

# 1 Demographic history and the efficacy of selection in 2 the globally invasive mosquito *Aedes aegypti*

3 Tyler V. Kent<sup>1,2</sup>, Daniel R. Schrider<sup>2</sup>, and Daniel R. Matute<sup>1</sup>

4

5 1: Department of Biology, College of Arts and Sciences, University of North Carolina, Chapel Hill, NC, USA

6 2: Department of Genetics, School of Medicine, University of North Carolina, Chapel Hill, NC, USA

7

8 \*Address correspondence: Daniel R. Schrider: [drs@unc.edu](mailto:drs@unc.edu) or Daniel R. Matute,

9 [dmatute@email.unc.edu](mailto:dmatute@email.unc.edu)

10

11 Running title: *Aedes aegypti* demographic patterns

## 12 **Abstract**

13 *Aedes aegypti* is the main vector species of yellow fever, dengue, zika and chikungunya. The  
14 species is originally from Africa but has experienced a spectacular expansion in its geographic  
15 range to a large swath of the world, the demographic effects of which have remained largely  
16 understudied. In this report, we examine whole-genome sequences from 6 countries in Africa,  
17 North America, and South America to investigate the demographic history of the spread of *Ae.*  
18 *aegypti* into the Americas its impact on genomic diversity. In the Americas, we observe patterns  
19 of strong population structure consistent with relatively low (but probably non-zero) levels of  
20 gene flow but occasional long-range dispersal and/or recolonization events. We also find  
21 evidence that the colonization of the Americas has resulted in introduction bottlenecks.  
22 However, while each sampling location shows evidence of a past population contraction and  
23 subsequent recovery, our results suggest that the bottlenecks in America have led to a  
24 reduction in genetic diversity of only ~35% relative to African populations, and the American  
25 samples have retained high levels of genetic diversity (expected heterozygosity of ~0.02 at  
26 synonymous sites) and have experienced only a minor reduction in the efficacy of selection.  
27 These results evoke the image of an invasive species that has expanded its range with  
28 remarkable genetic resilience in the face of strong eradication pressure.

29

## 30 Introduction

31 Invasive species pose a threat to native species and ecosystems as well as human health and  
32 agriculture (Chornesky and Randall 2003; Paini et al. 2016; Dueñas et al. 2021). While many  
33 invasive species spread by displacing or outcompeting native species, many instead take  
34 advantage of under-utilized niches, often similar to those in their home environments (Elton  
35 1958; Baker 1974; Sakai et al. 2003; Barrett 2015; Liu et al. 2020; Baker and Stebbins). One  
36 such pathway is developing an association with humans, either by exploiting anthropogenic  
37 changes to the ecosystem, or directly through the evolution of human preference (Hulme-  
38 Beaman et al. 2016). By adapting to these novel environments, invasive species face dramatic  
39 reductions in genetic diversity and the efficacy of selection, however they are nevertheless able  
40 to spread and outcompete native species, a long-studied conundrum dubbed the paradox of  
41 biological invasions (Sax and Brown 2000; Frankham 2005; Schrieber and Lachmuth 2017). For  
42 invasive species with human preferences, the shortcomings associated with the classic paradox  
43 of biological invasions may not hold true if they were adapted to human environments prior to  
44 their invasion, as the need to adapt to a novel environment may be greatly diminished or  
45 completely eliminated (Lee and Gelembiuk 2008; Hufbauer et al. 2012). The idea that some  
46 invasive species had previously adapted to humans and their environments is especially crucial  
47 for understanding the history and genomic consequences of invasive range expansions in  
48 human diseases and their vectors (Hufbauer et al. 2012; Powell 2019; Comeault et al. 2020).

49 Over the course of an invasion, populations of invasive species face dramatic  
50 demographic changes through their introduction and along their expansion front (Sakai et al.  
51 2003). These include introduction bottlenecks that reduce genetic diversity to an extent that  
52 depends on the severity of the bottlenecks, the number of bottlenecks and the diversity of  
53 alleles they sample, and the amount of gene flow across the species range (Nei et al. 1975;  
54 Dlugosch and Parker 2008; Blackburn et al. 2016). The expected reduction in effective  
55 population size, and subsequent loss of genetic diversity following the introduction of an  
56 invasive species will impact the species' ability to adapt (Barton and Partridge 2000; Brandvain  
57 and Wright 2016). During a bottleneck, the lower effective population size reduces the efficacy  
58 of selection relative to drift, allowing weakly deleterious mutations to drift to higher frequencies,  
59 and will also decrease the probability that beneficial mutations will fix in a population. Species  
60 facing anthropogenic selective pressures might experience such strong selection that  
61 adaptation can proceed in spite of reductions in genetic diversity; however, adaptation will still

62 be slowed if the introduced population's loss of diversity is severe enough that it will have to wait  
63 for new beneficial mutations (Orr and Unckless 2014; Osmond et al. 2019).

64 Importantly, the amount of gene flow between populations will also affect the efficacy of  
65 selection in contrasting ways: introducing new migrants to a population will increase the  
66 effective population size and thus the efficacy of selection, while the rate of influx of new  
67 mutations to a given locale can lead to a swamping effect, wherein selected alleles are  
68 overwhelmed by migration (Haldane 1956; Lenormand 2002; Yeaman 2022). For populations  
69 that are locally adapted, this gene flow is likely to primarily introduce maladapted alleles, but for  
70 populations with shared selective pressures like those from anthropogenic sources, gene flow  
71 may accelerate adaptation by introducing beneficial alleles that have arisen elsewhere in the  
72 global population that may not be present locally (Edelman and Mallet 2021; Yeaman 2022). In  
73 short, while theory predicts a decrease in the efficacy of selection and the rate of adaptation  
74 following a bottleneck and expansion, the degree to which this occurs in natural populations is  
75 less straightforward.

76 *Aedes aegypti*, the yellow fever mosquito, is a globally invasive species and the primary  
77 vector for the arboviruses dengue, yellow fever, chikungunya, and Zika. The species originated  
78 in Africa, where it spread during the African humid period, and differentiated into a generalist,  
79 *Ae. aegypti formosus* (hereafter *formosus*) and a human-adapted form, *Ae. aegypti aegypti*  
80 (hereafter *aegypti*) roughly 5,000 years ago (Brown et al. 2014; Rose et al. 2020; Rose et al.  
81 2023). The *aegypti* form shows a higher preference for human hosts and an increased tolerance  
82 for rainfall variation, likely as a response to its strong association with humans (Rose et al.  
83 2020). *aegypti* are relatively poor dispersers by their own flight, but their association with  
84 humans leads to long-distance dispersal, and high egg desiccation tolerance and dormancy  
85 means populations do not need to be immediately established, but may emerge years after their  
86 initial introduction (Machado-Allison and Craig 1972; Fischer et al. 2019; Mayilsamy 2019),  
87 possibly with diverse lineages from these egg banks (Kaj et al. 2001; Evans and Dennehy  
88 2005). *aegypti* has recently spread worldwide, first into the Americas shortly after European  
89 colonization, and later into Asia (Brown et al. 2014). Worldwide *aegypti* populations appear to  
90 harbour far less variation than populations of *formosus* or mixed ancestry in Africa, and  
91 potentially result from a single origin of human preference (Gloria-Soria et al. 2016; Lozada-  
92 Chávez et al. 2023). The expansion of *aegypti* out of Africa was likely complex, with high ship  
93 volume to the Americas during the human slave trade offering ample opportunities for multiple  
94 early introductions and continued gene flow (Powell et al. 2018). In recent decades, increasingly  
95 globalized trade may have allowed for additional long-distance dispersal events.

96           In the Americas, *aegypti* has had a substantial impact on human health and history  
97 (Tapia-Conyer et al. 2009; San Martín et al. 2010; Shepard et al. 2011; Cafferata et al. 2013).  
98 Outbreaks of yellow fever were first reported in the 17th century (Blake 1968; Bryan et al. 2004),  
99 although they may have begun earlier (Carter 1931), and continued outbreaks in the following  
100 centuries wreaked havoc on indigenous populations and naïve Europeans. Various epidemics  
101 of dengue, beginning in the 17th century (Brathwaite Dick et al. 2012), chikungunya, beginning  
102 in the 19th century (Brathwaite Dick et al. 2012), and more recently Zika in the 21st century  
103 (Chang et al. 2016) have occurred and continue to occur throughout the Americas, including the  
104 United States. During the 20th century, several countries in South America began an  
105 eradication program for *aegypti*, imposing intense pesticide pressure on populations, and by  
106 1962, 18 South American and Caribbean countries had reported eradication (American Health  
107 Organization 1997). Recolonization, recrudescence from unidentified populations, or a  
108 combination of both led to the reemergence of the vector across the continent (Kotsakiozi et al.  
109 2017). Insecticide resistance evolved rapidly and repeatedly throughout the Americas, with  
110 several haplotypes underlying resistance to multiple insecticides reported in the literature  
111 (Kawada et al. 2014; Al Nazawi et al. 2017; Haddi et al. 2017; Saavedra-Rodriguez et al. 2018;  
112 Fan et al. 2020; Love et al. 2023).

113           Introductions into the United States have been more recent, perhaps earliest in the  
114 Southeast. The first confirmed collection of the species was done in Savannah (GA) in 1828  
115 (specimen cataloged as *Culex taeniatus* Wiedemann; (Christophers 1960; Eisen and Moore  
116 2013) but earlier outbreaks of diseases consistent with yellow fever in Spanish Florida in 1649  
117 and the Northeast in 1668 suggest an earlier presence of the vector (Patterson 1992; Eisen and  
118 Moore 2013). These colonizations were then followed by subsequent spreads to the west, and  
119 multiple independent introductions reported in California with breeding populations first reported  
120 in the state in 2013 (Pless et al. 2017). Despite close monitoring of recent introductions and  
121 efforts to monitor and control populations throughout the Americas, previous genetic studies of  
122 the species have primarily focused on individual regions, high-level descriptions of population  
123 structure, and/or have used limited genomic information. Currently there lacks a detailed  
124 understanding of the demography of *aegypti* in North and South America, in particular the  
125 number and severity of *aegypti* introductions, the degree to which gene flow may aid in the  
126 spread and adaptation of populations throughout the continents, and how the efficacy of  
127 selection in introduced populations of *aegypti* has been affected by their demographic history.

128           Here we examine a set of 131 whole genomes from African and American *Ae. aegypti*  
129 populations to model the demographic history of the species' spread into the Americas, and

