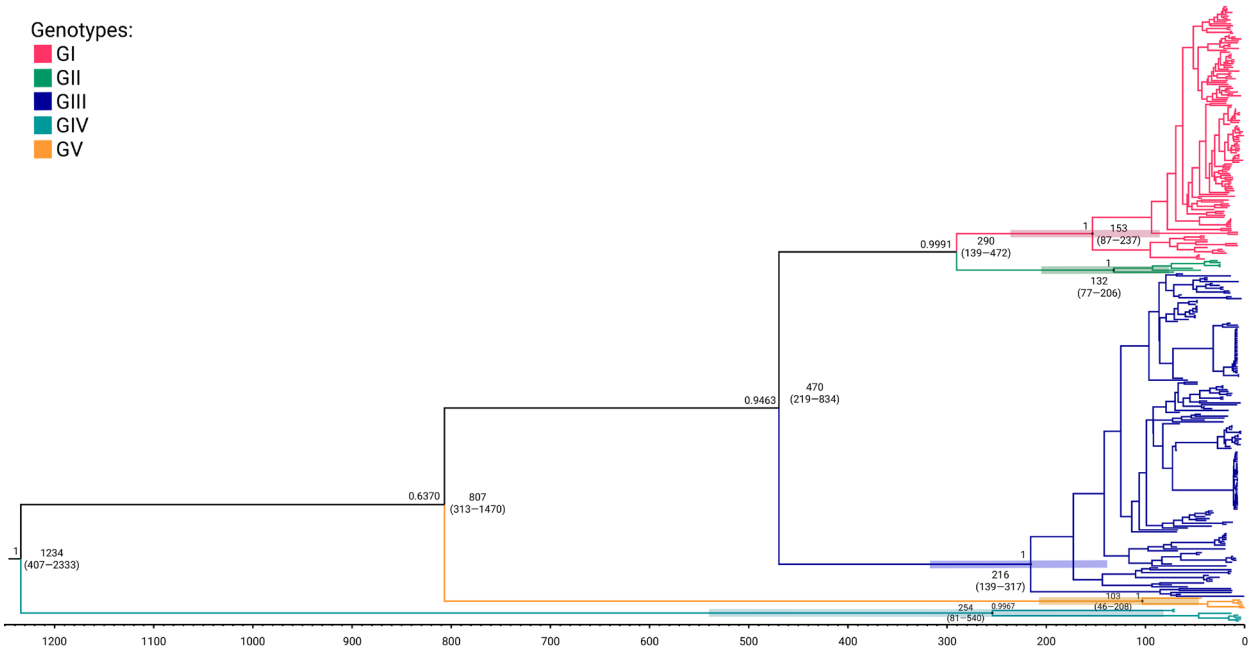


Figure 5 – Maximum clade credibility tree of JEV summarized from a posterior tree distribution produced by an MCMC run in BEAST. The numbers nearby the main nodes indicate divergence times before the most recent sample collected on February 20, 2022, with 95% highest posterior density (HPD) intervals within parantheses. The horizontal dark green bars depict 95% HPD of the divergence times of JEV genotypes.

340 Despite the relatively large number of complete ORF sequences used in our
341 study, the uncertainty in the assessment of the evolutionary timescale of the entire
342 tree spans a large time interval (95% HPD of root height, 407–2333 years). A little
343 more accurate estimates of time to MRCA of JEV GI-GV (95% HPD, 463–2100
344 years) were reported by Xu et al. (2023) with a sample size of 115 complete genome
345 sequences. In the other studies with smaller sample sizes, 95% HPD intervals for
346 tree roots were much wider: Gao et al. (2015) conclude that the MRCA of JEV GI-
347 GV is estimated to have occurred 3255 years ago (95% HPD, 978–6125 years); Pan
348 et al. (2011) studied JEV GI-GIV defined the MRCA age of the entire clade as 1695
349 years ago (95% HPD, 548–3153 years). More recent JEV emergence was described
350 by Mohammed et al. (2011) (95% HPD, 374–500 years ago), but the sample size
351 was the smallest one (n=35) which implies weak temporal structure in the data and
352 is also reflected in the substitution rate estimates (Figure 3).
353 Concerning the emergence of JEV genotypes, 95% HPD intervals of time to
354 MRCA of all JEV genotypes are overlapped (Figure 5). Therefore, we are unable to
355 reliably establish the chronology of the occurrence of genotypes. Besides, only the
356 GI and GIII are representative in terms of sample size, and, as a consequence, the
357 estimates of the other genotype ages (GII, GIV, GV) are unreliable. Generally,
358 considering the bounds of 95 HPDs obtained, JEV genotypes occurred 77–540 years
359 ago.
360 The yielded ages of JEV subtypes are consistent with the previous assessments
361 (Xu et al., 2023). However, the "minor" JEV genotypes, such as GII, GIV, GV, were
362 mainly scrutinized in the studies of E gene data sets. The estimates of time to MRCA
363 based on E gene sequences are more accurate (Schuh et al., 2013; Sikazwe et al.,
364 2022), whereas the substitution rate of E gene is significantly higher (95% HPD,
365 3.92×10^{-4}–6.52×10^{-4}) and MRCA ages are younger (e.g., the MRCA of JEV GI-GV,
366 95% HPD 216–921 years) (Schuh et al., 2013). "Rejuvenation" of the tree root along
367 with (or possibly due to) an increase in the evolutionary rate indicates the impact of
368 the priors on the inference of the evolutionary timescale. In other words, even if the
369 E gene of JEV evolved faster than the entire ORF gene, then the dates of the

370 emergence of JEV MRCA should not significantly differ between the analyses of
371 two loci in the case of the obligatory presence of temporal signal in both data sets.
372 In the cited works dedicated to the E gene analysis, only Sikazwe et al. (2022)
373 performed the RTT test, however, as a result of a test, JEV GV was removed from
374 the analysis as an outlier without a justified reason. Moreover, the R^2 and the
375 regression visualization were not reported as well as the substitution rate of the
376 common data set (JEV GI-GIV). Thus, the problem of the fitness of the E gene for
377 JEV date-stamped phylogeny reconstruction should be addressed in future studies.

