Adaptation to host cell environment during experimental evolution of Zika virus.

Vincent Grass[1], Emilie Hardy[1$], Kassian Kobert[2$], Soheil Rastgou Talemi[3$], Elodie Décembre[1], Coralie Guy[1], Peter V. Markov[2], Alain Kohl[4], Mathilde Paris[5], Anja Böckmann[6], Sara Muñoz-González[1], Lee Sherry[1], Thomas Höfer[3], Bastien Boussau[2]* and Marlène Dreux[1]*

[1] CIRI, Inserm, U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, École Normale Supérieure de Lyon, Univ Lyon, Lyon, 69007, France

[2] Laboratoire de Biométrie et Biologie Évolutive (LBBE), UMR CNRS 5558, Université Claude Bernard Lyon 1, Lyon, 69622, France

[3] Theoretical Systems Biology, German Cancer Research Center, Deutsches Krebsforschungszentrum (DKFZ) Heidelberg, 69120, Germany

[4] MRC-University of Glasgow Centre for Virus Research, Glasgow G61 1QH, UK

[5] Institut de Génomique Fonctionnelle de Lyon (IGFL), École Normale Supérieure de Lyon, Lyon, 69007, France

[6] Institut de Biologie et Chimie des Protéines, MMSB, Labex Ecofect, UMR 5086 CNRS, Université de Lyon, Lyon, 69007, France

[$] These authors contributed equally.

* corresponding authors: Marlène Dreux, CIRI, École Normale Supérieure de Lyon, 46 allée d'Italie, 69007, Lyon France, phone : +334 26 23 38 34 and Bastien Boussau, UMR CNRS 5558 - LBBE, UCB Lyon 1 - 43 bd du 11 novembre 1918, 69622 VILLEURBANNE cedex, France, phone : +336 27 64 29 11.

Email: marlene.dreux@ens-lyon.fr Bastien.Boussau@univ-lyon1.fr

Keywords: Zika virus; experimental evolution; viral adaptation; mathematical modelling; deep-sequencing analysis.

Abstract

Zika virus (ZIKV) infection can cause important developmental and neurological defects in Humans. Type I/III interferon responses control ZIKV infection and pathological processes, yet the virus has evolved various mechanisms to defeat these host responses. Here, we established a pipeline to delineate at high-resolution the genetic evolution of ZIKV in a controlled host cell environment. We uncovered that serially passaged ZIKV acquired increased infectivity and simultaneously developed a resistance to TLR3-induced restriction. We built a mathematical model that suggests that the increased infectivity is due to a reduced time-lag between infection and viral replication. We found that this adaptation is cell-type specific, suggesting that different cell environments may drive viral evolution along different routes. Deep-sequencing of ZIKV populations pinpointed mutations whose increased frequencies temporally coincide with the acquisition of the adapted phenotype. We functionally validated S455L, a substitution in ZIKV envelope (E) protein, recapitulating the adapted phenotype. Its positioning on the E structure suggests a putative function in protein refolding/stability. Taken together, our results uncovered ZIKV adaptations to the cellular environment leading to accelerated replication onset coupled with resistance to TLR3-induced antiviral response. Our work provides insights into Zika virus adaptation to host cells and immune escape mechanisms.

43 **Introduction**

44 Zika virus (ZIKV; *Flaviviridae*) is a mosquito-borne human pathogen related to other globally
45 relevant flaviviruses, including dengue, yellow fever, West Nile, Japanese encephalitis and
46 tick-borne encephalitis viruses. As is typical for flaviviruses, ZIKV has a 10.8 kb RNA genome
47 of positive polarity, encoding a polyprotein composed of 3 structural proteins (C, prM and E)
48 and 7 nonstructural (NS) proteins. The NS proteins are involved in the steps of RNA synthesis
49 and assembly of viral particles. Several NS proteins of flaviviruses interfere with host antiviral
50 responses, either by inhibition of the innate sensors or downstream signaling pathways[1-4].

51

52 For decades, ZIKV infections were either unrecognized or occurred only sporadically and were
53 associated with mild symptoms. However, ZIKV was detected in Brazil in 2015 and spread
54 rapidly, reaching infection rates exceeding 50%[5]. During the Brazilian ZIKV
55 outbreak, congenital infections led to fetal demise, microcephaly and other developmental
56 abnormalities (now grouped as Congenital Zika Syndrome), *e.g.,* visual and hearing
57 impairment, skeletal deformities, and possible Guillain–Barré syndrome in adults[5-9]. Severe
58 symptoms, including neural development defects and fetal demise, are linked to host antiviral
59 responses by type I and III interferons (IFN-I/III), which are also central for ZIKV control and
60 *in utero* transmission[10-15]. All cells possess signaling pathways designed to trigger the
61 production of IFN-I/III and IFN-stimulated genes (ISGs) upon viral infection. Their effects are
62 potent and wide-ranging: direct inhibition of the viral life cycle at multiple steps and jumpstart
63 of the adaptive immune response. These antiviral responses are induced by the recognition of
64 specific viral motifs by host sensors, such as Toll-like receptors (TLR) or RIG-I-like receptors
65 (RLR), that mobilize cascade signaling. As for other flaviviruses, TLR3-induced signaling
66 reduces ZIKV replication[16]. Nonetheless, like virtually all human pathogenic viruses, ZIKV has
67 evolved the ability to modulate and counteract the IFN-I/III signaling and other host responses[1-
68 4], likely through interactions with host proteins. The mutation rate of ZIKV is expected to be