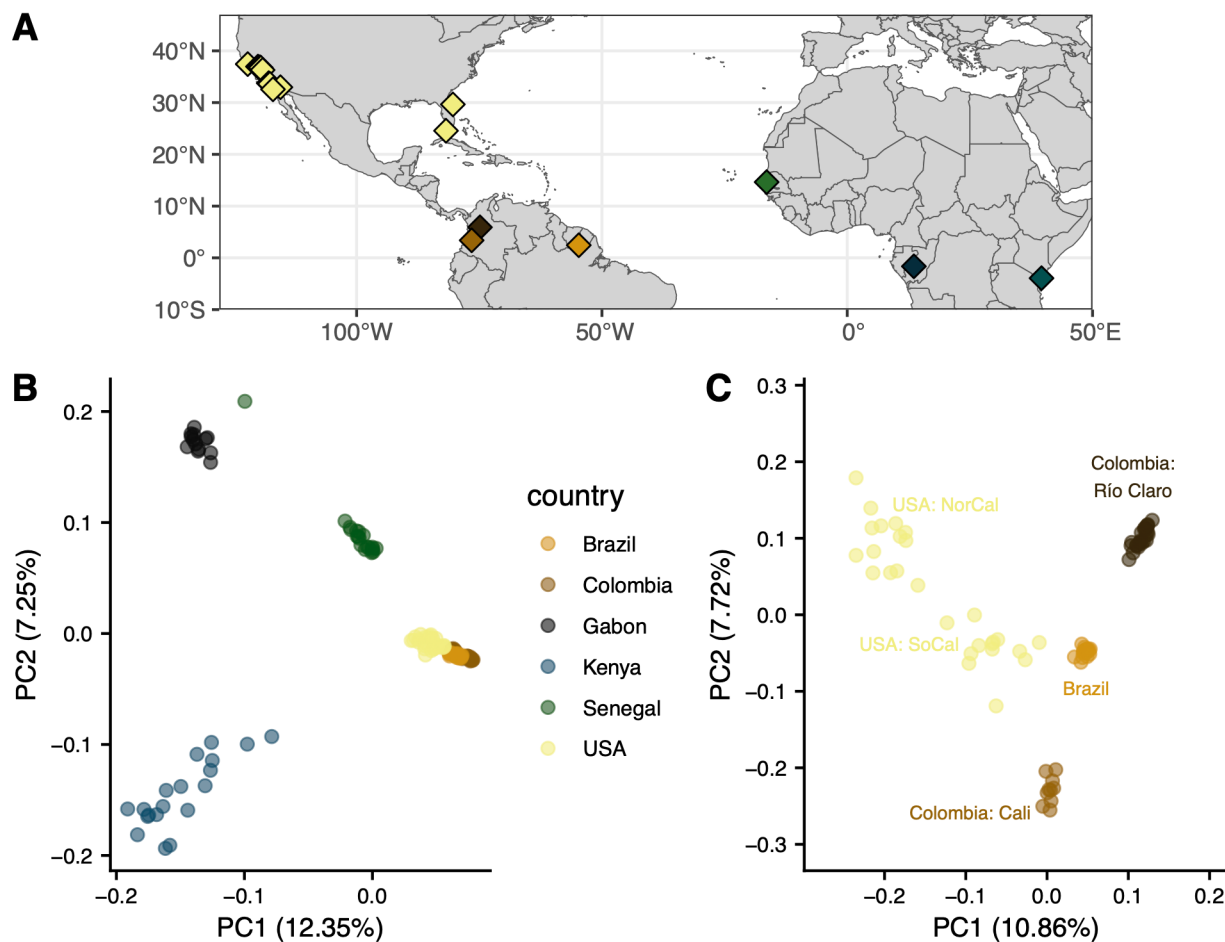
130 investigate its effects on genome-wide diversity and the efficacy of selection. First we  
131 investigate patterns of population structure across our dataset, within and among Africa and the  
132 Americas. We then infer the history of effective population size changes and split times between  
133 populations within and between the Americas and Africa. Finally, we investigate the efficacy of  
134 selection of all populations by examining levels of diversity genome-wide and in protein-coding  
135 genes and by inferring the distribution of fitness effects of new mutations. We find that the  
136 spread of *Ae. aegypti aegypti* to the Americas appears to be best characterized by multiple  
137 apparent introductions and limited subsequent gene flow at fine and coarse geographic scales,  
138 and that the history of bottlenecks and strong selection in the introduced range has had  
139 genome-wide impacts on diversity and the efficacy of selection. Despite a 33-40% reduction in  
140 neutral genetic diversity, *aegypti* in the Americas maintains high diversity and appears to have  
141 experienced only a modest reduction in the efficacy of selection. The limited effect of its  
142 introduction history and recent anthropogenic selection on diversity and the efficacy of selection  
143 illustrates a surprisingly resilient species at the genomic level—one that poses a threat to future  
144 eradication efforts. We discuss the implications of this history and its genomic impact in the  
145 context of invasive species and vector control, and adaptation (past, ongoing, and future) in the  
146 species.

## 147 **Results and Discussion**

148 We used publicly available sequencing data from several studies, recently collated in (Love et  
149 al. 2023). This dataset includes 27 specimens from California and Florida (Lee et al. 2019), 18  
150 from Santarem, Brazil, 13 from Franceville, Gabon, 19 from Kaya Bomu, Kenya, and 20 from  
151 Ngoye, Senegal (Rose et al. 2020), and 10 from Cali and 24 from Río Claro, Colombia (Love et  
152 al. 2023), for a total of 131 specimens, 79 of which are from the Americas. Hereafter we refer to  
153 specimens sampled from the same location as accessions, and to groups inferred to show  
154 evidence of recent shared ancestry as populations or clusters, depending on context. The  
155 African accessions used here were previously scored for human preference, with the Gabon  
156 accession exhibiting little-to-no human preference, the Kenya accession exhibiting  
157 intermediate/mixed preference, and the Senegal accession exhibiting strong human preferences  
158 (Rose et al. 2020), while the American accessions are assumed to exhibit strong human  
159 preferences.

160 *Aedes aegypti* is highly structured at fine and coarse scales

161 **Figure 1:** (A) Sampling locations of specimens used in this study. (B) Principal component



162 analysis (PCA) of all specimens highlighted by country of origin. (C) PCA of American samples  
163 highlighted by location.

164  
165  
166 To characterize the population structure of our samples, we first used principal component  
167 analysis, beginning with our full dataset. The first and second principal components,  
168 representing 12.35% and 7.25% of the variance, respectively, separate the African accessions  
169 from the Americas and from each other (Figure 1B; scree plot in Supplementary Figure 1). The  
170 African accessions separate distinctly from each other, with more apparent diversity in Kenya  
171 than in either Gabon or Senegal. The American samples cluster together, with the three South  
172 American accessions and the USA accessions clearly differentiated within this regional cluster.  
173 Among the African samples, Senegal appears to be most closely related to the American



174 samples, and particularly to samples from the US, consistent with previous studies identifying  
175 Senegal as a proxy for the ancestral *aegypti* form, and with a more recent invasion in the US.  
176 However, one Senegal specimen appears to cluster near to the Gabon cluster, which has been  
177 previously noted (Rose et al. 2020). This specimen could represent a recent migrant from a  
178 population closely related to the Gabon accession, or a generalist individual within a larger-than-  
179 recognized range of *formosus*. PC1 is consistent with separating accessions by either latitude or  
180 host preference, and interestingly, PC2 primarily separates Kenya from Gabon and Senegal.  
181 Kenya and Gabon both exhibit primarily low human preference, with the former composed of  
182 mixed preference, evident here with the spread of the Kenyan sample along both PCs, with  
183 some specimens closer in PC-space to American samples.

184         Within the Americas, principal components reveal more complex structure (Figure 1C;  
185 scree plot in Supplementary Figure 2). Accessions from Colombia and Brazil separate clearly,  
186 with Brazil closer to either Colombian accession than the Colombian accessions are to each  
187 other, as previously reported (Love et al. 2023), despite being relatively close in geographical  
188 space (Figure 1A); indeed, despite being separated by only ~500km, these two accessions are  
189 found nearly on opposite ends of PC2. As may be expected from geographically spread  
190 samples with few from a single location, the samples from the US are loosely clustered into two  
191 groups in the first two PCs, and three to four groups apparent in deeper PCs (Supplementary  
192 Figure 3), similar to what was previously reported with these samples (Lee et al. 2019). The first  
193 US cluster in PC 1 and 2 groups Northern California and some accessions from Central  
194 California in the upper left of Figure 1C. The second cluster, closer to Brazil in the centre of  
195 Figure 1C, consists of Southern California, a subset of accessions from Central California, and  
196 Florida. In PCs 3 to 5, Central and Northern California clearly separate into groups: a cluster of  
197 Menlo Park with Madera and Fresno, and a cluster of Clovis and Sanger (Supplementary Figure  
198 3). This separation is remarkable in that Clovis and Sanger are roughly 250 km apart, and on  
199 either side of Fresno with which they do not cluster. The Southern California and Florida cluster  
200 is largely maintained at PCs 3-5, with only the Florida sample from Vero Beach separating  
201 beginning at PC 4. The accessions from Exeter, California, located in the Central Valley south of  
202 Fresno, cluster closely with an accession from Key West, Florida, and near to the Southern  
203 California accessions. Similar to the Northern California cluster, structuring within the primarily  
204 Southern California cluster is hallmarked by a combination of fine and coarse scale geographic  
205 separation, and can likely be clustered into two to three groups within California, and two  
206 separate Florida accessions.



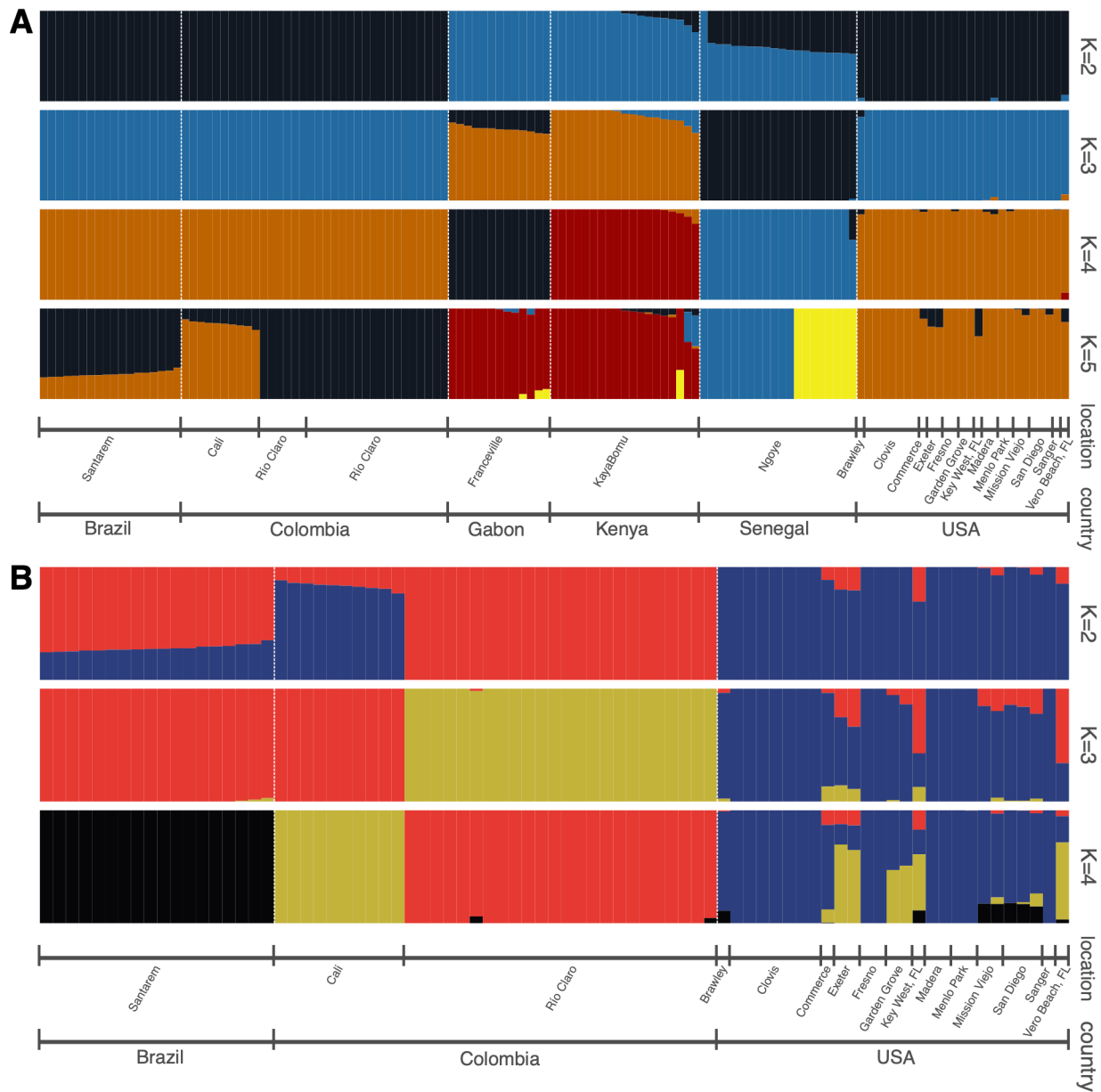
207 To more explicitly model population structure within our dataset, we used ADMIXTURE.  
208 ADMIXTURE clusters samples using allele frequencies into a pre-defined number of groups,  
209 modeling each individual as a mixture of the clusters. Using the whole sample set, the best  
210 number of clusters according to cross validation is  $K=2$ , where samples are largely split into  
211 African and American groups (Figure 2A). Several specimens in Kenya and nearly all in  
212 Senegal exhibit substantial admixture proportions, with the major component grouping with  
213 Gabon and the minor component with the Americas. All specimens from Gabon are entirely  
214 assigned to the same group (the major ancestry component in Africa), while the Brazilian and  
215 both Colombian accessions are all completely assigned to the opposing group—the minor  
216 ancestry component of Kenya and Senegal. Specimens from the US are nearly entirely  
217 composed of ancestry shared with South America and the minor ancestry component of Kenya  
218 and Senegal, though three specimens, one of each from Southern California, Central California,  
219 and Vero Beach, Florida, show small amounts of the Gabon-like ancestry.

220 Mixed ADMIXTURE group membership in Kenya and Senegal is consistent with  
221 domestic ancestry or human preference previously reported in these accessions. Senegal, apart  
222 from one unadmixed specimen assigned entirely to the Gabon cluster, varies between roughly  
223 50:50 and 70:30 split of Gabon-like to South American-like group memberships, while about half  
224 of Kenyan samples have a small amount of South American-like ancestry. Higher values of  $K$ ,  
225 although a slightly poorer fit according to cross validation (Supplementary Table 1), provide  
226 some context for the minor ancestry components in Senegal and Kenya. At  $K=3$  and above  
227 (Figure 2A), the minor ancestry in Kenya groups with the American accessions, while Senegal  
228 largely appears unadmixed and groups with some ancestry in Gabon, or with itself, until  $K=5$   
229 where it splits into two groups shared in minor proportions with some specimens in Gabon and  
230 Kenya. This is concordant with Senegal grouping most closely to the American accessions in  
231 PCA, but some Kenyan samples being nearly equidistant. While the extant accession in  
232 Senegal may represent the nearest known representative of an ancestral population to the  
233 American samples, these results from PCA and ADMIXTURE are consistent with ancestral  
234 population structure resulting ancestry shared among some African populations, including some  
235 domestic *aegypti* haplotypes that spread to the Americas (Slatkin and Pollack 2008). Back  
236 migration from the Americas to African populations (Brown et al. 2011; Powell and Tabachnick  
237 2013) has also been proposed and would be consistent with low levels of American-like  
238 ancestry in Kenya during a second wave of the shift to human preference in African cities (Rose  
239 et al. 2020).