378 **Conclusions**

379 1. All complete ORF data sets analyzed (JEV GI-GV, GI, GIII) yielded
380 strong temporal signal in BETS and can be used for the reconstructions of dated
381 phylogenies;

382 2. JEV represented by the clade comprising all five genotypes (GI-GV)
383 has the lowest substitution rate in comparison to its most closely related species and
384 other well-studied members of mosquito-borne flaviviruses: 2.41×10^{-4} s/s/y (95%
385 HPD, 1.67×10^{-4}–3.14×10^{-4} s/s/y) or approximately 2.5 nucleotide substitutions per
386 ORF (10,245 nucleotides) per year. To our knowledge, it is the first reliable
387 assessment of the JEV substitution rate confirmed by the formal Bayesian test of
388 temporal signal;

389 3. JEV GIII is characterized by the lowest substitution rate (6.17×10^{-5}
390 s/s/y (95% HPD, 3.84×10^{-4}–8.63×10^{-4} s/s/y) or approximately 0.6 nucleotide
391 substitutions per ORF per year) in comparison with GI and the common JEV clade.
392 It is the most accurate estimate we found in the literature;

393 4. Time to JEV MRCA (GI-GV clade) has the mean value of 1234 years
394 (95% HPD, 407–2333 years);

395 5. The mean values and appropriate 95% HPD intervals of divergence of
396 each JEV genotype are comparable and close. Based on the lower and upper bounds
397 of HPDs, JEV genotypes occurred in the period from 77 to 540 years ago.

References

Kass, R.E., Raftery, A.E., 1995. Bayes Factors. Journal of the American Statistical Association 90, 773-795. https://doi.org/10.1080/01621459.1995.10476572.

Mackenzie, J.S., 2005. Emerging zoonotic encephalitis viruses: lessons from Southeast Asia and Oceania. J Neurovirol 11, 434-440. https://doi.org/10.1080/13550280591002487.

Lindenbach, B.T., H.J.; Rice, C.M., 2007. Flaviviridae: The viruses and their replication. . Fields Virol., pp. 1101–1151.

Ferreira, M.A.R., Suchard, M.A., 2010. Bayesian analysis of elapsed times in continuous-time Markov chains. Canadian Journal of Statistics 36, 355-368. https://doi.org/10.1002/cjs.5550360302.

May, F.J., Davis, C.T., Tesh, R.B., Barrett, A.D.T., 2011. Phylogeography of West Nile Virus: from the Cradle of Evolution in Africa to Eurasia, Australia, and the Americas. Journal of Virology 85, 2964-2974. https://doi.org/10.1128/jvi.01963-10.

Mohammed, M.A., Galbraith, S.E., Radford, A.D., Dove, W., Takasaki, T., Kurane, I., Solomon, T., 2011. Molecular phylogenetic and evolutionary analyses of Muar strain of Japanese encephalitis virus reveal it is the missing fifth genotype. Infect Genet Evol 11, 855-862. https://doi.org/10.1016/j.meegid.2011.01.020.

Pan, X.L., Liu, H., Wang, H.Y., Fu, S.H., Liu, H.Z., Zhang, H.L., Li, M.H., Gao, X.Y., Wang, J.L., Sun, X.H., et al., 2011. Emergence of genotype I of Japanese encephalitis virus as the dominant genotype in Asia. J Virol 85, 9847-9853. https://doi.org/10.1128/JVI.00825-11.

Takhampunya, R., Kim, H.C., Tippayachai, B., Kengluecha, A., Klein, T.A., Lee, W.J., Grieco, J., Evans, B.P., 2011. Emergence of Japanese encephalitis virus genotype V in the Republic of Korea. Virol J 8, 449. https://doi.org/10.1186/1743-422X-8-449.

Ayres, D.L., Darling, A., Zwickl, D.J., Beerli, P., Holder, M.T., Lewis, P.O., Huelsenbeck, J.P., Ronquist, F., Swofford, D.L., Cummings, M.P., et al., 2012.