69 around 10^{-4} to 10^{-5} mutation per site per replication in accordance with other flaviviruses, since
70 the catalytic site of the NS5 polymerase is well-conserved among flaviviruses[17]. This mutation
71 rate ensures high genetic diversity within hosts, and adaptability of viral populations. Adaptive
72 mutations that improve the fitness of the virus can do so by improving the viral machinery,
73 optimizing the interactions with proviral host factors, or inhibiting antiviral factors. Given the
74 limited size of the viral genome, trade-offs between these three strategies, and how they play
75 out in the human and mosquito hosts, are expected.
76
77 Previous studies on ZIKV evolution focused on the ability of the virus to maintain robust
78 replication in the context of alternate human/mosquito hosts[18]. Nonetheless, clinical studies
79 have demonstrated the ability of arboviruses, including ZIKV, to replicate and last for several
80 months in a subset of infected patients (*e.g.,* detection of viral genome in plasma, urine and
81 semen)[19-24], underlining the need to better understand the possible outcome(s) of viral evolution
82 in the human host.
83
84 Molecular tracking of arbovirus evolution in the host is greatly complicated by the error rate of
85 the polymerases used for sequencing, which makes it difficult to distinguish between low
86 frequency genetic variants that have appeared during evolution and sequencing errors. To
87 overcome this challenge, methods for accurate identification of ultra-rare and low-frequency
88 genetic variants have recently been developed. Especially, the CirSeq method, which involved
89 generating of tandem repeats from circularized RNA templates, has proved successful to
90 analyze the viral genetic diversity of poliovirus[25] and of DENV[26], as well as the landscape of
91 transcription errors in eukaryotic cells[27].
92

93 Here, we adapted this methodology to the arbovirus ZIKV, to reveal how it can adapt to the
94 human host cell environment and to investigate the genetic interactions involved. We conducted
95 an in-depth analysis of the evolution of viral populations (*i.e.,* population of genomes isolated
96 from infected cell supernatants) through serial passages of a Brazilian ZIKV strain (collected
97 at the epidemics) in human cell cultures. This led us to uncover a phenotypic change linked to
98 higher viral spread, *via* increased specific infectivity, which is associated to viral resistance to
99 TLR3-induced antiviral responses. Bioinformatic analyses showed that specific ZIKV variants
100 increased in frequency in temporal association with this phenotypic adaptation, and the
101 corresponding mutations were functionally validated.
102
103 **Results**
104
105 **Acquisition of increased specific infectivity during experimental evolution.**
106 We performed experimental evolution in a human hepatocyte cell line (Huh7.5.1. cells) as they
107 have been previously described to be highly efficient in ZIKV production and non-responsive
108 to viral products, as those cells are deficient for different antiviral response sensors[28-31]. A
109 Brazilian patient isolate (ZIKV Pernambuco PE243; KX197192) was chosen, which means that
110 the starting inoculum is more diverse than if it had been derived from a biological clone.
111 Experimental evolution was performed by serial passaging of ZIKV: at each passage (up to 18
112 passages), viral populations harvested at 3 days post-infection were used to infect naïve cells
113 (**Fig. 1A**, schema on the left side). Quantification of infectious virus produced over the course
114 of the experimental evolution showed that viral production increased during serial viral
115 passaging and reached a plateau by days 21-to-27 (**Fig. S1A**). Increased viral productions were
116 observed for all 3 independent runs of experimental evolution and within a similar timeframe
117 (**Fig. S1A**). Likewise, quantifications of intracellular and extracellular viral RNA levels

118 confirmed the augmentation of ZIKV replication over time (**Fig. S1B**). Passaging of the virus
119 from runs 2 and 3 has been stopped once the plateau of viral production was reached.
120
121 To address how passaged viral populations adapted to human host cells, we first studied their
122 ability to initiate infection as compared to the parental virus by quantifying the specific
123 infectivity, defined as the probability for one physical virion to initiate infection (as described
124 in Materials and Methods). We demonstrated an increase of the specific infectivity of the viral
125 populations harvested over the course of all independent runs of experimental evolution (**Fig.**
126 **1A**). The maximum level of specific infectivity was observed by days 21-to-24 in the
127 independent runs of experimental evolution followed by a plateau. The trend of increased viral
128 production (**Fig. S1A**) appeared to be simultaneous and in proportion with the augmentation of
129 specific infectivity (**Fig. 1A**), suggesting that the ability of ZIKV to adapt in this defined host
130 environment primarily most likely occurred *via* an increased capacity to initiate infection.
131
132 **Increased specific infectivity of passaged virus enables resistance to TLR3-induced**
133 **antiviral response.**
134 Next, we tested the capacity of these passaged viral populations to propagate when submitted
135 to the host antiviral responses. Huh7.5.1 cells are known to be deficient for TLR3-induced
136 signaling[29], enabling specific induction of the antiviral response by complementation *via*
137 ectopic expression of wild type (WT) TLR3 (**Fig. S1C**). Treatment of the WT TLR3 expressing
138 cells by poly(I:C), a mimetic of the intermediate double-stranded RNA produced during viral
139 replication, led to a robust ISG upregulation at mRNA and protein levels, by using ISG15 and
140 MxA as representative ISGs. ISG15 is considered to be part of the early response, being
141 preferentially induced directly by transcriptional factors such as IRF3, downstream of sensor-
142 induced signaling, and by IFN-I/III receptor-induced signaling. MxA expression belongs to the