240           Within the Americas, the best number of groups according to cross validation is again  
241  $K=2$  (Figure 2B), and the patterns of admixture are identical to the patterns in the Americas  
242 when considering all samples at  $K=5$  (Figure 2A). At  $K=2$ , accessions from Brazil and from Cali,  
243 Colombia are shown as admixed, with reciprocal minor ancestry components, while Río Claro,  
244 Colombia is unadmixed of the major ancestry component in Brazil. Greater grouping between  
245 Brazil and Río Claro is consistent with results from PCA (Figure 1C) and divergence statistics  
246 (Supplementary Figures 4 and 5 for genome-wide  $F_{ST}$  and  $D_{xy}$ ) showing greater divergence  
247 within Colombia than between Río Claro and Brazil. The US is represented as partially admixed  
248 and primarily composed of the major ancestry component in Cali. The admixture in the US is  
249 isolated to the previously mentioned cluster present in the first two PCs of Southern California,  
250 Florida, and the Central California accession of Exeter (Figure 1C), and is present in some but  
251 not all of these individuals at  $K=2$ , varying from minor to sizable proportions. At higher values of  
252  $K$ , population structure remains qualitatively similar, with additional groupings adding to the  
253 diversity of admixture within the admixed US specimens, while the remaining US specimens  
254 compose their own group and South American accessions become unadmixed and quickly split  
255 among each other. This coarse-scale separation again reflects the two major clusters in the  
256 Americas, with little effect of geographic distance as evident in the separation within Colombia  
257 and the mixed fine and coarse-scale structure in the US.

258           In summary, both ADMIXTURE and PCA results from the Americas depart strongly from  
259 expectations under a simple model of geographic isolation by distance. This may lend support  
260 to the possibility that there have been several independent introductions to the Americas, each  
261 sampling a subset of the ancestral *aegypti* variation available in Africa. We examine this  
262 possibility more directly in the section below.

263

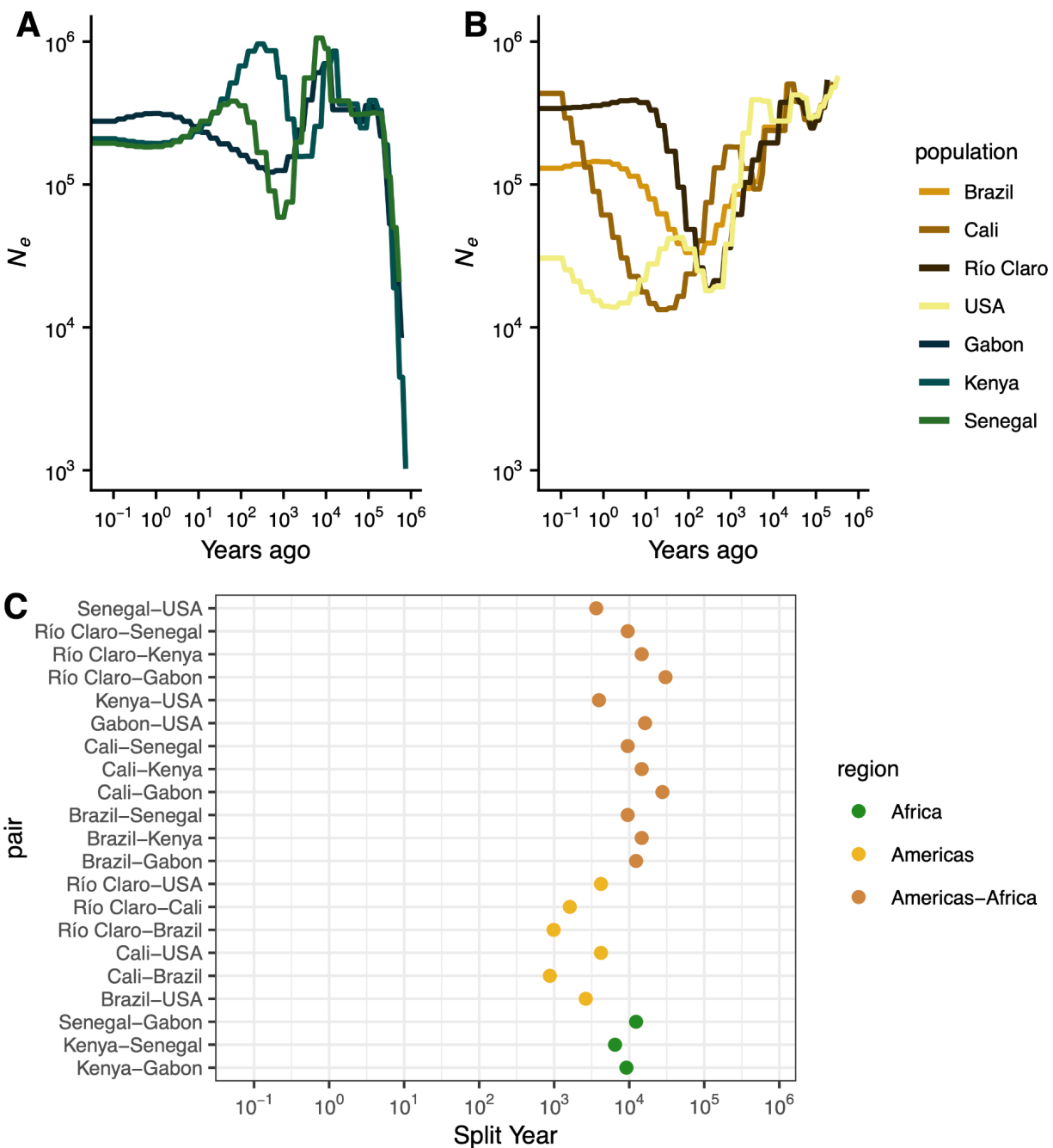


264  
 265 **Figure 2:** Admixture results for global samples (A) and for the Americas (B). The best  $K$ , obtained  
 266 via cross validation, is  $K=2$  for both (A) and (B), but we show results for higher  $K$  for additional  
 267 context.

268

269 American *aegypti* underwent multiple, strong bottlenecks

270



271

272

273

274

275

276 We sought to investigate the number and timing of introductions of *Ae. aegypti* among our

277 sampled accessions, which represent a range of events spanning from among the earliest to the

278 most recent proposed introductions to the Americas. We began by inferring the historical

279 effective population sizes among all countries using SMC++, making use of recent estimates for  
280 the generation time and mutation rate of *Ae. aegypti* (Rose et al. 2023). We report results  
281 obtained using the same model regularization parameter ( $rp=5$ ), which provided the most  
282 consistent fit across the seven groups examined in Figure 3 as evidenced by coalescent  
283 simulations (Supplementary Figures 8–14). These analyses suggest that the ancestral effective  
284 population size ( $N_e$ ) among all accessions must have been between 300,000 and 400,000 until  
285 about 20,000 years (300,000 generations) ago (Figure 3A & B). All accessions experienced a  
286 reduction in  $N_e$  around 10,000 years (150,000 generations) ago, coinciding with the end of the  
287 African humid period as previously reported (Rose et al. 2023). Kenya and Gabon recovered to  
288 near ancestral  $N_e$  over a span of a few thousand years, while Senegal appears to have suffered  
289 a deeper bottleneck than the other African accessions, dropping to roughly 80,000, with the  
290 lowest  $N_e$  point around 1,000 years (15,000 generations) ago, before recovering to about  
291 200,000. This bottleneck in Senegal occurred between 5,000 and 10,000 years ago, consistent  
292 with the founding and diversification of the *aegypti* domestic ancestry within Africa in the dry,  
293 variable Sahel after the African humid period (Rose et al. 2023). As an alternative approach to  
294 estimating population size histories, we also ran Stairway Plot 2 (Methods), and obtained  
295 qualitatively similar patterns as those obtained via SMC++ in terms of the differences relative  
296 severity of bottlenecks across accessions (Supplementary Figure 6).

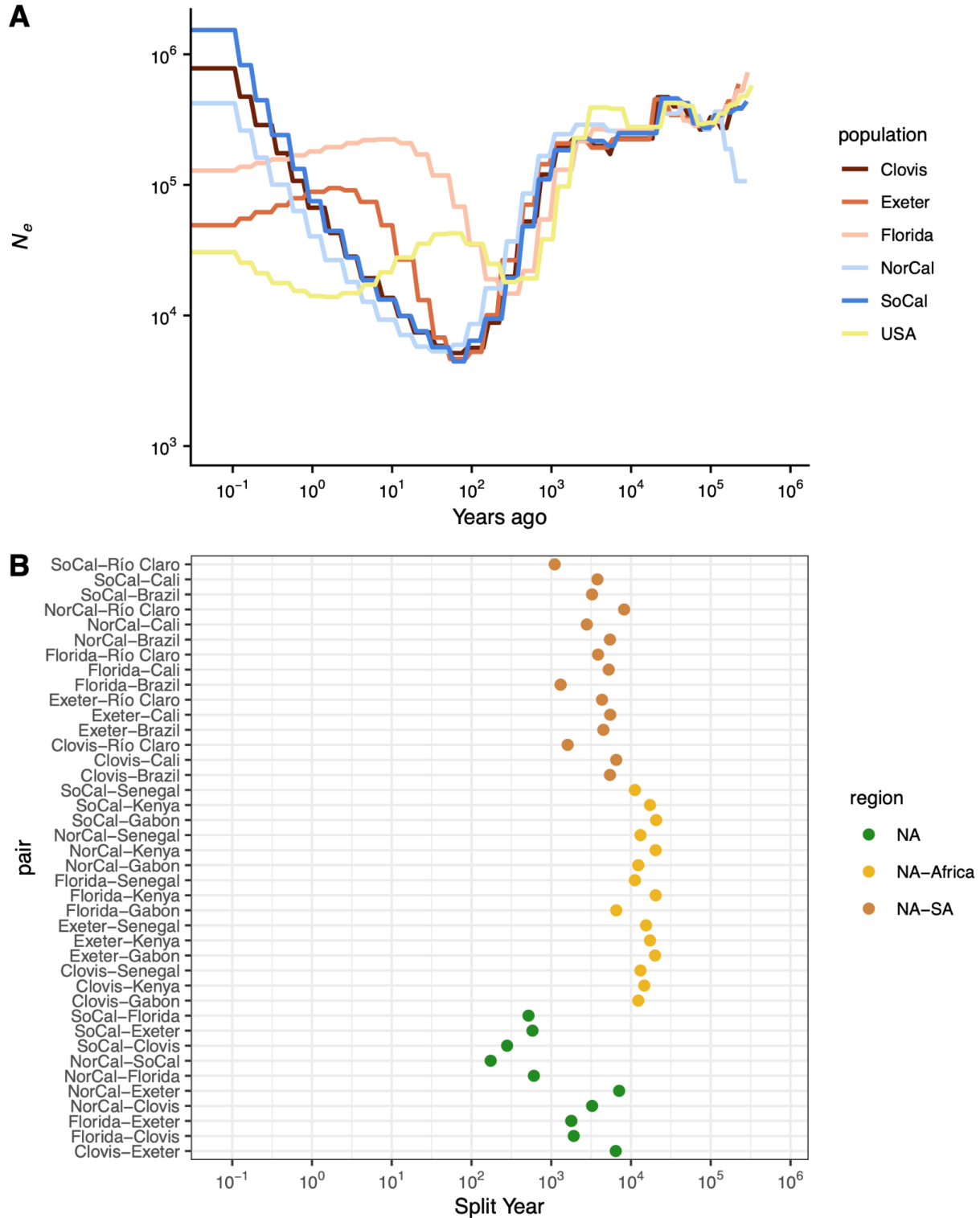
297         Among the American accessions, we infer strong, prolonged bottlenecks that vary in  
298 their timing and intensity. All American accessions began to decline between 5,000 and 10,000  
299 years ago, and continued to decline for thousands of years (Figure 3B). In the accession from  
300 Cali, Colombia, we infer a fluctuating  $N_e$  that was approximately stable until about 700 years  
301 ago, when its introduction bottleneck started and continued to decline until it reached a low  $N_e$  of  
302 13,312 only 20 years (300 generations) ago, followed by a recovery. The relatively larger  $N_e$  of  
303 the Cali accession that is inferred at the time when other American accessions began to exhibit  
304 a contraction may suggest a more diverse sampling of lineages, possibly through multiple  
305 introductions or migration after colonization; both of these possibilities would result in a period of  
306 decelerated coalescence and are consistent with the higher degree of admixture inferred in  
307 Figure 2B for this accession than the other South American accessions. In Río Claro and  
308 Brazil, bottlenecks beginning after the Africa humid period continued until their lowest  $N_e$  points  
309 of 18,639 and 33,225 occurred 250 and 150 years (3,750 and 2,250 generations) ago,  
310 respectively. In all three accessions we infer massive recoveries, to a present  
311 day  $N_e$  of 433,679, 340,611, and 129,800 in Cali, Río Claro, and Brazil, respectively. Methods  
312 based on the sequentially Markovian coalescent like SMC++ have limited ability to accurately

313 estimate the  $N_e$  in very recent time, however the qualitative result of rapid recoveries in these  
314 accessions occurs within the span of generations in the past where these methods have  
315 reasonable accuracy (Patton et al. 2019), at least in Río Claro and Brazil, and is consistent with  
316 monitoring and disease prevalence suggesting a rebound after the collapse of Pan-American  
317 eradication efforts (reviewed in (Webb 2016)), potentially through recolonization (Monteiro et al.  
318 2014). While the bottlenecks in each locality are inferred to have occurred at different times over  
319 a span of roughly 200 years, we cannot be confident that they all represent different events;  
320 however, we can be confident that each of the South American accessions experienced strong  
321 bottlenecks and recoveries in the very recent past. Introduction bottlenecks in the South  
322 American accessions resulted in reductions in  $N_e$  of 94%, 89%, and 96% in Río Claro, Brazil,  
323 and Cali, respectively, which appear to have occurred no later than 4,000 generations or 250  
324 years ago.