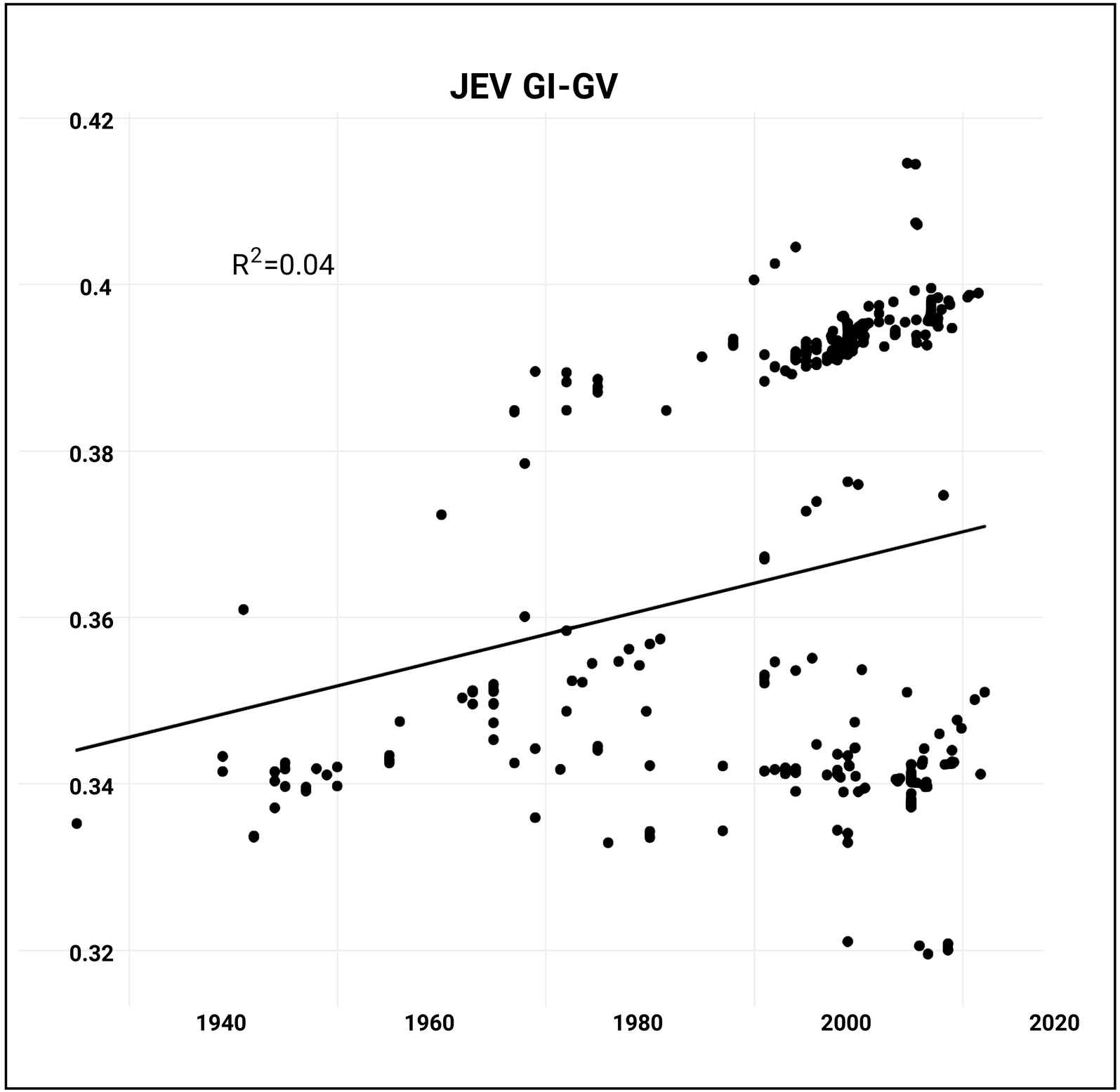
428 BEAGLE: an application programming interface and high-performance computing
429 library for statistical phylogenetics. Syst Biol 61, 170-173.
430 https://doi.org/10.1093/sysbio/syr100.
431 Ravanini, P., Huhtamo, E., Ilaria, V., Crobu, M.G., Nicosia, A.M., Servino, L.,
432 Rivasi, F., Allegrini, S., Miglio, U., Magri, A., et al., 2012. Japanese encephalitis
433 virus RNA detected in Culex pipiens mosquitoes in Italy. Euro Surveill 17.
434 https://doi.org/10.2807/ese.17.28.20221-en.
435 Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment
436 software version 7: improvements in performance and usability. Mol Biol Evol 30,
437 772-780. https://doi.org/10.1093/molbev/mst010.
438 McMullen, A.R., Albayrak, H., May, F.J., Davis, C.T., Beasley, D.W.C.,
439 Barrett, A.D.T., 2013. Molecular evolution of lineage 2 West Nile virus. J Gen Virol
440 94, 318-325. https://doi.org/10.1099/vir.0.046888-0.
441 Schuh, A.J., Ward, M.J., Brown, A.J., Barrett, A.D., 2013. Phylogeography of
442 Japanese encephalitis virus: genotype is associated with climate. PLoS Negl Trop
443 Dis 7, e2411. https://doi.org/10.1371/journal.pntd.0002411.
444 Larsson, A., 2014. AliView: a fast and lightweight alignment viewer and
445 editor for large data sets. Bioinformatics 30, 3276-3278.
446 https://doi.org/10.1093/bioinformatics/btu531.
447 Gao, X., Liu, H., Li, M., Fu, S., Liang, G., 2015. Insights into the evolutionary
448 history of Japanese encephalitis virus (JEV) based on whole-genome sequences
449 comprising the five genotypes. Virol J 12, 43. https://doi.org/10.1186/s12985-015-
450 0270-z.
451 Moureau, G., Cook, S., Lemey, P., Nougairede, A., Forrester, N.L.,
452 Khasnatinov, M., Charrel, R.N., Firth, A.E., Gould, E.A., de Lamballerie, X., 2015.
453 New insights into flavivirus evolution, taxonomy and biogeographic history,
454 extended by analysis of canonical and alternative coding sequences. PLoS One 10,
455 e0117849. https://doi.org/10.1371/journal.pone.0117849.
456 Williams, D.T., Diviney, S.M., Niazi, A.U., Durr, P.A., Chua, B.H., Herring,
457 B., Pyke, A., Doggett, S.L., Johansen, C.A., Mackenzie, J.S., 2015. The Molecular