143 late response regulated mostly through IFN-I/III receptor-induced signalling[32] (**Fig. S1D-E**). In
144 contrast, the parental Huh7.5.1 cells, without ectopic WT TLR3, did not respond to poly(I:C)
145 (**Fig. S1D-E**). This demonstrated that WT TLR3 expression renders our cell model responsive
146 to the TLR3 agonist poly(I:C). Specificity of the TLR3-induced ISG response was also
147 confirmed by the absence of ISG upregulation upon poly(I:C) treatment of cells expressing
148 TLR3 with a deletion of the Toll/interleukin-1 receptor (TIR) domain of the cytosolic tail
149 (ΔTIR-TLR3), necessary for recruiting the downstream TIR domain-containing adaptor
150 inducing IFNβ (TRIF)[33] (**Fig. S1C-D**).
151 By using this set up, we showed that poly(I:C)-induced TLR3 signaling greatly decreases
152 replication of the parental virus, but not that of passaged viral populations harvested at day 51
153 (**Fig. 1B**). Importantly, viral RNA levels measured at 72 hours post-infection were similar in
154 the TLR3-activated cells *versus* the non-activated cells for the passaged virus (**Fig. 1B**, right
155 curves). We quantified the inhibition of viral replication by activated-TLR3 signaling for viral
156 populations harvested at different time points (**Fig. S1F**). In keeping with the results shown in
157 **Fig. 1B**, TLR3-induced antiviral response inhibited the parental virus replication by up to 90%
158 at 72 hours post-infection. The resistance to TLR3-mediated inhibition was already observed
159 for the viral populations harvested at 12 days and at the different MOIs applied (**Fig. S1F**), thus
160 likely independent of the infection of one cell by infectious units containing several virions[34].
161
162 The resistance to TLR3-induced antiviral response can result from a faster onset of infection
163 by the passaged viral populations. In such a scenario, ongoing replication before the
164 establishment of a robust antiviral response in host cells would out-compete the latter. In
165 agreement with this hypothesis, when comparing similar infection set-up (MOI 0.05) - in the
166 absence of TLR3-induced response - the replication rate of passaged viral populations was
167 faster compared to the parental virus (**Fig. 1B**, comparing the slope from 24-to-48 hours post-

168 infection). Similar observations were made for viral populations harvested at distinct late time
169 points of the experimental evolution, including at days 27, 33 and 36 (**Fig. S1G** and *data not*
170 *shown*). These observations suggest that the passaged viral populations could outcompete host
171 antiviral response by an earlier replication onset post-inoculation.
172
173 **Increased viral replication depends on the targeted cell type but is independent of the ISG**
174 **response**
175 Faster replication, and thus a faster expression of viral proteins compared to host antiviral
176 effectors may be cell-type specific. To test this hypothesis, we assessed the infection efficiency
177 of passaged virus across different cell environments. First, to assess the infection speed by the
178 passaged viral population compared to the parental virus, we quantified the size of infectious
179 foci formed in a given timeframe, as reflecting the propagation speed *via* rounds of infection.
180 Foci formed upon infection of Huh7.5.1 cells by the long-term passaged viral population were
181 significantly larger than those formed by parental virus (**Fig. 1C,** left panels and **Fig. S2A**,
182 upper panels). This confirmed the increased infection speed of the viral populations in the cell
183 type used for experimental evolution of ZIKV. The opposite was observed in the simian Vero
184 cells (**Fig. 1C and S2A,** left panels). To further analyze this cell type-specific phenotype, we
185 performed a kinetic analysis of viral replication. The replication rate of passaged viral
186 populations significantly increased (approx. 10-fold) compared to parental virus in Huh7.5.1
187 cells (**Fig. 1D**, upper panel), with similar ZIKV RNA levels observed early after infection,
188 likely reflecting the viral input. In contrast, the replication rates of the adapted viral population
189 did not increase compared to parental virus in Vero cells (**Fig. 1E,** upper panel). The kinetic
190 analysis of the ISG response (representative MxA and ISG56 mRNAs) demonstrated a
191 difference between the cell types (**Fig. 1D-E,** lower panels and **Fig S2C-D**). An early ISG
192 response was triggered upon infection of Huh7.5.1 cells by parental virus, yet only observed

193 with MOI ≥ 0.1, and vanished for the viral populations harvested later in the course of the
194 experimental evolution (**Fig. 1D and S1G-H**). In Vero cells, infection by either viral
195 populations led to similar ISG responses (**Fig. 1E** and **S2D**). This differential profile of ISG
196 response between the two cell types might result from the absence in Vero cells of a response
197 to an activating signal contained in the supernatants harvested at early time-points of the
198 experimental evolution, or alternatively from a qualitatively and/or quantitatively different cell
199 entry pathways in these cell types.
200 To discriminate these possibilities, we broadened the phenotypic analysis to alternative human
201 cell types. Similar to Huh7.5.1 cells, the replication rate of adapted viral populations was clearly
202 increased in HEK-293 cells compared to the parental virus (**Fig. 1F**, upper panel). Unlike
203 Huh7.5.1 cells, neither parental virus nor adapted viral populations induced an ISG response in
204 HEK-293 cells (**Fig. 1F** and **S2E**), indicating that higher replication rate of adapted viral
205 populations *versus* parental virus can occur independently of an ISG response. The results also
206 demonstrated that the increased propagation rate of the adapted viral populations is not
207 restricted to a unique cell type. To further confirm that the increased infection rate by the
208 adapted viral population is independent from the early ISG response, we compared the viral
209 replication kinetics in both human U6A cells deficient for STAT2 (*i.e.,* an ISG transcription
210 regulator of the signaling pathway induced by IFN-I and III), and U6A cells complemented for
211 STAT2 expression, referred to as STAT2-U6A cells (**Fig. 1G,** upper panels). STAT2
212 expression potentiates the ISG response upon ZIKV infection with an accelerated response
213 observed at 48 hours post-infection in STAT2-U6A cells as opposed to the poor response of
214 the corresponding U6A cells (**Fig. 1G** and **S2F**) and consistently, decreased viral replication in
215 STAT2-U6A cells compared to the corresponding U6A cells (**Fig. 1G**, comparing the upper
216 panels). In both the U6A cells and STAT2-U6A cells, the levels of both viral replication and
217 the ISG response were similar for the adapted *versus* the parental viral populations (**Fig. 1G**).