325 As evidenced by the PCA and ADMIXTURE results, the United States accessions  
326 appear to represent a complex structuring of populations. However, grouping the US accessions  
327 into a single population when running SMC++ results in reductions in inferred  $N_e$  similar to those  
328 in Cali, pointing to substantial bottlenecks in the underlying structured populations. We therefore  
329 split the US into five approximate populations representing Southern California, Northern  
330 California, Florida, Clovis and Sanger, CA, and Exeter, CA, following clustering in PCA space  
331 (Supplementary Figure 3) and results from previous studies (Lee et al. 2019). All populations  
332 experienced a substantial bottleneck in the recent past, all beginning roughly 10,000 years ago  
333 as in the South American populations, but all peaking at a similar  $N_e$  of about 5,000 only 55  
334 years (825 generations), except for Florida which peaked about 240 years ago at about 15,000  
335 (Figure 4A). All populations similarly are inferred to have recovered to between 400,000 and  
336 1,500,000 at the present day, except for Florida and Exeter which are inferred to have  
337 plateaued 5 to 10 years ago around 100,000, although this lower recovery may be an artifact of  
338 having only two samples in each of these populations. Similarly, the samples from Florida  
339 represent divergent lineages (Supplementary Figure 3), which may inflate  $N_e$  estimates and bias  
340 the timing of bottlenecks. Among California clusters, the timing of the start and peak of each  
341 bottleneck is similar, though the Northern California cluster is slightly shifted to the more recent  
342 past, consistent with previous reports of a later, independent introduction relative to Southern  
343 California (Lee et al. 2019). As a whole, US populations experienced deeper and more recent  
344 bottlenecks than the South American populations, consistent with later introductions, and an  
345 earlier introduction in Florida than in California, though the differences in  $N_e$  between the US  
346 clusters are minimal. The apparent earlier introduction in Florida with a lowest  $N_e$  roughly 240

347 years ago is after historical records of diseases matching the symptoms of Yellow Fever first  
348 appeared in Spanish Florida and the Northeast (Patterson 1992; Eisen and Moore 2013), but  
349 shortly before reports of the species in Georgia (Christophers 1960; Eisen and Moore 2013) and  
350 dengue outbreaks throughout port cities in the Eastern US (Brathwaite Dick et al. 2012).  
351





352

353

354

355

356

**Figure 4:** (A) Effective population size histories for approximate USA populations, with the estimated history of the combined accessions from the previous plot. (B) Split times for all approximate USA populations with all other accessions.

357 We also used SMC++ to infer population splits between all pairwise accessions. This  
358 model infers a split time using the within- and cross-population coalescent rates assuming a  
359 clean split model. At a broad scale, we find that split times involving African samples are older  
360 than split times within the Americas, though split times between Africa and the Americas span a  
361 range of time older and more recent than splits within Africa (Figure 3C). Among the African  
362 accessions, Senegal and Kenya are inferred to be most closely related, followed by Kenya and  
363 Gabon, and Senegal and Gabon. This order of population splits is consistent with previously  
364 published behavioural results (Rose et al. 2020), and support the notion that the Senegalese  
365 and a portion of the Kenyan accessions appear to represent a derived domestic *aegypti*  
366 ancestry, while the Gabon accession have maintained the ancestral generalist *formosus* form.  
367 The split times among the African accessions span a few thousand years, from 6,500 years ago  
368 for Kenya-Senegal, to 9,175 for Kenya-Gabon, and 12,300 years ago for Senegal-Gabon.  
369 Assuming a single origin for human preference, mixed *aegypti* and *formosus* ancestry in the  
370 Kenyan accession would likely be the result of recent migration from an *aegypti* population more  
371 closely related to Senegal. In a clean split model not accounting for this recent migration, the  
372 estimated split time would be closer to the present, rather than correctly inferring an older split  
373 followed by recent migration. Alternatively, as previously mentioned, the domestic ancestry in  
374 these accessions may have sampled different diverse haplotypes in the ancestral domestic  
375 population, yielding the same signal, consistent with ADMIXTURE results.

376 Between Africa and the Americas, we infer a broad range of population split times.  
377 Among the South American accessions, we infer nearly identical split times in increasing order  
378 to Senegal (split times between the three South American accessions and Senegal were each  
379 estimated to be 9,545 years ago), Kenya (split times all estimated to be 14,678 years ago), and  
380 Gabon (27,699 years to Cali and 30,479 years to Río Claro), except for the Brazil-Gabon split  
381 which is estimated to be slightly more recent than the Brazil-Kenya split (12,343 years ago;  
382 Figure 3B). These split times far predate their introduction to the Americas, and are on the order  
383 of or older than the splits within Africa, suggesting a deep split between extant African  
384 populations and the ancestral domestic population that eventually invaded the Americas. In  
385 contrast, split times between the US and Africa are much more recent, though still far predating  
386 the American introduction (3,636 years ago with Senegal, 3,953 years ago with Kenya, and  
387 16,256 years ago with Gabon); aside from a more recent split, this result could also be  
388 explained by migration between the African populations and the lineages leading to the US  
389 accessions after the split between the South American and African samples.

390           Within the Americas, estimated split times vary by thousands of years. In South America,  
391 our ordering of estimated population splits is consistent with results from PCA: Cali and Brazil  
392 most recently split (875 years ago), followed by Río Claro and Brazil (983 years ago), and finally  
393 Río Claro and Cali (1,613 years ago). These splits are well before their introductions to the  
394 Americas, and differences among them will reflect differential inheritance of ancestral variation  
395 and any migration after their introduction, but strongly suggest independent introductions to  
396 each location. Split times between South America and the US are deeper: estimated splits  
397 between the US and both Río Claro and Cali are identical (4,203 years ago), while the split  
398 between the US and Brazil is more recent (2,636 years ago). Again, the differences between  
399 these split times may reflect ancestral sorting, and the extent of recent migration (if any).  
400 However, in this case our estimates may also be affected by the presence of population  
401 structure in the US.

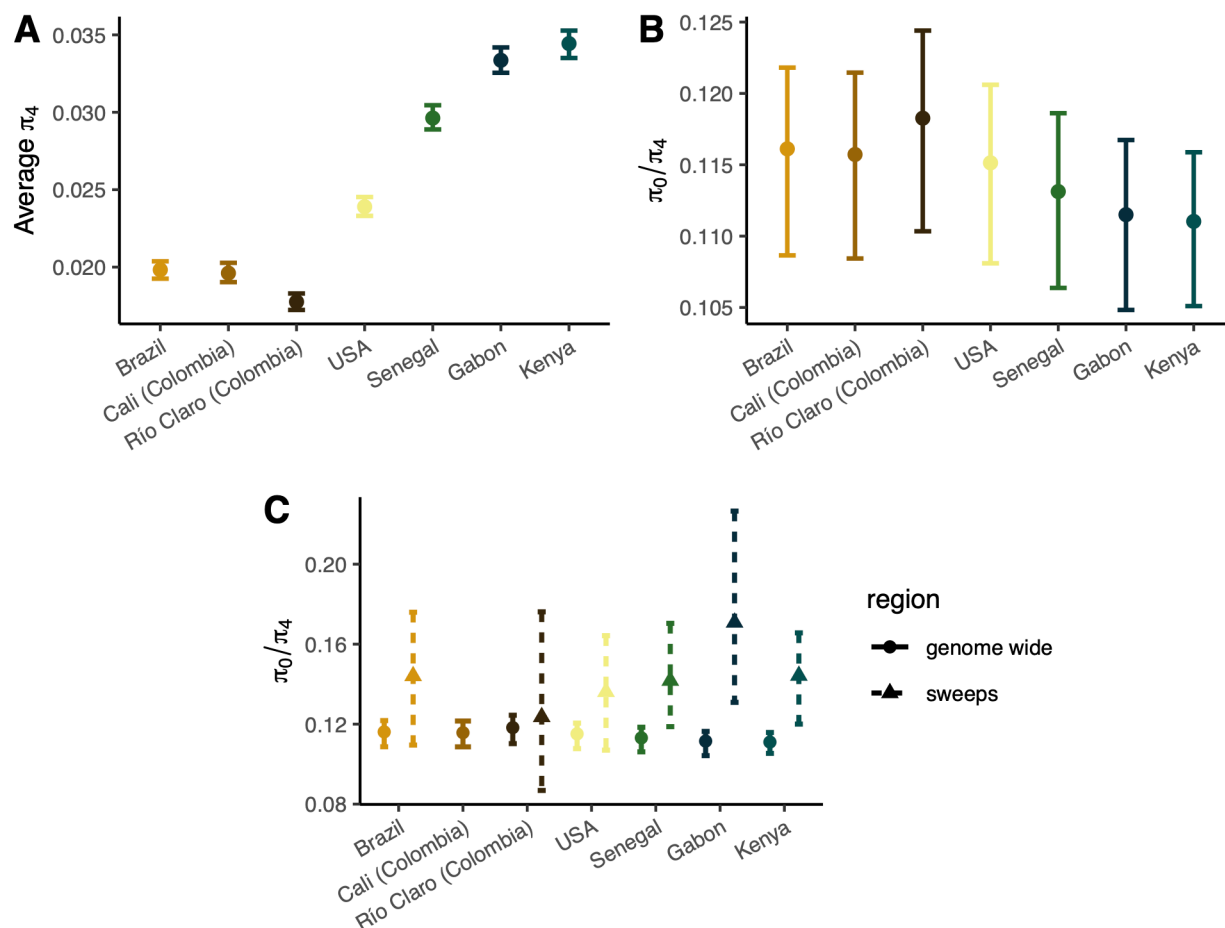
402           With the potential impact of population structure in mind, we also considered split times  
403 for five different clusters of the US accessions to all other accessions (Figure 4B). When  
404 considering these US partitions separately, the estimated split times are much deeper than  
405 when considering the US as a whole. The splits between US clusters and African accessions  
406 are slightly older than those estimated between South American and African accessions.  
407 However, the split between Florida and Gabon is estimated to be substantially more recent at  
408 6,536 years ago, potentially related to the more recent splits between Brazil and Gabon (Figure  
409 3B) and Florida and Brazil (Figure 4B). For most US clusters compared to Cali and Brazil, the  
410 estimated split times are 4,500 to 6,500 years ago; however, Florida, Southern California, and  
411 Northern California have more recent split estimates, driving the US splits to these accessions  
412 as a whole towards the present. Southern and Northern California both have estimated splits to  
413 Cali about half as long ago as the rest of the US (3,791 and 2,790 years respectively), and  
414 Southern California also has a more recent estimate to Brazil (3,250 years ago). Florida stands  
415 out with a much more recent split estimate to Brazil (1,310 years ago), more recent even than  
416 its split times to Exeter and Clovis. While these estimates are much older than reasonable  
417 introduction times to the Americas, they may reflect a similar ancestral source population, or an  
418 unsampled South American source population for Florida related to the accession presented  
419 here, especially in concordance with the high levels of admixture observed in the Floridian  
420 samples (Figure 2B). In total, the estimated split times between the US partitions and those in  
421 South America do not suggest an origin for the sampled US accessions in related South  
422 American populations.