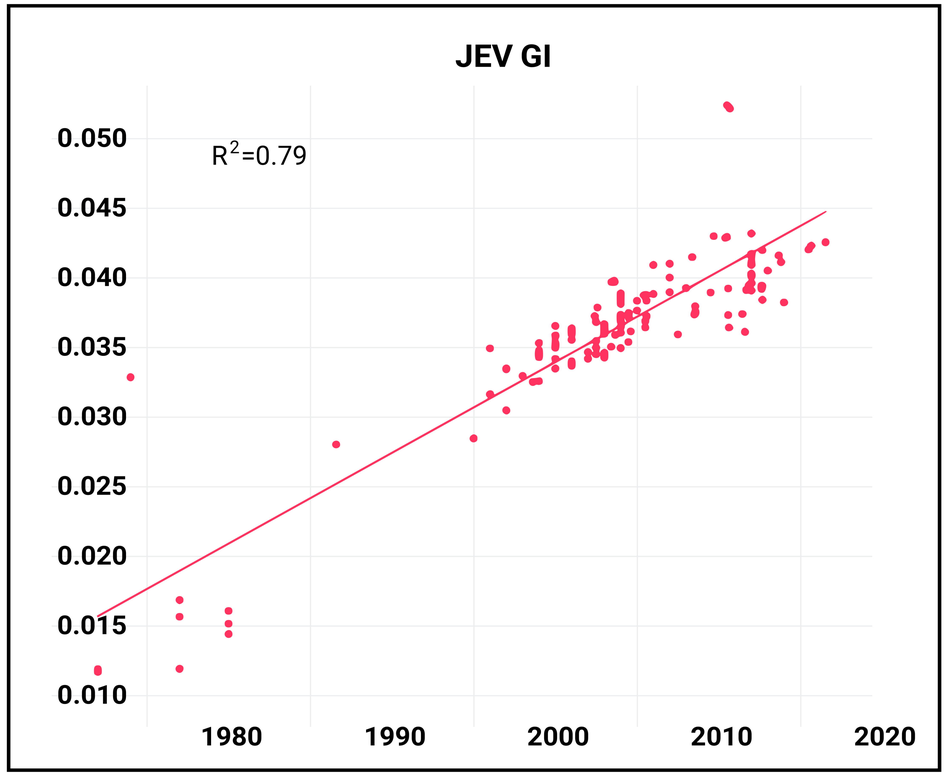
458 Epidemiology and Evolution of Murray Valley Encephalitis Virus: Recent
459 Emergence of Distinct Sub-lineages of the Dominant Genotype 1. PLoS Negl Trop
460 Dis 9, e0004240. https://doi.org/10.1371/journal.pntd.0004240.
461 Engel, D., Jost, H., Wink, M., Borstler, J., Bosch, S., Garigliany, M.M., Jost,
462 A., Czajka, C., Luhken, R., Ziegler, U., et al., 2016. Reconstruction of the
463 Evolutionary History and Dispersal of Usutu Virus, a Neglected Emerging Arbovirus
464 in Europe and Africa. mBio 7, e01938-01915. https://doi.org/10.1128/mBio.01938-
465 15.
466 Ricklin, M.E., Garcia-Nicolas, O., Brechbuhl, D., Python, S., Zumkehr, B.,
467 Nougairede, A., Charrel, R.N., Posthaus, H., Oevermann, A., Summerfield, A., 2016.
468 Vector-free transmission and persistence of Japanese encephalitis virus in pigs. Nat
469 Commun 7, 10832. https://doi.org/10.1038/ncomms10832.
470 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermiin,
471 L.S., 2017. ModelFinder: fast model selection for accurate phylogenetic estimates.
472 Nat. Methods 14, 587-589. https://doi.org/10.1038/nmeth.4285.
473 Simon-Loriere, E., Faye, O., Prot, M., Casademont, I., Fall, G., Fernandez-
474 Garcia, M.D., Diagne, M.M., Kipela, J.M., Fall, I.S., Holmes, E.C., et al., 2017.
475 Autochthonous Japanese Encephalitis with Yellow Fever Coinfection in Africa. N
476 Engl J Med 376, 1483-1485. https://doi.org/10.1056/NEJMc1701600.
477 Sun, J.W., D., Zhong, H., Guan, D., Zhang, H., Tan, Q., Zhou, H., Zhang, M.,
478 Ning, D., Zhang, B., Ke, C., et al., 2017. Returning ex-patriot Chinese to
479 Guangdong, China, increase the risk for local transmission of Zika virus. J Infect 75,
480 356-367. https://doi.org/10.1016/j.jinf.2017.07.001.
481 Rambaut, A., Drummond, A.J., Dong, X., Baele, G., Suchard, M.A., 2018.
482 Posterior summarization in Bayesian phylogenetics using tracer 1.7. Syst. Biol. 67,
483 901-904. https://doi.org/10.1093/sysbio/syy032.
484 Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchene, S., Fourment, M.,
485 Gavryushkina, A., Heled, J., Jones, G., Kuhnert, D., De Maio, N., et al., 2019.
486 BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis.