218 Taken together, our results obtained in cell types failing to respond to ZIKV infection by ISG
219 upregulation, such as HEK-293 cells, or STAT2-U6A cells competent for ISG induction,
220 suggested that the increased infection rate of the adapted viral populations is independent of the
221 extent of ISG upregulation.
222 We also tested the rate of viral replication in macrophages differentiated from monocytes
223 isolated from healthy human blood donors (**Fig. 1H** and **S2B**), representing an *in vitro* cell
224 model closely related to the cell type targeted *in vivo* by ZIKV[35-38]. Interestingly, we observed
225 that the propagation of the passaged viral populations was significantly abrogated in this
226 cellular model, compared to the parental virus (**Fig. 1H**). Conversely, ISG induction was readily
227 detected in response to the parental virus, but not for the passaged viral population (**Fig. 1H**
228 **and S2G**).
229 These comparative analyses with different cell types suggested that the viral adaptations leading
230 to higher infection rate are independent of the ISG response, and that the ability of passaged
231 viral populations to initiate infection most likely involves viral-host interactions and entry
232 pathways that depend on the cellular environment.
233
234 **Genetic diversity and evolution of the viral populations determined by deep-sequencing**
235 **analysis.**
236 To delve into the mechanism underlying viral adaptation, we analyzed the genomic diversity
237 of the viral populations harvested in the course of evolutionary experiments. Sanger sequencing
238 and next-generation deep-sequencing methods provide generally consensus sequences and
239 cannot accurately detect low frequency variants, which can nevertheless be very important
240 functionally. To bypass this limitation we adapted the "CirSeq" method that reduces next-
241 generation sequencing errors[25] to sequence viral variants at various time points of our three
242 runs of experimental evolution. We used total RNA derived from approx. 5x10E6 to 5x10E7

243 secreted infectious particles per sample. Bioinformatic analyses revealed that CirSeq was
244 successful, with tandem repeat occurrence ranging from 59% to 94% (mean 84%) per read, and
245 repeat sizes ranging from 33 to 97 bases (mean size 50 to 60 bases). In addition, coverage
246 quantification revealed that our methodology using 200-PE runs with HiSeq 2500XL led to a
247 read number/depth as high as 2x10E5 (**Fig. 2A** and **S3A**). Minimal coverage, as defined as read
248 numbers below 1000, was detected at only 55 positions out of all the analyzed libraries and the
249 mean coverage per position across all analyzed experimental conditions ranged from
250 approximatively 10E4 to 10E5 (**Fig. S3B**), thus allowing an in-depth analysis of viral
251 populations. Moreover, the profiles of viral genome coverage across all analyzed samples were
252 similar (**Fig. 2A and S3A**). These results thus validate the reproducibility of our experiments
253 and imply that composition and/or secondary structure of certain viral genome segments had a
254 limited impact on the sequencing depth. Next, we analyzed the mutation frequencies over the
255 course of the three independent runs of experimental evolution (**Fig. 2B-C** and **S3C**). We
256 validated that the consensus sequence obtained by deep-sequencing of the parental inoculum
257 used for our experimental evolutions perfectly matches the clinical isolate reference sequence
258 (GenBank ID KX197192). To identify mutations whose frequency significantly varied through
259 time, for each run, we computed position-wise standard deviations of the frequency of the most
260 frequent variant. As shown in **Fig. S4A**, the standard deviations are distributed in two
261 categories: *i/* a high density of very low frequency variants, likely corresponding to
262 polymorphisms due to the error-prone feature of the NS5 polymerase and/or generated during
263 library preparation, despite the CirSeq protocol and *ii/* low density of polymorphisms showing
264 higher standard deviations, corresponding to mutations that have reached a high frequency
265 during at least one time-point of one experiment. Variants with standard deviation above a
266 threshold of 0.1 or 0.02 showed that 4 or 7 variants, respectively, had high standard deviations
267 in two or three runs (**Fig. S4B-C).** We provide a representation of variant frequencies at day 18