423            Within the US, estimated split times again suggest a complex introduction history (Figure  
424 4B). There is a strong stratification in split times between a few US clusters—Florida and  
425 Northern California both have deep split times to the Central California accessions in  
426 Clovis/Sanger and Exeter, similar to the deep split between Clovis/Sanger and Exeter. These  
427 deeper splits span a broad range of time (1,785 to 7,100 years ago) with Florida more closely  
428 related to Clovis/Sanger and Exeter, while Northern California and Exeter is the deepest split.  
429 While deep splits between Florida and Central California may not be surprising, and are  
430 consistent with independent introductions to California and Florida, the deep splits between  
431 some of the more geographically close California samples are more unexpected. For example,  
432 the accessions of Clovis/Sanger and that of Exeter are separated by about 80km (split time  
433 estimate: 6,444 years ago), and Clovis/Sanger and Fresno (part of the Northern California  
434 cluster), are only separated by 16km (split time estimate: 3,255 years ago). These findings  
435 suggest that there has been long distance dispersal of related haplotypes within California,  
436 along with multiple introductions of lineages to several cities in the state that appear to have  
437 experienced little-to-no gene flow since their introduction and up until the time of sampling.

438            Among the more closely-related set of accessions in the US, estimated split times leave  
439 their origins less clear. The most recent split is between Northern and Southern California at  
440 174 years ago, suggesting either an introduction from elsewhere within the Americas for both  
441 sets of accessions, as has been suggested at least for the Southern California populations (Lee  
442 et al. 2019), or a split somewhere within the Americas and subsequent dispersal to different  
443 regions. Southern California and the Clovis accession are estimated to have split similarly  
444 recently at 280 years ago, again suggesting they may have originated through a split from a  
445 population already within the Americas, or possibly long-distance migration along the California  
446 Highway 99 corridor. The remaining splits, between both Northern and Southern California and  
447 Florida, and Southern California and Exeter occur within a timespan ranging from 519 to 606  
448 years ago. While these timeframes are older than a likely introduction, they remain much more  
449 recent than splits within South America or between American and African accessions,  
450 suggesting divergence within the Americas is not unlikely, but also consistent with separate  
451 introductions followed by gene flow from the same source population.

452 Introduced populations suffer minimal reductions in selection efficacy  
453 despite substantially reduced diversity

454

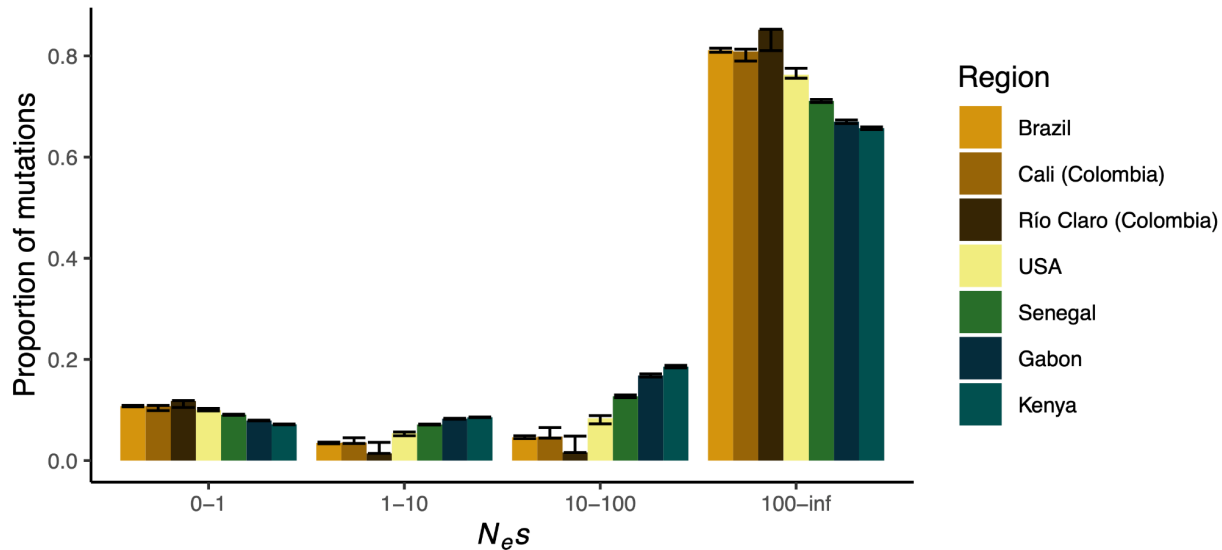


455 **Figure 5:** Diversity and measures of the efficacy of selection. (A) neutral diversity,  $\pi_4$ . (B)  $\pi_0/\pi_4$   
456 with bootstrapped errors for all accessions. American accessions show a subtle decrease in the  
457 efficacy of selection. (C)  $\pi_0/\pi_4$  in regions around sweeps (note that Colombia was combined for  
458 sweep scans, so is reported as a combined value for Río Claro and Cali, and as such, estimated  
459 neutral diversity is likely overestimated in the sweep regions, yielding an artificially lower  $\pi_0/\pi_4$ ).  
460

461  
462 Given the strong bottlenecks experienced by all sampled introduced accessions, we sought to  
463 quantify the impact of the introduction history and previously described history of recent strong  
464 selection (Love et al. 2023) on genome-wide levels of diversity. We first estimated genome-wide  
465 neutral diversity using 4-fold degenerate sites ( $\pi_4$ ) (Figure 5A). Among African accessions,

466 neutral diversity ranged from about 0.03 in Senegal to nearly 0.035 in Kenya. This level of  
467 diversity is similar to that reported in *Anopheles gambiae* (Corbett-Detig et al. 2015) and  
468 somewhat higher than previous genomic estimates of all sites (Rose et al. 2020; Love et al.  
469 2023). As expected given the demographic history of the accessions, neutral diversity in the  
470 introduced range was much lower than in the ancestral range: nearly 0.02 in Brazil and Cali,  
471 Colombia, and ~0.018 in Río Claro, Colombia. Diversity in the combined USA set of accessions  
472 was still less than 0.025, despite being inflated by population structure. While a 33-40%  
473 reduction in diversity relative to Senegal is substantial, it is perhaps less of a reduction than  
474 expected given the large bottlenecks and eradication efforts sustained by these accessions. In  
475 an absolute sense and relative to other species (Leffler et al. 2012; Buffalo 2021), neutral  
476 diversity near 0.02 in the Americas is high, posing potential problems for future control efforts.

477 The combined effect of the introduction bottleneck(s) and recent selective pressures in  
478 the American *Ae. aegypti* populations should feed back into its ability to adapt because both of  
479 these phenomena are expected to reduce the efficacy of selection. We sought to test this using  
480 two indirect measures of the efficacy of selection. First, we estimated the ratio of 0-fold to 4-fold  
481 diversity in each accession (Figure 5B). Assuming negative selection against new deleterious  
482 mutations is persistent, an increase in the  $\pi_0/\pi_4$  ratio reflects a reduction in the efficacy of  
483 selection, i.e. an increase in the proportion of effectively neutral mutations as seen by selection  
484 has allowed weakly deleterious mutations to accumulate. We find African accessions to have a  
485 relatively low  $\pi_0/\pi_4$  ratio (~0.111 in Kenya and Gabon, and ~0.113 in Senegal). Consistent with  
486 a reduction in the efficacy of selection, the introduced accessions have higher ratios, near 0.116  
487 in Brazil and Cali, and 0.118 in Río Claro, while the grouped USA accessions were near 0.115  
488 (here artificially deflated due to population structure). These values are consistent with those  
489 reported in other insect species with no known history of introductions or eradication efforts  
490 (Chen et al. 2017). This decrease in the efficacy of selection is subtle, though perhaps  
491 unexpected given that neutral diversity in the introduced range remains high in an absolute  
492 sense, and in a relative sense was not as reduced as much as may have been expected.  
493



494  
495 **Figure 6:** Distribution of fitness effects (DFE) of new mutations for 0-fold degenerate sites relative  
496 to 4-fold degenerate sites for all regions. Error bars represent 95% bootstrapped confidence  
497 intervals for each bin.

498  
499 Inferring changes to the efficacy of selection via  $\pi_0/\pi_4$  in nonequilibrium populations can  
500 be complicated by differences in the recovery time to equilibrium between nonsynonymous and  
501 synonymous alleles (Brandvain and Wright 2016). As such, an increase in  $\pi_0/\pi_4$  of effectively  
502 neutral mutations as we find here may not solely reflect a reduced efficacy of selection. To  
503 address this, we can also use explicit modeling of the distribution of fitness effects (DFE) of new  
504 mutations. When we infer the DFE in each accession, we find clear differences between the  
505 African and American accessions—American accessions, and to a lesser extent Senegal,  
506 exhibit both an increase in strongly selected mutations and nearly neutral mutations (Figure 6).  
507 The simultaneous shift to low and high  $N_e s$  categories is evident in all introduced accessions,  
508 including the grouped USA accessions where the effect is attenuated by population structure  
509 (Andersson et al. 2023). The shift is primarily driven by a reduction in the low and intermediate  
510 effect classes, with minor increases in the nearly neutral class and large increases in the  
511 strongly selected class. In the nearly neutral category, we see increases in the American  
512 accessions of 1.7-3% over Senegal, and 3.5-4.5% over Kenya, representing a small but  
513 significant decrease in the efficacy of natural selection in introduced populations. Consequently,  
514 we do observe a small decrease in the efficacy of selection consistent with the increase in the  
515 ratio, though the biggest difference in the introduced accessions is in the unexpected shift  
516 towards more strongly deleterious mutations. However, it is possible that we have limited power  
517 to differentiate between moderately (10-100  $N_e s$ ) and strongly deleterious mutations (100+  $N_e s$ )



518 without knowledge of the true underlying DFE (Kousathanas and Keightley 2013); grouping  
519 these categories yields increases in the strongly selected categories for the American  
520 accessions of only 1-2% over the African accessions (Supplementary Figure 7).

521 The efficacy of selection is additionally affected by selective sweeps (Smith and Haigh  
522 1974), and in *Ae. aegypti*, particularly at sites under strong selection such as those underlying  
523 insecticide resistance (Love et al. 2023). While demography and selection both contribute to  
524 reductions in  $N_e$ , selective sweeps only affect diversity in the regions linked to selected sites  
525 (Smith and Haigh 1974), meaning that the impact on both diversity and the efficacy of selection  
526 are heterogeneous along the genome. To elucidate the contribution of selective sweeps to the  
527 reduction in the efficacy of selection genome-wide, we estimated  $\pi_0/\pi_4$  in the 10kb surrounding  
528 the top 1% of previously identified sweep signals (Love et al. 2023). We find the  $\pi_0/\pi_4$  ratio is  
529 substantially increased in regions surrounding sweeps in all accessions (Figure 5C). The  
530 Colombian accessions do not show as much of an increase as in other accessions, though this  
531 is likely due to the Colombian accessions having been grouped together in the previous work,  
532 dampening the effects of independent sweeps. This result is consistent with a substantial draft  
533 effect in the species, though here we are unable to quantify the amount of the genome affected.  
534 Future work is needed to determine the relative contribution of demographic and selective  
535 processes in shaping genome-wide diversity and the efficacy of selection, with important  
536 implications for *Ae. aegypti*'s evolved response to control efforts.

537

## 538 Conclusions

539 Here we investigated the demographic history of *Aedes aegypti* with respect to its introduction  
540 from Africa to the Americas. Our results suggest extensive population structure at fine and  
541 coarse scales, with some limited evidence for admixture in distant accessions in South America  
542 and between South America and some accessions in the US, while many accessions in close  
543 proximity exhibit deep divergence with no evidence of recent admixture. In particular,  
544 accessions from Colombia and Brazil appear to be largely genetically distinct, with the  
545 accession in Cali, Colombia more closely related to the accession from Santarem, Brazil than to  
546 the accession in Río Claro, Colombia, despite the relatively short geographic distance between  
547 the Colombian sampling locations. Similarly, accessions in California overlap each other in  
548 geographic space in the Central Valley, despite having estimated divergence times hundreds of  
549 years ago. The geographic overlap near Fresno, California separates the broad Northern  
550 California genetic cluster from the Southern California cluster, the latter of which includes

551 accessions from Florida, further highlighting the complexities of scale in population structure of  
552 the species. Together, our observed patterns of population structure and admixture imply that  
553 gene flow between regions may be limited, but occasional new (or re-) colonizations likely occur  
554 through long distance dispersal, as is likely within California. Our admixture results are more  
555 broadly consistent with the sorting of diverse ancestries during the introduction bottlenecks,  
556 supported by deep estimated divergence times. In the context of the spread of beneficial  
557 mutations, such as those conferring insecticide resistance, our results suggest that adaptation  
558 may occur through parallel mutation instead of through the spread of single-origin mutations  
559 across large geographic ranges, though future work is needed to more directly test this  
560 hypothesis at loci of interest.