487 PLoS Comput. Biol 15, e1006650. https://doi.org/10.1371/journal.pcbi.1006650.

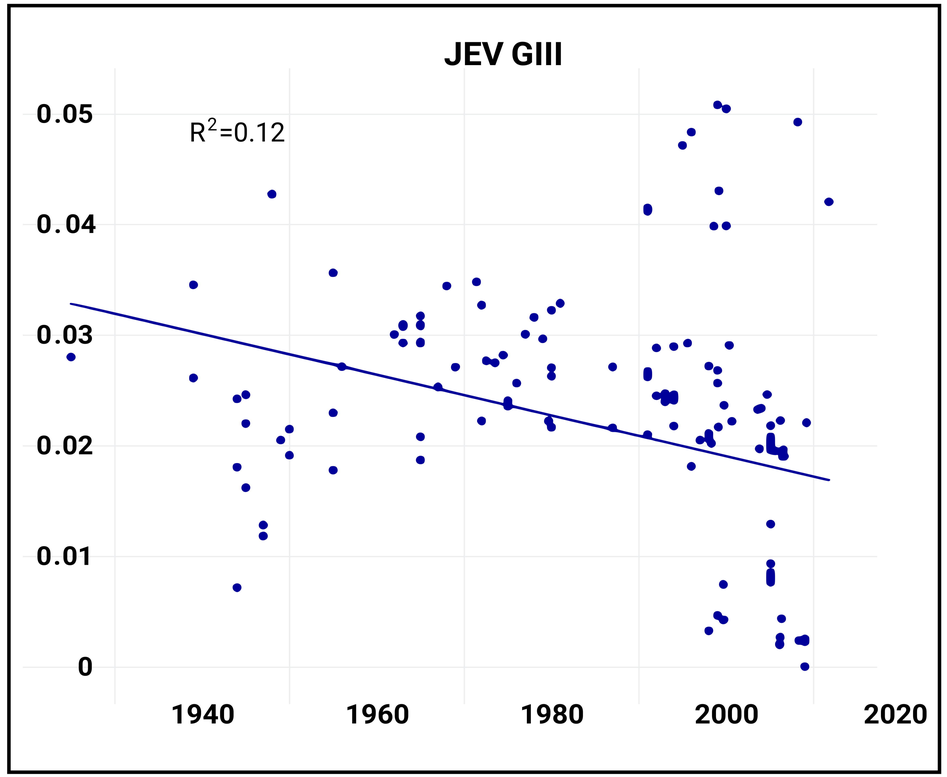
488 Delatorre, E., de Abreu, F.V.S., Ribeiro, I.P., Gomez, M.M., Dos Santos,
489 A.A.C., Ferreira-de-Brito, A., Neves, M., Bonelly, I., de Miranda, R.M., Furtado,
490 N.D., et al., 2019. Distinct YFV Lineages Co-circulated in the Central-Western and
491 Southeastern Brazilian Regions From 2015 to 2018. Front Microbiol 10, 1079.
492 https://doi.org/10.3389/fmicb.2019.01079.
493 Ladreyt, H., Durand, B., Dussart, P., Chevalier, V., 2019. How Central Is the
494 Domestic Pig in the Epidemiological Cycle of Japanese Encephalitis Virus? A
495 Review of Scientific Evidence and Implications for Disease Control. Viruses 11.
496 https://doi.org/10.3390/v11100949.
497 Liu, R.F., He, Z.J., Mei, P., Xi, J.C., Cao, X.D., Yuan, L.H., Lu, J.H., 2019.
498 Risk assessment and genomic characterization of Zika virus in China and its
499 surrounding areas. Chin Med J (Engl) 132, 1645-1653.
500 https://doi.org/10.1097/CM9.0000000000000317.
501 Duchene, S., Lemey, P., Stadler, T., Ho, S.Y.W., Duchene, D.A.,
502 Dhanasekaran, V., Baele, G., 2020. Bayesian Evaluation of Temporal Signal in
503 Measurably Evolving Populations. Molecular Biology and Evolution.
504 https://doi.org/10.1093/molbev/msaa163.
505 Auerswald, H., Maquart, P.O., Chevalier, V., Boyer, S., 2021. Mosquito Vector
506 Competence for Japanese Encephalitis Virus. Viruses 13.
507 https://doi.org/10.3390/v13061154.
508 Douglas, J., Zhang, R., Bouckaert, R., 2021. Adaptive dating and fast
509 proposals: Revisiting the phylogenetic relaxed clock model. PLoS Comput Biol 17,
510 e1008322. https://doi.org/10.1371/journal.pcbi.1008322.
511 Martin, D.P., Varsani, A., Roumagnac, P., Botha, G., Maslamoney, S., Schwab,
512 T., Kelz, Z., Kumar, V., Murrell, B., 2021. RDP5: a computer program for analyzing
513 recombination in, and removing signals of recombination from, nucleotide sequence
514 datasets. Virus Evol 7, veaa087. https://doi.org/10.1093/ve/veaa087.
515 Srihi, H., Chatti, N., Ben Mhadheb, M., Gharbi, J., Abid, N., 2021.
516 Phylodynamic and phylogeographic analysis of the complete genome of the West