268 that highlights the similarity of the frequency profiles between independent experiments (**Fig.**
269 **2C**).
270
271 **Identification of candidate adaptive mutations in ZIKV E and NS4B proteins.**
272 We further analyzed the mutations that passed the standard deviation thresholds, considering
273 them as candidate adaptive mutations (**Fig. 2D-F and S4B-C**). Among the 4 variants that
274 passed the 0.1 threshold, 3 positions do so in a number (n) = 3 of experiments (positions 1786,
275 2341, 7173); 1 additional position does so in n = 2 (position 2194) (**Fig. 2D-E**). Two additional
276 positions are found when the threshold is set to 0.02 for n = 3 (positions 5663, 10007) (**Fig.**
277 **2F**). To know if some of these variants were already present at the beginning of the
278 experimental evolution, we sequenced the parental viral population. We compared the
279 frequency of our variants of interest to the frequency of all other variants in the library, and
280 computed percentiles. The minor variant frequencies were 12.5% for C1786T, 0.06% for
281 C2194T, 0.3% for C2341T, 3.3% for T5663C, 0.3% for T7173C, 2.0% for C10007T. This
282 places variants C1786T, T5663C, and C10007T among the 1% most frequent variants, C2341T
283 between the 33 and 34% most frequent variants, T7173C between the 44 and 45% most frequent
284 variants, and C2194T between the 97 and 98% most frequent variants (*i.e.,* between the 1 and
285 2% rarest variants). Variants C1786T, T5663C, and C10007T were thus present in the initial
286 inoculum at the highest frequency. The particularly high frequency of non-synonymous
287 C1786T mutation, leading to the reversion at position 270 of E envelope protein from V to A
288 and present in the clinical isolate (ZIKV PE243, KX197192)[39] likely results from the initial
289 viral amplification of the clinical isolate in Vero cells. Variant C2194T was particularly rare,
290 and variants C2341T and T7173C had unremarkable frequencies at the start of the experiment,
291 in the middle of the distribution of variant frequencies. Among the variants, C2341T, T7173C,
292 C1786T and C2194T are non-synonymous mutations within the coding part of the ZIKV

293 polyprotein and are thus most likely to have phenotypic effects on the virus. The frequencies
294 of mutations C1786T and C2194T do not vary in a manner correlated with the increased specific
295 infectivity of the evolved virus and thus are unlikely to be linked to this phenotype. Mutations
296 C2341T and T7173C may however be linked to the evolved phenotype.
297 The frequency of the C2341T mutation increased from days 6-to-12 (**Fig. 2D**), thus preceding
298 both the augmented specific infectivity and the acquisition of resistance to TLR3-induced
299 antiviral response (**Fig. 1A and S1A-B**). This corresponds to a substitution of amino acid S-to-
300 L, at position 455 of the E envelope protein. The T7173C mutation also became the majority
301 variant by day 12, 12 and 18 in the three independent experiments (**Fig. 2D**). This corresponds
302 to a Y-to-H substitution at position 87 in NS4B. Together, our bioinformatic analyses identified
303 two point mutations in E and NS4B proteins that might be involved in the phenotype of viral
304 adaptation observed during experimental evolution.
305
306 **The S455L mutation in the E viral protein recapitulates both the increased specific**
307 **infectivity and the resistance to TLR3 antiviral response.**
308 To functionally validate the two candidate adaptive mutations, the S455L substitution in the E
309 envelope protein and the Y87H substitution in NS4B were introduced alone, or in combination,
310 into recombinant ZIKV. First, we generated a ZIKV molecular clone (referred to as 'ref no
311 mut') corresponding to the clinical isolate used for experimental evolution by introducing R99G
312 and Y122H substitution in NS1 in the previously reported molecular clone BeH819015[40]. Of
313 note, we showed the absence of significant impact of these NS1 mutations on viral spread and
314 specific infectivity (**Fig. S5A-C**).
315
316 The efficiency of transfection of the different mutants was comparable to that of the reference,
317 as determined by ZIKV intracellular levels at 6 hours post-transfection (**Fig. S5G**). While the

318 Y87H substitution in NS4B showed no significant difference as compared to the reference clone
319 (ref no mut), the S455L substitution in the E protein significantly increased the rate of viral
320 replication when present alone or in combination with Y87H in NS4B (**Fig. 3A**). The S455L
321 substitution in E also recapitulated the enhanced specific infectivity observed for viral
322 populations evolved in the course of the evolution experiments (compare **Fig. 3B** to **1A**).
323
324 We examined the speed of viral propagation *via* new rounds of infection by measuring the size
325 of infectious foci formed in a limited timeframe (as in **Fig. 1C)**. We observed a clear increase
326 upon infection with the E S455L mutant as compared to the reference in Huh7.5.1 cells (**Fig.**
327 **3C** and **S5H,** left panels). In contrast, in Vero cells, the size of infectious foci formed upon
328 infection by the same E S455L mutant was reduced as compared to the reference (**Fig. 3C** and
329 **S5H,** right panels). This is consistent with the observations for the viral populations adapted in
330 the course of experimental evolution described earlier (**Fig. 1C**). We demonstrated that, similar
331 to the adapted viral populations obtained in the experimental evolution, the E S455L mutant
332 resists inhibition by TLR3-induced antiviral response (**Fig. 3D**, right curves), as opposed to the
333 reference clone (**Fig. 3D**, left curves). These results demonstrate that the S455L substitution,
334 located in proximity to the trans-membrane domain of ZIKV E protein and interacting with
335 membrane lipid and/or with the viral protein M (**Fig. S7**), recapitulates the phenotype of high
336 infectivity observed with the adapted viral population.
337
338 Given that the E S455L mutation causes improved viral replication both in the presence and
339 absence of TLR3-induced signaling, we asked whether accelerated onset of replication could
340 provide a unifying mechanism behind these observations. To this end, we modeled this
341 mechanism mathematically and compared the model and experimental data. The model (**Fig.**
342 **4A**) describes the infection of susceptible cells, which, after a time delay, enter a phase of