561         Regarding the number and timing of introductions in the Americas, we find through  
562 inferences of historical effective population sizes and population split times that populations in  
563 both South and North America likely originated through multiple introductions. We note that  
564 inferring the precise timing of introduction events with genomic data is challenging, and our  
565 estimates should not be treated as exact, however, our divergence estimates suggest American  
566 populations split thousands of years ago, long before the proposed introduction of *aegypti* to the  
567 Americas via the slave trade (Rose et al. 2023). Similarly, while the differences in bottleneck  
568 strength and timing, the general lack of admixture in the Americas, and estimated split times  
569 predating possible introduction times by hundreds of years strongly suggest there have been  
570 multiple origins in South America, we cannot directly ascertain whether populations that were  
571 declared eradicated were in fact eliminated and recolonized from elsewhere, or if instead  
572 recovery occurred from the expansion of a smaller number of individuals surviving in refugia.  
573 However, the fact that we do not observe especially strong bottlenecks that would be expected  
574 for a population recovering from near eradication (discussed below), perhaps makes the  
575 recolonization hypothesis more likely. Similarly, for accessions in the US, estimated split times  
576 and bottleneck severities and timings point to independent origins relative to the sampled South  
577 American accessions, although we cannot identify the introduction route. Notably, some pairs of  
578 North and South American clusters exhibit markedly more recent estimated split times than  
579 other such pairs (e.g. Southern California and Clovis relative to Río Claro, and Florida relative to  
580 Brazil). These more recent split time estimates, which we obtained under a model that does not  
581 allow for migration, could potentially be underestimates caused by post-split gene flow between  
582 some North and South American clusters; this notion is consistent with our evidence of relatively  
583 low but nonzero admixture proportions shown in Figure 2B. Likewise relative to African  
584 accessions, the markedly recent split time estimates between both Florida and Brazil from

585 Gabon could reflect recent gene flow to the Americas, back migration to Africa, or shared  
586 admixture in Florida and Brazil from Gabon that predates their introduction to the Americas.  
587 Future work is needed to parse out the unique specifics of gene flow and American  
588 introductions of individual groups of samples.

589 We find all introduced accessions have similar levels of diversity, reduced by only ~33-  
590 40% from the ancestral levels. Diversity in these accessions therefore remains high ( $\pi \approx 0.02$  at  
591 synonymous sites), reflecting a model where  $N_e$  only dropped to the tens of thousands in South  
592 American populations before quickly recovering. We note that population structure in the US  
593 means the diversity estimates and measures of the efficacy of selection we report here are  
594 inflated, with underlying populations possibly exhibiting lower diversity than in South America  
595 given the recent low  $N_e$  in each US accession of about 5,000 (Figure 4A). Nevertheless,  
596 diversity levels in the US when grouped into a single population are still estimated to be reduced  
597 relative to ancestral levels. We additionally find evidence of only subtle reductions in the efficacy  
598 of selection relative to the ancestral range, as reflected in an increased  $\pi_0/\pi_4$  and an increase in  
599 the proportion of weakly deleterious mutations in the DFE. We do additionally find evidence that  
600 the efficacy of selection around selective sweeps is similar between African and American  
601 accessions, suggesting that the genome-wide signal of a weakly increased  $\pi_0/\pi_4$  in the  
602 Americas may be due primarily to the demographic history, though future work is required to  
603 tease apart the relative contributions of demography and selection to the genomic landscape of  
604 diversity in the species, and how deleterious variation may be distributed. This genomic  
605 resilience given the introduction bottlenecks, rapid expansions, and history of eradication  
606 attempts underscores the challenge faced by ongoing and future control efforts. Despite their  
607 demography, American populations of *Ae. aegypti aegypti* all exhibit high diversity and a strong  
608 efficacy of selection, making the path of rapid adaptation to ongoing and future anthropogenic  
609 control efforts likely even in the absence of strong gene flow.

## 610 **Methods**

### 611 **Data filtering**

612 Here we make use of a previously curated and filtered whole genome sequencing dataset of  
613 131 samples from (Love et al. 2023), which includes 27 samples from California and Florida  
614 from (Lee et al. 2019), 18 samples from Santarém, Brazil, 13 samples from Franceville, Gabon,  
615 19 samples from Kaya Bomu, Kenya, and 20 samples from Ngoye, Senegal taken from (Rose

616 et al. 2020), as well as 10 samples from Cali, Colombia and 24 samples from Río Claro,  
617 Colombia. Here we use the VCF from (Love et al. 2023), with detailed methods for alignment,  
618 variant calling and filtering described therein. Briefly, Colombian samples were sequenced using  
619 paired-end 150bp Illumina reads, quality checked with FASTQC 0.11.9 (Andrews 2010) and  
620 trimmed with Trimmomatic 0.39 (Bolger et al. 2014). Reads from all specimens were aligned to  
621 the AaegL5 reference genome using bwa-mem2 version 2.1 (Vasimuddin et al. 2019), then  
622 genotyped with GATK HaplotypeCaller version 4.1.9.0 (DePristo et al. 2011; Poplin et al. 2018),  
623 using the EMIT\_ALL\_CONFIDENT\_SITES flag. SNPs were filtered using scikit-allel version  
624 1.3.2 (DOI:10.5281/zenodo.4759368). The dataset from Love et al. (2023) that we used here  
625 imposed the following filters:

- 626 1) For the Colombian accession, specimens with mean coverage below 15X or fewer than  
627 70% of reads mapping were removed.
- 628 2) For the full sample, indels were removed, and SNPs were filtered following GATK's best  
629 practices (<https://gatk.broadinstitute.org/hc/en-us/articles/360035890471-Hard-filtering-germline-short-variants>). Specifically, sites were retained if they passed 5 of the 6  
630 following filters: i) variant quality by depth (QD) greater than or equal to 2, ii) variant  
631 strand bias via Fisher's exact test (FS) less than 40, iii) variant strand bias via symmetric  
632 odds ratio test (SQR) less than 4, iv) mapping quality (MQ) greater than or equal to 40,  
633 v) mapping quality rank sum test (MQRS) between -5 and 5, inclusive, and vi) site  
634 position within reads rank sum test (RPRS) between -3 and 3, inclusive.
- 635 3) All SNPs that were not called in at least 75% of specimens in at least 5 of the 6 countries  
636 were removed.
- 637 4) All non-diallelic SNPs were removed.
- 638 5) We removed regions previously identified as repetitive (Matthews et al. 2018) and  
639 regions that were not uniquely mappable.

640  
641 In addition to these filters used by Love et al., for analyses that required regions of the  
642 genome not likely to have experienced strong linked or direct selection (See Population  
643 Structure and Historical effective population size inference and split times), we further filtered  
644 the genome to remove all sites within 100 kilobases (kb) of annotated genes and for locations of  
645 centromeres as reported in (Matthews et al. 2018). This mask is referred to as the intergenic  
646 sites mask in later analyses.

647 For diversity statistics that require an accurate count of quality invariant sites, we further  
648 removed invariant sites with more than 25% of samples missing base calls using bcftools  
649 version 1.16 (Danecek et al. 2021), as well as sites with less than a depth of 4 or greater than a

650 depth of 30. All subsetting or filtering of VCFs for specific masks or regions was performed  
651 using bedtools v2.30.0 (Quinlan and Hall 2010) unless otherwise noted. Some analyses  
652 required sites of 0-fold or 4-fold degeneracy, which were called using a python script  
653 (<https://github.com/tvkenet/Degeneracy>), the L5.0 version of the *Ae. aegypti* genome and  
654 annotation, NCBI accession GCF\_002204515.2 (Matthews et al. 2018).  
655

## 656 Population structure and admixture

657 To infer the population structure of *Ae. aegypti*, we used two different approaches. First, we  
658 used PCA clustering as a high-level description of the underlying structure. We used the  
659 bedtools tool *intersect* to remove sites in the aforementioned intergenic sites mask from the full  
660 VCF file and converted to plink format using the `--make-bed` function in plink v1.90b3.45 (Purcell  
661 et al. 2007). Sites were thinned for linkage disequilibrium (LD) by removing sites in 100kb  
662 windows exceeding an  $r^2$  of 0.5, with a step size of 10kb. Principal components were then  
663 calculated for the full dataset and separately for samples from the Americas. Second, we  
664 estimated population structure more explicitly with ADMIXTURE (Alexander et al. 2009), using  
665 the same LD-pruned SNP set for all samples and for the Americas with  $k=2-10$ . The K number  
666 of populations with the best predictive accuracy was determined using ADMIXTURE's cross-  
667 validation with the `--cv` flag, and we present results for several K around this value.

## 668 Historical effective population size inference and split times

669 In order to obtain a view of the history of bottlenecks and expansions in *Ae. aegypti*, we inferred  
670 the historical effective population size of each accession. Inference of historical  $N_e$  makes use of  
671 information in the site frequency spectrum (SFS) (Liu and Fu 2020a) and/or uses local  
672 sequence information to infer a distribution of coalescent times with the sequentially Markovian  
673 coalescent (SMC) (Mather et al. 2020). SMC-based methods have been shown to have the  
674 highest accuracy over times from several hundred to tens of thousands of generations ago,  
675 while SFS-based methods are most accurate in the recent past (Patton et al. 2019), assuming a  
676 large enough sample size (Terhorst and Song 2015). Because of our sample size constraints,  
677 unphased data, and the use of folded SFS's, we use SMC++ (Terhorst et al. 2017), which has  
678 been shown to have high accuracy for a broad time frame (Terhorst et al. 2017; Patton et al.  
679 2019), and we provide inferences from the SFS-based Stairway Plot 2 (Liu and Fu 2020a) in the  
680 supplement for qualitative assurance of size histories. For each method, we scaled generations

681 to years and set the per generation mutation rate using estimates previously reported in (Rose  
682 et al. 2023): 0.067 years per generation and a mutation rate of  $4.85 \times 10^{-9}$ .

683 We used SMC++ v1.15.4 (Terhorst et al. 2017), and masked the genome using our  
684 intergenic sites mask. We ran inference on all accession as described in the previous section,  
685 as well as dividing the United States into 5 putative populations determined by their clustering in  
686 PCA space and corroboration with (Lee et al. 2019): Southern California (Commerce, Mission  
687 Viejo, Garden Grove, Brawley, San Diego), Northern California (Madera, Menlo Park, Fresno),  
688 Clovis and Sanger, Exeter, and Florida (see Supplement for plots of PC's 1-5). We converted a  
689 VCF for each chromosome to SMC++ format with the `vcf2smc` function in SMC++ using our  
690 mask and choosing random individuals as “distinguished” lineages. SMC++ makes use of the  
691 information captured in multiple samples by combining the coalescent times and allelic states  
692 between haplotypes of a set of one or more “distinguished” individuals, as in other SMC  
693 methods, with the allele frequencies of a set of “unlabeled samples” conditioned on the  
694 information from the distinguished individuals (Terhorst et al. 2017). For accessions with more  
695 than 5 samples, we used 5 distinguished individuals, for accessions with 3-5 samples we used 2  
696 distinguished individuals, and for accessions with only 2 samples we chose a single  
697 distinguished individual, with the remaining individual specimens from an accession considered  
698 as unlabeled samples. We ran the `estimate` function in SMC++ on these data, using 10 knots, a  
699 window size of 10, and regularization penalties of 4, 5, and 6 for inference over the past  
700 800,000 years. Because the choice of regularization penalty resulted in qualitatively different  $N_e$   
701 trajectories for some accessions (see Results), we ran neutral simulations in *msprime* version  
702 1.2.0 (Baumdicker et al. 2022) to estimate the site frequency spectra under the inferred  
703 demographic histories to independently compare the fit of SMC++ runs under the three different  
704 regularization parameters. Simulated SFSs were extracted from the simulated tree sequence  
705 using `tskit` version 0.5.6 (Kelleher et al. 2019) and compared to the observed SFS for each  
706 accession using a multinomial likelihood calculation following (Beichman et al. 2017); we  
707 present results for all accessions with a regularization penalty of 5, which was a better fit for our  
708 South American accessions and Gabon, and did not qualitatively alter the results for Kenya and  
709 Senegal (See Results and Supplementary Figures 8-14). We used the same parameters for the  
710 USA accessions, though we did not assess the fit of the SFS for these accessions as the  
711 sample sizes are too low to reliably use SFS-based methods and following the other  
712 accessions, the qualitative results are unlikely to be altered.