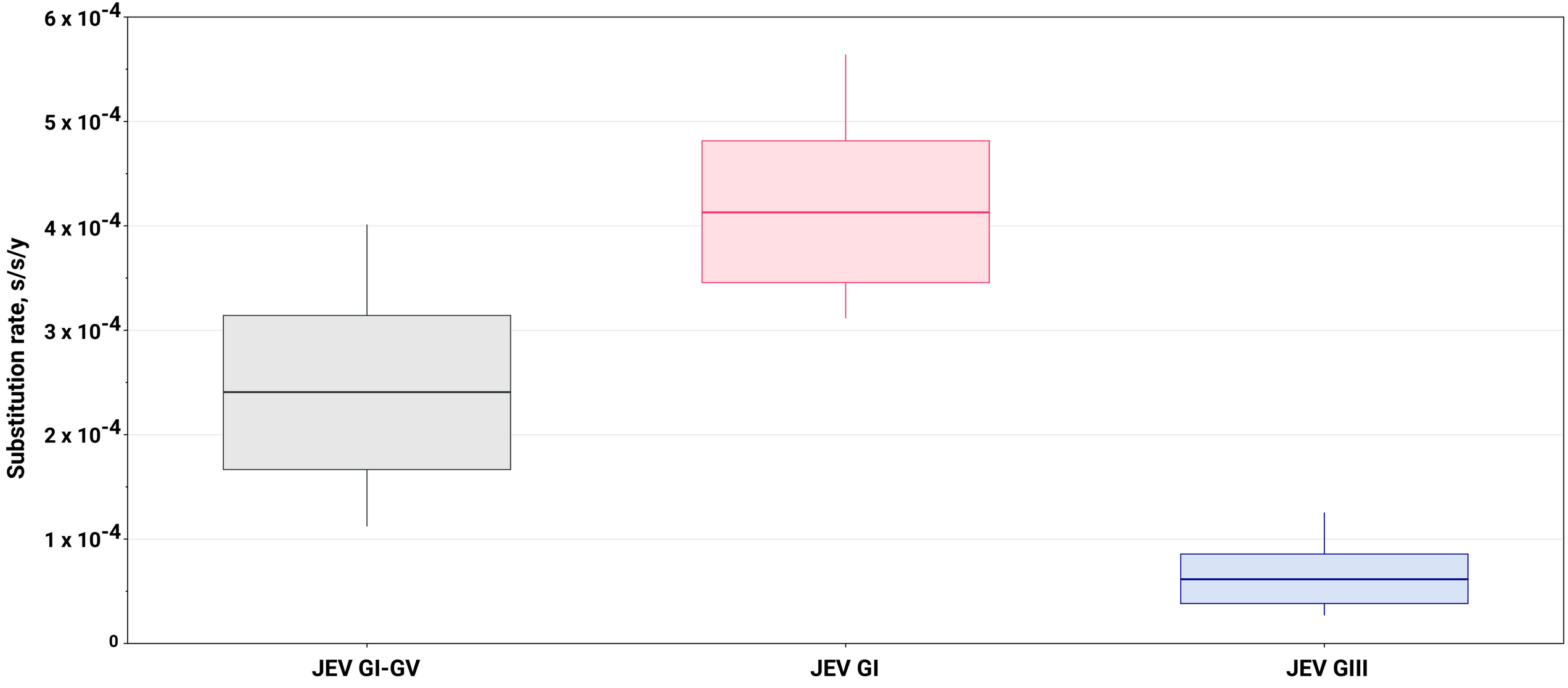
517 Nile virus lineage 2 (WNV-2) in the Mediterranean basin. BMC Ecol Evol 21, 183.
518 https://doi.org/10.1186/s12862-021-01902-w.
519 Vannice, K.S., Hills, S.L., Schwartz, L.M., Barrett, A.D., Heffelfinger, J.,
520 Hombach, J., Letson, G.W., Solomon, T., Marfin, A.A., Japanese encephalitis
521 vaccination experts, p., 2021. The future of Japanese encephalitis vaccination: expert
522 recommendations for achieving and maintaining optimal JE control. NPJ Vaccines
523 6, 82. https://doi.org/10.1038/s41541-021-00338-z.
524 Zecchin, B., Fusaro, A., Milani, A., Schivo, A., Ravagnan, S., Ormelli, S.,
525 Mavian, C., Michelutti, A., Toniolo, F., Barzon, L., et al., 2021. The central role of
526 Italy in the spatial spread of USUTU virus in Europe. Virus Evol 7, veab048.
527 https://doi.org/10.1093/ve/veab048.
528 Cheng, Y., Tran Minh, N., Tran Minh, Q., Khandelwal, S., Clapham, H.E.,
529 2022. Estimates of Japanese Encephalitis mortality and morbidity: A systematic
530 review and modeling analysis. PLoS Negl Trop Dis 16, e0010361.
531 https://doi.org/10.1371/journal.pntd.0010361.
532 Moreira Salles, A.P., de Seixas Santos Nastri, A.C., Ho, Y.L., Vilas Boas
533 Casadio, L., Emanuel Amgarten, D., Justo Arevalo, S., Soares Gomes-Gouvea, M.,
534 Jose Carrilho, F., de Mello Malta, F., Rebello Pinho, J.R., 2022. Updating the
535 Phylodynamics of Yellow Fever Virus 2016-2019 Brazilian Outbreak With New
536 2018 and 2019 Sao Paulo Genomes. Front Microbiol 13, 811318.
537 https://doi.org/10.3389/fmicb.2022.811318.
538 Sikazwe, C., Neave, M.J., Michie, A., Mileto, P., Wang, J., Cooper, N., Levy,
539 A., Imrie, A., Baird, R.W., Currie, B.J., et al., 2022. Molecular detection and
540 characterisation of the first Japanese encephalitis virus belonging to genotype IV
541 acquired in Australia. PLoS Negl Trop Dis 16, e0010754.
542 https://doi.org/10.1371/journal.pntd.0010754.
543 Stica, C.J., Barrero, R.A., Murray, R.Z., Devine, G.J., Phillips, M.J., Frentiu,
544 F.D., 2022. Global Evolutionary History and Dynamics of Dengue Viruses Inferred
545 from Whole Genome Sequences. Viruses 14. https://doi.org/10.3390/v14040703.