343 productive viral replication. Produced infectious virions can then infect other susceptible cells.
344 Individual cells in this multiscale model will have their own time course of ZIKV replication,
345 depending on the time of infection (**Appendix 1, mathematical model**). We determined the
346 kinetic parameters of the model (including infection rate, delay to productive replication,
347 replication rate, virion production rate and infection-induced cell death rate) by fitting the
348 model to two sets of experimental data (**Fig. 3 and S1G**). The effects of both evolution and
349 introduction of the E S455L point mutation were modeled by allowing the delay to productive
350 replication to be different from parental and reference strains, respectively, keeping all other
351 parameters identical. The model captured the data (**Fig. 4B-C**) with a well-constrained set of
352 parameter values **(Fig. S6)**. The delay to productive replication was shortened from around 8
353 hours for the (parental or reference) controls to 6 hours or less for the adapted or mutated strains
354 (**Fig. 4D**). Our model thus suggests that the E S455L substitution allows a reduced time-lag
355 between infection and viral replication, sufficient to cause increased viral spread, and probably
356 explaining the resistance to TLR3-induced antiviral response.
357
358 **Discussion**
359
360 Experimental evolution in a stable and controlled cell environment uncovered ZIKV
361 adaptations leading to the avoidance of induced host antiviral response *via* an increased specific
362 infectivity. Bioinformatic analysis of the viral genome evolution pinpointed two candidate
363 mutations whose frequencies increased concomitantly with the acquisition of the adapted
364 phenotype. We further showed that viral adaptation augmented infection independently of the
365 ISG response and in specific cellular environments, thus suggesting a modulation of the host-
366 virus interaction involved at the early step of ZIKV infection. Consistently, we uncovered a key
367 determinant in the envelope (E) protein that both augments the probability for one physical

368 virus to initiate infection in certain cell environments and mediates resistance to the TLR3-
369 induced antiviral response.
370
371 **Proposed mechanism for the resistance to antiviral response**
372 Recent reports have shown the ability of *Flaviviridae* to acquire adaptive mutations over the
373 course of long passaging in cell culture[41,42]. Based notably on the sequencing of molecular
374 clones, these studies concluded that adaptation to the host cell environment could occur through
375 multiple, independent mutational pathways. In the present study, we found instead that in
376 independent experiments the same mutations increased in frequency after a few passages in a
377 controlled cellular environment. This difference between previous studies[41,42] and our results
378 might be explained by a strong selective advantage provided by mutations present at a low
379 frequency in the initial viral populations. Previous studies in the context of the infection by
380 hepatitis C virus, known to lead to high diversity upon chronic infection in patients, showed
381 that mutation frequencies underwent waves of increases and decreases during experimental
382 evolution[41]; in some experiments, waves can be observed in the frequency of some variants
383 (*e.g.,* C2341T and C1786T; **Fig. 2D-F**).
384 Since the acquisition of both increased specific infectivity and resistance to TLR3-induced
385 antiviral response were temporally associated during the experimental evolution, and knowing
386 that the E S455L point mutation reproduces both phenotypes, we propose that the two
387 phenotypes are causally linked. Like other flaviviruses, ZIKV has evolved inhibitory
388 mechanisms against antiviral responses, including blockage of the response to type I and III
389 IFNs by NS5-mediated degradation of STAT2[1-4]. Our results suggest that viral adaptation
390 occurred *via* a reduction of the time delay prior to the onset of productive infection rather than
391 by directly modulating virus-mediated inhibition of the host antiviral detection. Firstly, we
392 showed that improved viral replication of the adapted populations was independent of the

393 amplitude of the ISG response, as demonstrated using HEK-293 where ISG upregulation is
394 absent. Secondly, the comparison of the U6A-STAT2 cells *versus* the corresponding STAT2-
395 deficient U6A cells indicated that the adapted viral populations did not differentially inhibit
396 ISG responses induced by Jak/STAT signaling compared to the parental virus. Thirdly, kinetic
397 analysis demonstrated that the adapted viral populations showed enhanced resistance to TLR3-
398 mediated inhibition at later time-points post-infection, and with higher viral input. Fourthly,
399 our mathematical model of ZIKV infection and replication showed that a reduced time-lag prior
400 to the establishment of productive infection may suffice to cause increased virus spread, and
401 hence associated resistance to the antiviral response. In accordance with the model, we
402 demonstrated increased specific infectivity of the adapted viral populations and E S455L
403 mutated recombinant virus, suggesting adaptation of host-ZIKV interaction at the early step of
404 infection.
405 As this viral adaptation is cell type-restricted, the adapted viral populations likely become fitter
406 by interacting with host factors preferentially present or conversely at a limiting level in certain
407 cell types at early stages of infection. These host factors can include some machineries of the
408 entry pathway such as endosome acidification and/or expression of cell type restricted receptor
409 (*e.g.,* DC-sign primarily expressed in dendritic and myeloid cells)[16,43-45]. Overall, our results
410 indicate that the observed viral adaptation most likely results from an improved ability to
411 rapidly and efficiently establish viral replication. Hence, faster accumulation of viral products
412 in newly infected cells can overwhelm and/or bypass the host antiviral factors.
413
414 **Identification of a key determinant in ZIKV E responsible for increased specific**
415 **infectivity.**
416 Bioinformatic analysis identified the mutations S455L in E and Y87H in NS4B as reproducibly
417 associated to the adapted phenotype. These mutations were at low frequency at the beginning