713 In addition to inference of  $N_e$  over time, SMC++ allows for estimation of population split  
714 times via a clean split model using inferred within- and cross-coalescent rates from SMC++



715 inference within and between populations. To estimate split times between accessions, we used  
716 the split function in SMC++ for every accession pair. We followed the same procedure  
717 described above to determine distinguished individuals for each accession. We then converted  
718 a VCF for each chromosome to SMC++ format using the `vcf2smc` function, and separately for  
719 each accession using the distinguished individuals in the focal accession. We treated all  
720 remaining individuals from both accessions as undistinguished lineages. We then ran the split  
721 function in SMC++ using these joint SMC++ files and the model fits for each accession from the  
722 previous estimate step.

723 In addition to SMC++  $N_e$  history inference, we ran Stairway Plot 2 (Liu and Fu 2020b) on  
724 the country-level groups of accessions (splitting Colombia into Río Claro and Cali) using an  
725 intergenic sites SNP set filtered as described in Data Filtering, but masking 10kb around genes.  
726 For each accession, we calculated the folded SFS using `scikit-allele`  
727 (DOI:10.5281/zenodo.4759368), then ran Stairway Plot 2 using the mutation rate and  
728 generation time estimates from (Rose et al. 2023) as with SMC++, and a sequence length,  $L$ , of  
729 66,563,773, representing the number of filtered variant and invariant sites used to generate the  
730 SFS.

731

## 732 Diversity statistics and the distribution of fitness effects

733 We calculated pairwise genetic diversity as  $\pi$  for each accession using `pixy` (Korunes and  
734 Samuk 2021) in windows of 500kb, following (Love et al. 2023), separately for all sites, 0-fold  
735 sites, and 4-fold sites. Genome-wide averages were recalculated using the raw `pixy` output as  
736 the sum of differences (`count_diffs`) divided by the sum of comparisons (`count_comparisons`).  
737 We also report genome wide average  $F_{ST}$  and  $D_{xy}$  for all sites (Supplementary figures 4 and 5).  
738 Error bars for all genome-wide averages were calculated as the 95% bootstrapped confidence  
739 interval with 100 bootstraps over windows in R. *Ae. aegypti* has faced substantial eradication  
740 efforts, and as such has undergone strong, recent positive selection (Love et al. 2023), which  
741 should lead to genomic heterogeneity in the efficacy of selection due to hitchhiking (Smith and  
742 Haigh 1974). As such, we also estimated the ratio  $\pi_0/\pi_4$  in regions immediately surrounding  
743 selective sweeps. For diversity statistics around selective sweeps, we filtered the Sweepfinder  
744 results from Love et al. 2023 for the top 1% of CLR windows, and calculated the statistic of  
745 interest in windows surrounding these 10kb outlier windows.

746 Because the behaviour of statistics like  $\pi_0/\pi_4$  is unclear in nonequilibrium populations,  
747 we additionally estimated the distribution of fitness effects of new mutations (DFE) in each



748 accession, which is thought to be more robust to nonequilibrium dynamics (Brandvain and  
749 Wright 2016). To estimate the DFE, we first calculated the folded site frequency spectrum using  
750 the `sfs_folded` function in `scikit-allel` for 0-fold, 4-fold, and intergenic sites defined as sites more  
751 than 100kb up- and downstream of annotated genes. We estimated the DFE for 0-fold  
752 degenerate sites relative to both 4-fold degenerate and intergenic sites using DFE-alpha version  
753 2.15 (Keightley and Eyre-Walker 2007) using default parameters, and obtaining  $N_e$ s bins using  
754 the `prop_muts_in_s_ranges` function. 95% confidence intervals were calculated by  
755 bootstrapping the site frequency spectra by site using the `random.choice` function in `numpy`  
756 (Harris et al. 2020) to randomly sample the genotype arrays and re-estimating the DFE with  
757 1,000 replicates.

758  
759

## 760 **Acknowledgements**

761 We would like to thank members of the Matute and Schrider labs, Tom Booker, and Julia  
762 Kreiner for useful feedback and discussion. D.R.S. was supported by NIH award  
763 R35GM138286. D.R.M. was supported by NIH award R35GM148244. The funders had no role  
764 in study design, data collection and analysis, decision to publish, or preparation of the  
765 manuscript.

766

## 767 **Data and Code Availability**

768 All sequencing data used in this paper are publicly available (see Love et al. 2023). All code  
769 used for the analyses presented in this paper can be found at  
770 [https://github.com/tvKent/aedes\\_demography/](https://github.com/tvKent/aedes_demography/).

## 771 **References**

772 Alexander DH, Novembre J, Lange K. 2009. Fast model-based estimation of ancestry in  
773 unrelated individuals. *Genome Res.* 19:1655–1664.  
774 Al Nazawi AM, Aqili J, Alzahrani M, McCall PJ, Weetman D. 2017. Combined target site (kdr)

- 775 mutations play a primary role in highly pyrethroid resistant phenotypes of *Aedes aegypti*  
776 from Saudi Arabia. *Parasit. Vectors* 10:161.
- 777 American Health Organization P. 1997. The feasibility of eradicating *Aedes aegypti* in the  
778 Americas. *Rev. Panam. Salud Publica* 1:68–72.
- 779 Andersson BA, Zhao W, Haller BC, Brännström Å, Wang X-R. 2023. Inference of the distribution  
780 of fitness effects of mutations is affected by single nucleotide polymorphism filtering  
781 methods, sample size and population structure. *Mol. Ecol. Resour.* 23:1589–1603.
- 782 Andrews S. 2010. FastQC: A Quality Control Tool for High Throughput Sequence Data.  
783 Available from: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- 784 Baker HG. 1974. The evolution of weeds. *Annu. Rev. Ecol. Syst.* 5:1–24.
- 785 Baker HG, Stebbins GL. Genetics of colonizing species, proceedings. *International Union of*  
786 *Biological.*
- 787 Barrett SCH. 2015. Foundations of invasion genetics: the Baker and Stebbins legacy. *Mol. Ecol.*  
788 24:1927–1941.
- 789 Barton N, Partridge L. 2000. Limits to natural selection. *Bioessays* 22:1075–1084.
- 790 Baumdicker F, Bisschop G, Goldstein D, Gower G, Ragsdale AP, Tsambos G, Zhu S, Eldon B,  
791 Ellerman EC, Galloway JG, et al. 2022. Efficient ancestry and mutation simulation with  
792 msprime 1.0. *Genetics* 220: iyab229.
- 793 Beichman AC, Phung TN, Lohmueller KE. 2017. Comparison of Single Genome and Allele  
794 Frequency Data Reveals Discordant Demographic Histories. *G3* 7:3605–3620.
- 795 Blackburn TM, Lockwood JL, Cassey P. 2016. The influence of numbers on invasion success.  
796 In: *Invasion Genetics*. Chichester, UK: John Wiley & Sons, Ltd. p. 25–39.
- 797 Blake JB. 1968. Yellow fever in eighteenth century America. *Bull. N. Y. Acad. Med.* 44:673–686.
- 798 Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence  
799 data. *Bioinformatics* 30:2114–2120.
- 800 Brandvain Y, Wright SI. 2016. The Limits of Natural Selection in a Nonequilibrium World. *Trends*  
801 *Genet.* 32:201–210.
- 802 Brathwaite Dick O, San Martín JL, Montoya RH, del Diego J, Zambrano B, Dayan GH. 2012.  
803 The history of dengue outbreaks in the Americas. *Am. J. Trop. Med. Hyg.* 87:584–593.
- 804 Brown JE, Evans BR, Zheng W, Obas V, Barrera-Martinez L, Egizi A, Zhao H, Caccone A,  
805 Powell JR. 2014. Human impacts have shaped historical and recent evolution in *Aedes*  
806 *aegypti*, the dengue and yellow fever mosquito. *Evolution* 68:514–525.
- 807 Brown JE, McBride CS, Johnson P, Ritchie S, Paupy C, Bossin H, Lutomiah J, Fernandez-  
808 Salas I, Ponlawat A, Cornel AJ, et al. 2011. Worldwide patterns of genetic differentiation  
809 imply multiple “domestications” of *Aedes aegypti*, a major vector of human diseases. *Proc.*  
810 *Biol. Sci.* 278:2446–2454.

- 811 Bryan CS, Moss SW, Kahn RJ. 2004. Yellow fever in the Americas. *Infect. Dis. Clin. North Am.*  
812 18:275–292.
- 813 Buffalo V. 2021. Quantifying the relationship between genetic diversity and population size  
814 suggests natural selection cannot explain Lewontin’s Paradox. *Elife* 10: e67509.
- 815 Cafferata ML, Bardach A, Rey-Ares L, Alcaraz A, Cormick G, Gibbons L, Romano M, Cesaroni  
816 S, Ruvinsky S. 2013. Dengue Epidemiology and Burden of Disease in Latin America and  
817 the Caribbean: A Systematic Review of the Literature and Meta-Analysis. *Value Health Reg*  
818 *Issues* 2:347–356.
- 819 Carter HR. 1931. Yellow Fever: An Epidemiological and Historical Study of Its Place of Origin.  
820 William & Wilkins Company
- 821 Chang C, Ortiz K, Ansari A, Gershwin ME. 2016. The Zika outbreak of the 21st century. *J.*  
822 *Autoimmun.* 68:1–13.
- 823 Chen J, Glémin S, Lascoux M. 2017. Genetic Diversity and the Efficacy of Purifying Selection  
824 across Plant and Animal Species. *Mol. Biol. Evol.* 34:1417–1428.
- 825 Chornesky EA, Randall JM. 2003. The Threat of Invasive Alien Species to Biological Diversity:  
826 Setting a Future Course. *Ann. Mo. Bot. Gard.* 90:67–76.
- 827 Christophers SR. 1960. *Aedes aegypti* (L.) the yellow fever mosquito: its life history, bionomics  
828 and structure. Cambridge University Press
- 829 Comeault AA, Wang J, Tittes S, Isbell K, Ingley S, Hurlbert AH, Matute DR. 2020. Genetic  
830 Diversity and Thermal Performance in Invasive and Native Populations of African Fig Flies.  
831 *Mol. Biol. Evol.* 37:1893–1906.
- 832 Corbett-Detig RB, Hartl DL, Sackton TB. 2015. Natural selection constrains neutral diversity  
833 across a wide range of species. *PLoS Biol.* 13:e1002112.
- 834 Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T,  
835 McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools.  
836 *Gigascience* 10: giab008.
- 837 DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel  
838 G, Rivas MA, Hanna M, et al. 2011. A framework for variation discovery and genotyping  
839 using next-generation DNA sequencing data. *Nat. Genet.* 43:491–498.
- 840 Dlugosch KM, Parker IM. 2008. Founding events in species invasions: genetic variation,  
841 adaptive evolution, and the role of multiple introductions. *Mol. Ecol.* 17:431–449.
- 842 Dueñas M-A, Hemming DJ, Roberts A, Diaz-Soltero H. 2021. The threat of invasive species to  
843 IUCN-listed critically endangered species: A systematic review. *Global Ecology and*  
844 *Conservation* 26:e01476.
- 845 Edelman NB, Mallet J. 2021. Prevalence and Adaptive Impact of Introgression. *Annu. Rev.*  
846 *Genet.* 55:265–283.
- 847 Eisen L, Moore CG. 2013. *Aedes* (*Stegomyia*) *aegypti* in the continental United States: a vector  
848 at the cool margin of its geographic range. *J. Med. Entomol.* 50:467–478.