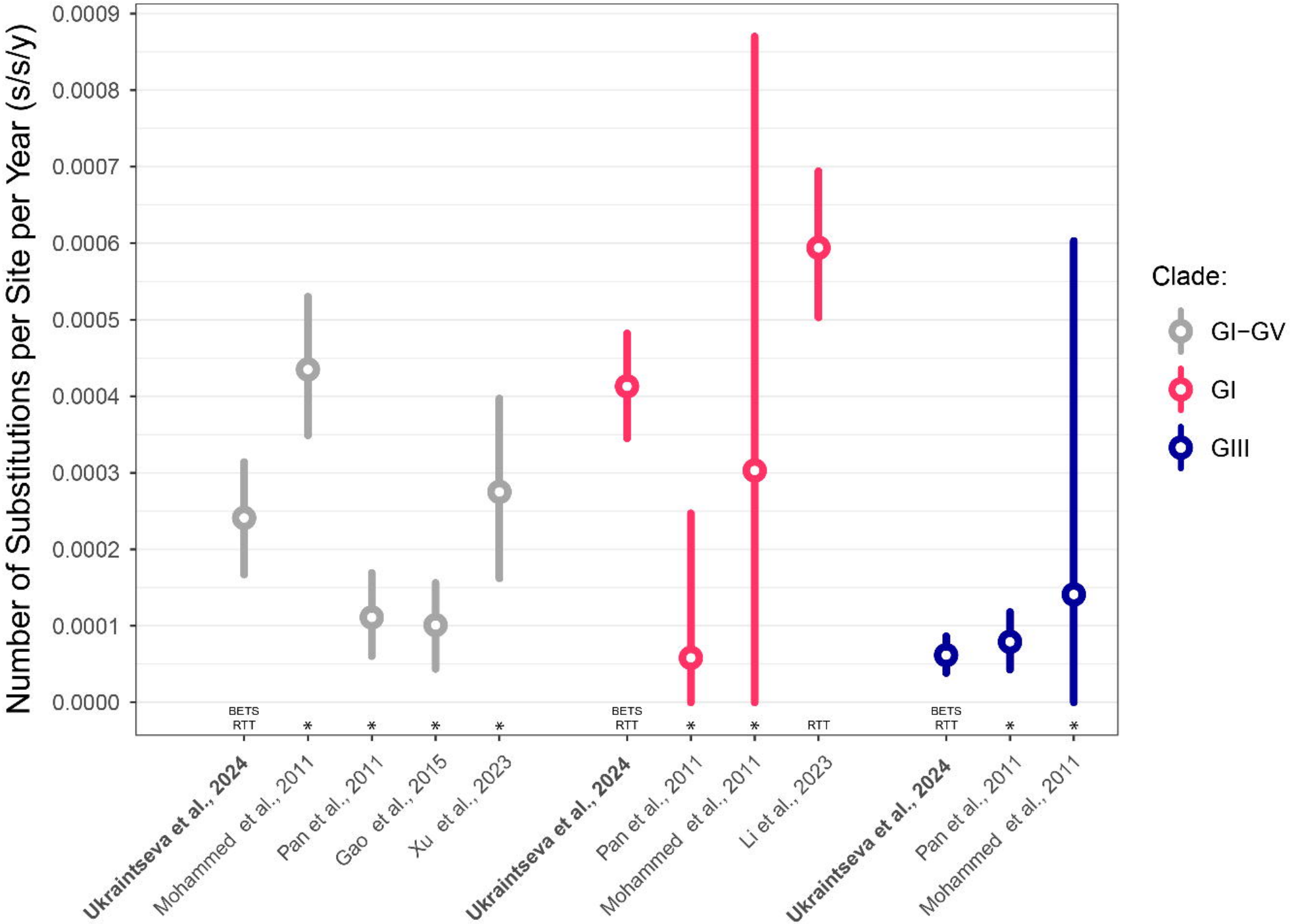
546 Van den Eynde, C., Sohier, C., Matthijs, S., De Regge, N., 2022. Japanese
547 Encephalitis Virus Interaction with Mosquitoes: A Review of Vector Competence,
548 Vector Capacity and Mosquito Immunity. Pathogens 11.
549 https://doi.org/10.3390/pathogens11030317.
550 Featherstone, L.A., Rambaut, A., Duchene, S., Wirth, W., 2023. Clockor2:
551 Inferring global and local strict molecular clocks using root-to-tip regression.
552 bioRxiv, 2023.2007.2013.548947. https://doi.org/10.1101/2023.07.13.548947.
553 Li, F., Feng, Y., Wang, G., Zhang, W., Fu, S., Wang, Z., Yin, Q., Nie, K., Yan,
554 J., Deng, X., et al., 2023. Tracing the spatiotemporal phylodynamics of Japanese
555 encephalitis virus genotype I throughout Asia and the western Pacific. PLoS Negl
556 Trop Dis 17, e0011192. https://doi.org/10.1371/journal.pntd.0011192.
557 McGuinness, S.L., Lau, C.L., Leder, K., 2023. The evolving Japanese
558 encephalitis situation in Australia and implications for travel medicine. J Travel Med
559 30. https://doi.org/10.1093/jtm/taad029.
560 Postler, T.S., Beer, M., Blitvich, B.J., Bukh, J., de Lamballerie, X., Drexler,
561 J.F., Imrie, A., Kapoor, A., Karganova, G.G., Lemey, P., et al., 2023. Renaming of
562 the genus Flavivirus to Orthoflavivirus and extension of binomial species names
563 within the family Flaviviridae. Arch Virol 168, 224. https://doi.org/10.1007/s00705-
564 023-05835-1.
565 Xu, G., Gao, T., Wang, Z., Zhang, J., Cui, B., Shen, X., Zhou, A., Zhang, Y.,
566 Zhao, J., Liu, H., et al., 2023. Re-Emerged Genotype IV of Japanese Encephalitis
567 Virus Is the Youngest Virus in Evolution. Viruses 15.
568 https://doi.org/10.3390/v15030626.
569