418 of the experiment; then, presumably strong fitness benefits allowed them to rise in frequency
419 through selection. The functional analysis by insertion of the candidate mutations into a ZIKV
420 molecular clone demonstrated that the S455L substitution in E controls both improved viral
421 spread and resistance to TLR3-induced antiviral signaling, probably through accelerated
422 replication onset according to our mathematical model.
423 The cellular study model has reduced complexity compared to whole organisms (*e.g.,* unique
424 cell type, restricted diversity of antiviral sensing and absence of adaptive response and physico-
425 chemical constraints such as blood flow). This enabled the discovery of viral adaptation
426 improving fitness in a stable and well-defined cell environment[46,47]. At the organism scale,
427 interferences between several selective pressures could impede the acquisition of some
428 optimized interaction(s). This may explain why the adaptive mutation at position 455 of the E
429 envelope protein is found only in 2 out of 519 ZIKV full genome sequences isolated from
430 patients (NCBI Virus Variation Resource on December 21 2020) and is not maintained in a
431 mutational scanning using a different cell type and strain[48]. Alternatively, adaptive mutations
432 may enable the virus to infect different cell types with differential efficiency. In this scenario,
433 the most frequent ZIKV variants would be well adapted to infect the major target cell types,
434 while low-frequency variants would be better at infecting specific cell types. In accordance
435 with this, comparison between various cell types showed that the adapted phenotype is cell
436 type-restricted, indicating that the targeted host factor(s) is differentially involved in viral
437 replication depending on the cell type. This also suggests cell type-specific trade-offs for the
438 virus: adaptation to a particular cell environment limits infectivity in other cellular
439 environments. Consistently, we observed that the adapted virus had a reduced ability to infect
440 the monocyte-derived macrophages, possibly as a result of different cell entry mechanisms.
441 Such trade-offs is in line with the results of Duggal et al. who found that mutations appeared

442 during passages in Vero cells reduce the pathogenicity of the virus in mouse[49]; it is unclear
443 differences in cell types or in host species underlie this observation.
444
445 The position 455 is located at the C-terminus of the stem region, next to the transmembrane
446 domain of the E protein[50-52]. In the structure of E (5ire.pdb reference)[52], the position 455 is
447 located near the membrane surface, in an environment of nearly exclusively polar and
448 hydrophobic residues (**Fig. S7**). Its side chain hydroxyl does not show interactions with the
449 transmembrane helix of the M protein facing it, but is likely involved in lipid headgroup
450 interactions. A mutation from polar to hydrophobic in this region thus has the potential to
451 change the membrane anchorage and insertion of the protein. *In silico* estimation of the
452 structural changes induced by the S455L substitution suggested that this adaptive mutation
453 might enhance the stability of the E protein *via* its interaction with apposed segments in M or,
454 more likely, with specific membrane lipids (**Fig. S7**). Consistent with the known regulatory
455 function of this segment in flaviviruses[50-55], we propose that the E S455L substitution modulates
456 the membrane fusion process and/or E protein membrane incorporation. Thus, albeit future
457 analysis is required to test this hypothesis, the structural property of the E S455L substitution
458 is in agreement with the demonstration of an ability to better initiate de novo infection (*i.e.,*
459 increased specific infectivity) of the E adaptive mutant. Therefore, we propose a working model
460 of viral adaptation *via* an optimized interaction with the host machinery involved at an early
461 stage of infection, via E membrane interactions, and likely by modulating its function in
462 membrane fusion. In turn, higher infectivity would result in the accelerated accumulation of
463 viral products and/or altered entry pathway that overcome or bypass host antiviral responses.
464 In conclusion, we studied viral adaptation in a stable and controlled cellular environment
465 leading us to link phenotype to mutation. The methods of deep sequencing and bioinformatics
466 set up here allowed the identification of mutations at low frequency arising across the entire

467 viral genome, as expected for RNA viruses with a high error rate during replication. The profiles
468 of frequency increase of the variants suggested that they were *bona fide* variants, not
469 experimental artifacts of our sequencing protocol. Focusing on the mutations that reached high
470 frequency and whose increased frequency is temporally associated to the adapted phenotype,
471 we functionally validated one of them by recapitulating the observed phenotypes, thus showing
472 the power of our approach combining viral evolution, mathematic modeling and functional
473 assays.

474 **Methods**

475

476 **Biological materials**

477 Huh7.5.1 cells (kindly provided by Dr F.V. Chisari; Scripps Research Institute), Vero E6 cells
478 (kindly provided by Dr M Bouloy; Institut Pasteur) and HEK-293 cells (ATCC CRL-1573)
479 were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% Fetal
480 Bovine Serum (FBS), penicillin (100 units per mL; (U)/mL), streptomycin (100 µg/mL), non-
481 essential amino acids (100 nM) and Hepes (10 mM) at 37°C/5% CO_2, as previously described[56].
482 The fibrosarcoma U6A cells (kindly provided by Dr M. Köster; Helmholtz-Zentrum für
483 Infektionsforschung) expressing or not human STAT2[57] were cultured in DMEM supplemented
484 with 10% FBS, penicillin (100 U/mL), streptomycin (100 µg/mL), non-essential amino acids
485 (100 nM), hepes (10 mM) and sodium pyruvate (1mM) at 37 °C in 5 % CO_2. All cell culture
486 reagents were purchased from Life Technologies. Monocytes were isolated from blood from
487 healthy adult human volunteers, obtained according to procedures approved by the
488 'Etablissement Français du sang' (EFS) Committee. PBMCs were isolated using Ficoll-
489 Hypaque density centrifugation and monocytes were positively selected using anti-CD14
490 microbeads (MACS Miltenyi Biotec) according to the manufacturer's instructions. Then,
491 2x10e5 CD14+ cells were differentiated into monocyte-derived macrophages in RPMI 1640
492 culture medium with 10% FBS, penicillin (100 U/mL), streptomycin (100 µg/mL), non-
493 essential amino acids (100 nM), hepes (10 mM), sodium pyruvate (1mM), 0.05 mM of β-
494 mercaptoethanol, Glutamine (2 mM) and supplemented with 100 ng/mL human M-CSF
495 cytokine (Peprotech). Three days post-isolation, half of the media was refreshed by similar
496 media supplemented with 100 ng/mL human M-CSF cytokine. Six days later, the differentiation
497 was assessed by FACS analysis. The infections of macrophages derived from monocytes were
498 performed at day 7 post-isolation. The U6A and HEK-293 cells were infected by the indicated