- 849 Elton CS. 1958. *The Ecology of Invasions by Animals and Plants*. Springer, US.
- 850 Evans MEK, Dennehy JJ. 2005. Germ banking: bet-hedging and variable release from egg and  
851 seed dormancy. *Q. Rev. Biol.* 80:431–451.
- 852 Fan Y, O’Grady P, Yoshimizu M, Ponlawat A, Kaufman PE, Scott JG. 2020. Evidence for both  
853 sequential mutations and recombination in the evolution of kdr alleles in *Aedes aegypti*.  
854 *PLoS Negl. Trop. Dis.* 14:e0008154.
- 855 Fischer S, De Majo MS, Di Battista CM, Montini P, Loetti V, Campos RE. 2019. Adaptation to  
856 temperate climates: Evidence of photoperiod-induced embryonic dormancy in *Aedes*  
857 *aegypti* in South America. *J. Insect Physiol.* 117:103887.
- 858 Frankham R. 2005. Resolving the genetic paradox in invasive species. *Heredity* 94:385.
- 859 Gloria-Soria A, Ayala D, Bheecarry A, Calderon-Arguedas O, Chadee DD, Chiappero M,  
860 Coetzee M, Elahee KB, Fernandez-Salas I, Kamal HA, et al. 2016. Global genetic diversity  
861 of *Aedes aegypti*. *Mol. Ecol.* 25:5377–5395.
- 862 Haddi K, Tomé HVV, Du Y, Valbon WR, Nomura Y, Martins GF, Dong K, Oliveira EE. 2017.  
863 Detection of a new pyrethroid resistance mutation (V410L) in the sodium channel of *Aedes*  
864 *aegypti*: a potential challenge for mosquito control. *Sci. Rep.* 7:46549.
- 865 Haldane JB. 1956. The relation between density regulation and natural selection. *Proc. R. Soc.*  
866 *Lond. B Biol. Sci.* 145:306–308.
- 867 Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E,  
868 Taylor J, Berg S, Smith NJ, et al. 2020. Array programming with NumPy. *Nature* 585:357–  
869 362.
- 870 Hufbauer RA, Facon B, Ravigné V, Turgeon J, Foucaud J, Lee CE, Rey O, Estoup A. 2012.  
871 Anthropogenically induced adaptation to invade (AIAI): contemporary adaptation to human-  
872 altered habitats within the native range can promote invasions. *Evol. Appl.* 5:89–101.
- 873 Hulme-Beaman A, Dobney K, Cucchi T, Searle JB. 2016. An Ecological and Evolutionary  
874 Framework for Commensalism in Anthropogenic Environments. *Trends Ecol. Evol.* 31:633–  
875 645.
- 876 Kaj I, Krone SM, Lascoux M. 2001. Coalescent theory for seed bank models. *J. Appl. Probab.*  
877 38:285–300.
- 878 Kawada H, Oo SZM, Thaug S, Kawashima E, Maung YNM, Thu HM, Thant KZ, Minakawa N.  
879 2014. Co-occurrence of point mutations in the voltage-gated sodium channel of pyrethroid-  
880 resistant *Aedes aegypti* populations in Myanmar. *PLoS Negl. Trop. Dis.* 8:e3032.
- 881 Keightley PD, Eyre-Walker A. 2007. Joint inference of the distribution of fitness effects of  
882 deleterious mutations and population demography based on nucleotide polymorphism  
883 frequencies. *Genetics* 177:2251–2261.
- 884 Kelleher J, Wong Y, Wohns AW, Fadil C, Albers PK, McVean G. 2019. Inferring whole-genome  
885 histories in large population datasets. *Nat. Genet.* 51:1330–1338.
- 886 Korunes KL, Samuk K. 2021. pixy: Unbiased estimation of nucleotide diversity and divergence

- 887 in the presence of missing data. *Mol. Ecol. Resour.* 21:1359–1368.
- 888 Kotsakiozi P, Gloria-Soria A, Caccone A, Evans B, Schama R, Martins AJ, Powell JR. 2017.  
889 Tracking the return of *Aedes aegypti* to Brazil, the major vector of the dengue, chikungunya  
890 and Zika viruses. *PLoS Negl. Trop. Dis.* 11:e0005653.
- 891 Kousathanas A, Keightley PD. 2013. A comparison of models to infer the distribution of fitness  
892 effects of new mutations. *Genetics* 193:1197–1208.
- 893 Lee CE, Gelembiuk GW. 2008. Evolutionary origins of invasive populations. *Evol. Appl.* 1:427–  
894 448.
- 895 Lee Y, Schmidt H, Collier TC, Conner WR, Hanemaaijer MJ, Slatkin M, Marshall JM, Chiu JC,  
896 Smartt CT, Lanzaro GC, et al. 2019. Genome-wide divergence among invasive populations  
897 of *Aedes aegypti* in California. *BMC Genomics* 20:204.
- 898 Leffler EM, Bullaughey K, Matute DR, Meyer WK, Ségurel L, Venkat A, Andolfatto P, Przeworski  
899 M. 2012. Revisiting an old riddle: what determines genetic diversity levels within species?  
900 *PLoS Biol.* 10:e1001388.
- 901 Lenormand T. 2002. Gene flow and the limits to natural selection. *Trends Ecol. Evol.* 17:183–  
902 189.
- 903 Liu C, Wolter C, Xian W, Jeschke JM. 2020. Most invasive species largely conserve their  
904 climatic niche. *Proc. Natl. Acad. Sci. U. S. A.* 117:23643–23651.
- 905 Liu X, Fu Y-X. 2020a. Author Correction: Stairway Plot 2: demographic history inference with  
906 folded SNP frequency spectra. *Genome Biol.* 21:305.
- 907 Liu X, Fu Y-X. 2020b. Stairway Plot 2: demographic history inference with folded SNP frequency  
908 spectra. *Genome Biol.* 21:280.
- 909 Love RR, Sikder JR, Vivero RJ, Matute DR, Schrider DR. 2023. Strong Positive Selection in  
910 *Aedes aegypti* and the Rapid Evolution of Insecticide Resistance. *Mol. Biol. Evol.* 40:  
911 msad072.
- 912 Lozada-Chávez AN, Lozada-Chávez I, Alfano N, Palatini U, Sogliani D, Elfekih S, Degefa T,  
913 Sharakhova MV, Badolo A, Patchara S, et al. 2023. Molecular signature of domestication in  
914 the arboviral vector *Aedes aegypti*. *bioRxiv* 2023.03.13.532092.
- 915 Machado-Allison CE, Craig GB. 1972. Geographic Variation in Resistance to Desiccation in  
916 *Aedes aegypti* and *A. atropalpus* (Diptera: Culicidae). *Ann. Entomol. Soc. Am.* 65:542–547.
- 917 Mather N, Traves SM, Ho SYW. 2020. A practical introduction to sequentially Markovian  
918 coalescent methods for estimating demographic history from genomic data. *Ecol. Evol.*  
919 10:579–589.
- 920 Matthews BJ, Dudchenko O, Kingan SB, Koren S, Antoshechkin I, Crawford JE, Glassford WJ,  
921 Herre M, Redmond SN, Rose NH, et al. 2018. Improved reference genome of *Aedes*  
922 *aegypti* informs arbovirus vector control. *Nature* 563:501–507.
- 923 Mayilsamy M. 2019. Extremely Long Viability of *Aedes aegypti* (Diptera: Culicidae) Eggs Stored  
924 Under Normal Room Condition. *J. Med. Entomol.* 56:878–880.



- 925 Monteiro FA, Schama R, Martins AJ, Gloria-Soria A, Brown JE, Powell JR. 2014. Genetic  
926 diversity of Brazilian *Aedes aegypti*: patterns following an eradication program. *PLoS Negl.*  
927 *Trop. Dis.* 8:e3167.
- 928 Nei M, Maruyama T, Chakraborty R. 1975. The Bottleneck Effect and Genetic Variability in  
929 Populations. *Evolution* 29:1–10.
- 930 Orr HA, Unckless RL. 2014. The population genetics of evolutionary rescue. *PLoS Genet.*  
931 10:e1004551.
- 932 Osmond MM, Otto SP, Martin G. 2019. Genetic Paths to Evolutionary Rescue and the  
933 Distribution of Fitness Effects Along Them. *Genetics* 214: 493-510.
- 934 Paini DR, Sheppard AW, Cook DC, De Barro PJ, Worner SP, Thomas MB. 2016. Global threat  
935 to agriculture from invasive species. *Proc. Natl. Acad. Sci. U. S. A.* 113:7575–7579.
- 936 Patterson KD. 1992. Yellow fever epidemics and mortality in the United States, 1693-1905. *Soc.*  
937 *Sci. Med.* 34:855–865.
- 938 Patton AH, Margres MJ, Stahlke AR, Hendricks S, Lewallen K, Hamede RK, Ruiz-Aravena M,  
939 Ryder O, McCallum HI, Jones ME, et al. 2019. Contemporary Demographic Reconstruction  
940 Methods Are Robust to Genome Assembly Quality: A Case Study in Tasmanian Devils.  
941 *Mol. Biol. Evol.* 36:2906–2921.
- 942 Pless E, Gloria-Soria A, Evans BR, Kramer V, Bolling BG, Tabachnick WJ, Powell JR. 2017.  
943 Multiple introductions of the dengue vector, *Aedes aegypti*, into California. *PLoS Negl.*  
944 *Trop. Dis.* 11:e0005718.
- 945 Poplin R, Ruano-Rubio V, DePristo MA, Fennell TJ, Carneiro MO, Van der Auwera GA, Kling  
946 DE, Gauthier LD, Levy-Moonshine A, Roazen D, et al. 2018. Scaling accurate genetic  
947 variant discovery to tens of thousands of samples. *bioRxiv* 201178. doi:  
948 <https://doi.org/10.1101/201178>
- 949 Powell JR. 2019. An Evolutionary Perspective on Vector-Borne Diseases. *Front. Genet.*  
950 10:1266.
- 951 Powell JR, Gloria-Soria A, Kotsakiozi P. 2018. Recent History of *Aedes aegypti*: Vector  
952 Genomics and Epidemiology Records. *Bioscience* 68:854–860.
- 953 Powell JR, Tabachnick WJ. 2013. History of domestication and spread of *Aedes aegypti*—a  
954 review. *Mem. Inst. Oswaldo Cruz* 108 Suppl 1:11–17.
- 955 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de  
956 Bakker PIW, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and  
957 population-based linkage analyses. *Am. J. Hum. Genet.* 81:559–575.
- 958 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic  
959 features. *Bioinformatics* 26:841–842.
- 960 Rose NH, Badolo A, Sylla M, Akorli J, Otoo S, Gloria-Soria A, Powell JR, White BJ, Crawford  
961 JE, McBride CS. 2023. Dating the origin and spread of specialization on human hosts in  
962 *Aedes aegypti* mosquitoes. *Elife* e83524.

- 963 Rose NH, Sylla M, Badolo A, Lutomiah J, Ayala D, Aribodor OB, Ibe N, Akorli J, Otoo S, Mutebi  
964 J-P, et al. 2020. Climate and Urbanization Drive Mosquito Preference for Humans. *Curr.*  
965 *Biol.* 30:3570–3579.e6.
- 966 Saavedra-Rodriguez K, Maloof FV, Campbell CL, Garcia-Rejon J, Lenhart A, Penilla P,  
967 Rodriguez A, Sandoval AA, Flores AE, Ponce G, et al. 2018. Parallel evolution of vgsc  
968 mutations at domains IS6, IIS6 and IIIS6 in pyrethroid resistant *Aedes aegypti* from Mexico.  
969 *Sci. Rep.* 8:6747.
- 970 Sakai AK, Allendorf FW, Holt JS, Lodge DM, Molofsky J, With KA, Baughman S, Cabin RJ,  
971 Cohen JE, Ellstrand NC, et al. 2003. The Population Biology of Invasive Species. *Annual*  
972 *Review of Ecology and Systematics.* 32:1, 305-332
- 973 San Martín JL, Brathwaite O, Zambrano B, Solórzano JO, Bouckennooghe A, Dayan GH,  
974 Guzmán MG. 2010. The epidemiology of dengue in the Americas over the last three  
975 decades: a worrisome reality. *Am. J. Trop. Med. Hyg.* 82:128–135.
- 976 Sax DF, Brown JH. 2000. The paradox of invasion. *Glob. Ecol. Biogeogr.* 9:363–371.
- 977 Schrieber K, Lachmuth S. 2017. The Genetic Paradox of Invasions revisited: the potential role  
978 of inbreeding × environment interactions in invasion success. *Biol. Rev. Camb. Philos. Soc.*  
979 92:939–952.
- 980 Shepard DS, Coudeville L, Halasa YA, Zambrano B, Dayan GH. 2011. Economic impact of  
981 dengue illness in the Americas. *Am. J. Trop. Med. Hyg.* 84:200–207.
- 982 Slatkin M, Pollack JL. 2008. Subdivision in an ancestral species creates asymmetry in gene  
983 trees. *Mol. Biol. Evol.* 25:2241–2246.
- 984 Smith JM, Haigh J. 1974. The hitch-hiking effect of a favourable gene. *Genet. Res.* 23:23–35.
- 985 Tapia-Conyer R, Méndez-Galván JF, Gallardo-Rincón H. 2009. The growing burden of dengue  
986 in Latin America. *J. Clin. Virol.* 46 Suppl 2:S3–S6.
- 987 Terhorst J, Kamm JA, Song YS. 2017. Robust and scalable inference of population history from  
988 hundreds of unphased whole genomes. *Nat. Genet.* 49:303–309.
- 989 Terhorst J, Song YS. 2015. Fundamental limits on the accuracy of demographic inference  
990 based on the sample frequency spectrum. *Proc. Natl. Acad. Sci. U. S. A.* 112:7677–7682.
- 991 Vasimuddin M, Misra S, Li H, Aluru S. 2019. Efficient Architecture-Aware Acceleration of BWA-  
992 MEM for Multicore Systems. In: 2019 IEEE International Parallel and Distributed  
993 Processing Symposium (IPDPS). p. 314–324.
- 994 Webb JLA Jr. 2016. *Aedes aegypti* suppression in the Americas: historical perspectives. *Lancet*  
995 388:556–557.
- 996 Yeaman S. 2022. Evolution of polygenic traits under global vs local adaptation. *Genetics* 220:  
997 iyab134.