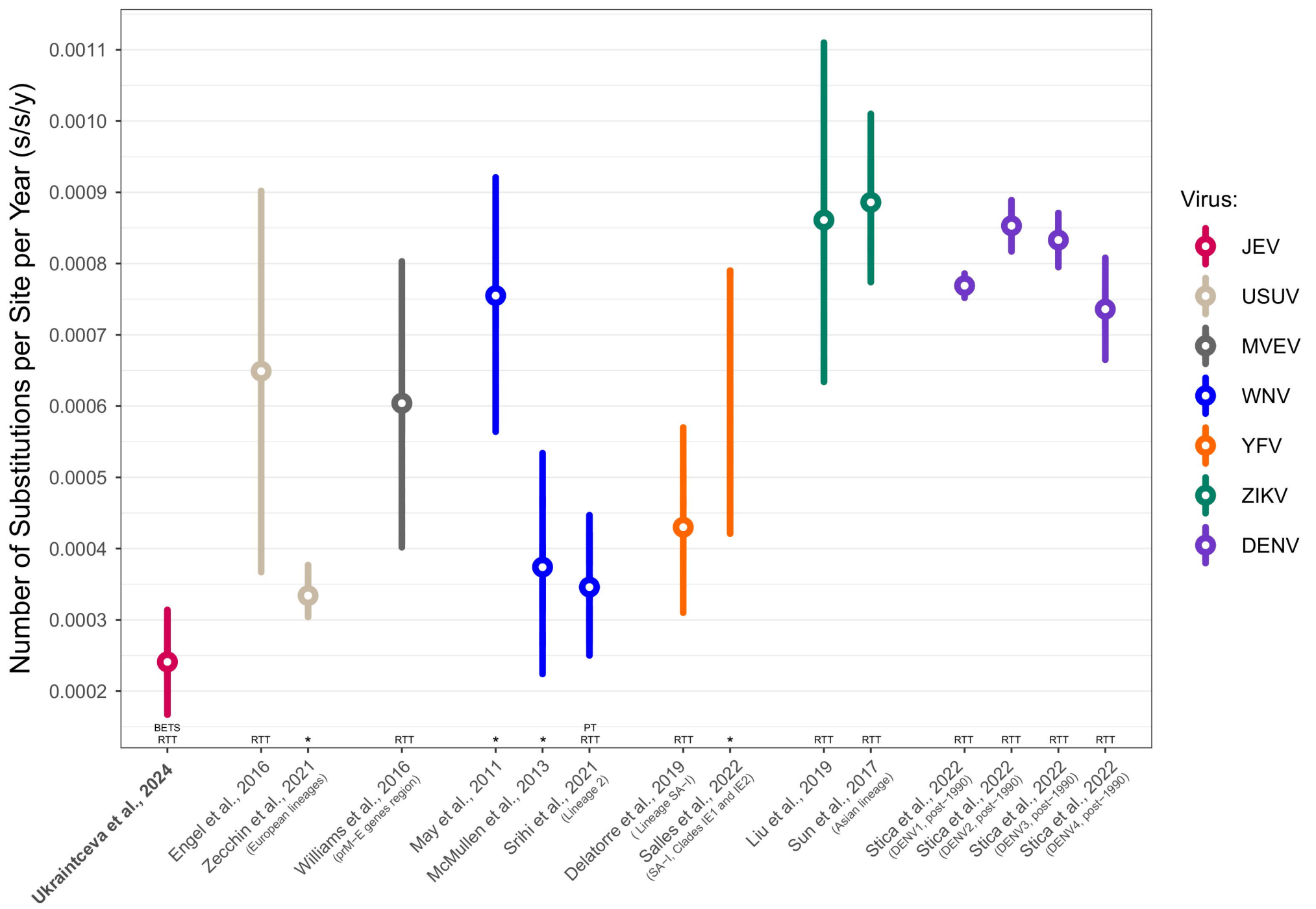

JEV GI-GV
R2=0.04
0.42
0.4
0.38
0.36
0.34
0.32
1940
1960
1980
2000
2020

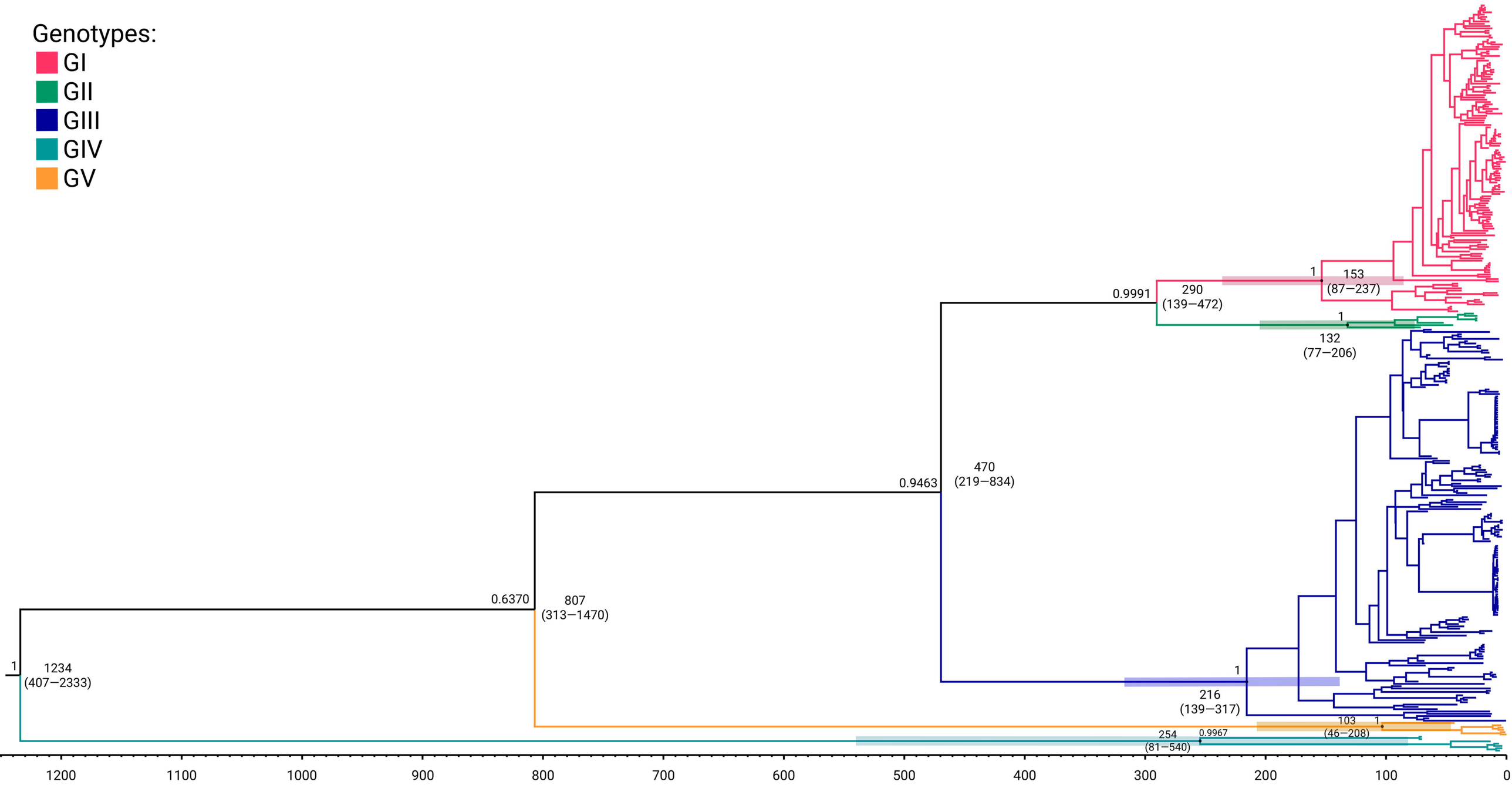

JEV GI
R2=0.79
0.050
0.045
0.040
0.035
0.030
0.025
0.020
0.015
0.010
1980
1990
2000
2010
2020


JEV GIII
R2=0.12
0.05
0.04
0.03
0.02
0.01
0
1940
1960
1980
2000
2020


Substitution rate, s/s/y
6 x 10-4
5 x 10-4
4 x 10-4
3 x 10-4
2 x 10-4
1 x 10-4
0
JEV GI-GV
JEV GI
JEV GIII


Number of Substitutions per Site per Year (s/s/y)
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
BETS
RTT
*
RTT
Ukraintseva et al., 2024
Mohammed et al., 2011
Pan et al., 2011
Gao et al., 2015
Xu et al., 2023
Li et al., 2023
Clade:
GI–GV
GI
GIII


Number of Substitutions per Site per Year (s/s/y)
0.0011
0.0010
0.0009
0.0008
0.0007
0.0006
0.0005
0.0004
0.0003
0.0002
BETS
RTT
RTT
*
RTT
*
*
PT
RTT
RTT
*
RTT
RTT
RTT
RTT
RTT
RTT
Ukraintceva et al., 2024
Engel et al., 2016
Zecchin et al., 2021
(European lineages)
Williams et al., 2016
(prM–E genes region)
May et al., 2011
McMullen et al., 2013
Srihi et al., 2021
(Lineage 2)
Delatorre et al., 2019
(Lineage SA–I)
Salles et al., 2022
(SA–I, Clades IE1 and IE2)
Liu et al., 2019
Sun et al., 2017
(Asian lineage)
Stica et al., 2022
(DENV1, post–1990)
Stica et al., 2022
(DENV2, post–1990)
Stica et al., 2022
(DENV3, post–1990)
Stica et al., 2022
(DENV4, post–1990)
Virus:
JEV
USUV
MVEV
WNV
YFV
ZIKV
DENV


Genotypes:
GI
GII
GIII
GIV
GV
1
1234
(407–2333)
0.6370
807
(313–1470)
0.9463
470
(219–834)
0.9991
290
(139–472)
1
153
(87–237)
1
132
(77–206)
1
216
(139–317)
103
(46–208)
1
254
(81–540)
0.9967
1200
1100
1000
900
800
700
600
500
400
300
200
100
0