499 viral populations with MOI of 0.1. Primary macrophages derived from monocytes were infected
500 with MOI of 0.2.
501 To establish a controlled set up for the evolutionary experiment and in line with our strategy to
502 work in a cell environment non-responsive to viral products, we used Huh7.5.1. In addition, to
503 enable the subsequent study of the influence of activated antiviral response on the evolved
504 ZIKV populations, such as TLR3-induced signaling by poly(I:C) treatment, we transduced
505 Huh7.5.1 cells with a WT TLR3 and, as reference control, a mutant TLR3 invalidated for
506 signaling via a deletion of the Toll/interleukin-1 receptor (TIR) domain of the cytosolic tail
507 (ΔTIR-TLR3)[29,33]. The retroviral-based vectors expressing WT TLR3 and ΔTIR-TLR3 were
508 produced in HEK-293T cells and used to transduce Huh7.5.1 cells, as we previously
509 described[58]. Transduced cells were selected by adding blasticidin at 10 µg/ml (Invivogen).
510
511 **Reagents**
512 The antibodies used for immunostaining were mouse anti-E glycoprotein (4G2) kindly
513 provided by P. Despres (PIMIT, Université de La Réunion-INSERM France) and anti-Mouse
514 IgG (H+L) Alexa Fluor 555 Secondary Antibody (Life Technologies). The antibodies used for
515 western blot detection were anti-ISG15 (Santa Cruz), anti-flag (clone M2; Sigma-Aldrich) and
516 mouse anti-actin (clone AC74; Sigma-Aldrich). The antibodies used for FACS analysis were
517 APC-conjugated anti-CD163 (clone REA812, MACS Miltenyi Biotec), PE-conjugated mouse
518 anti-HLA-DR (clone LN3, Invitrogen); PE-Cy7-conjugated mouse anti-CD11b (clone ICRF44,
519 BD Biosciences), Pacific Blue-conjugated mouse anti-CD14 (clone 63D3, Biolegend). Other
520 reagents include TLR3 agonist poly(I:C) (Invivogen); Hoescht (Life Technologies) and cDNA
521 synthesis and qPCR kits (Life Technologies). The reagents for the preparation of the libraries
522 include: Vivaspin unit 100 MWA (Sartorius); Pico Total RNA kit (Agilent), T4 RNA ligase
523 (New England Biolabs); Superscript III (Life Technology); Ultra DNA Library Prep kit for

524 Illumina (New England Biolabs): AMPure XP beads (Agencourt); High Sensitivity DNA kit
525 (Agilent).
526
527 **Analysis of intracellular and extracellular RNA levels**
528 RNAs were isolated from cells or supernatants harvested in guanidinium thiocyanate citrate
529 buffer (GTC) by phenol/chloroform extraction procedure as previously described[56]. The
530 efficiency of RNA extraction and reverse transcription-real-time quantitative PCR (RT-qPCR)
531 was controlled by the addition of carrier RNAs encoding Xef1α (xenopus transcription factor
532 1α) *in vitro* transcripts in supernatants diluted in GTC buffer. ZIKV RNA, Xef1α and
533 intracellular MxA, ISG15, ISG56 and glyceraldehyde-3-phosphate dehydrogenase (GAPDH)
534 mRNA levels were determined by RT-qPCR using cDNA synthesis and qPCR kits (Life
535 Technologies) and analyzed using StepOnePlus Real-Time PCR system (Life Technologies),
536 using previously described primers[31]. Primers were designed to the region or prM protein
537 (nucleotid position: 689-to-782) of the ZIKV genome and are described in the **Table S1**.
538 Extracellular and intracellular ZIKV RNA levels were normalized for Xef1α and GAPDH RNA
539 levels, respectively.
540
541 **Analysis of extracellular infectivity**
542 Infectivity titers in supernatants were determined by end-point dilution using Huh7.5.1 cells.
543 Foci forming unit (ffu) were detected 48 hours after infection by anti-E envelope protein
544 specific immunofluorescence using 4G2 clone, as previously described[31]. Briefly, Huh7.5.1
545 cells were fixed with 4% PFA and permeabilized by incubation for 7 minutes in PBS containing
546 0.1% Triton. Cells were then blocked in PBS containing 3% BSA for 15 minutes and incubated
547 for 1 hour with mouse anti-E (clone 4G2) hybridoma supernatant, 1:100-dilution in PBS
548 containing 1% BSA. After 3 washes with PBS, cells were incubated 1 hour with secondary

549 Alexa 555-conjugated anti-mouse antibody (1µg/mL) and Hoechst dye (200ng/ml) in PBS
550 containing 1% BSA. Foci countings were performed using Zeiss Axiovert 135 or Olympus
551 CKX53 microscopes.
552
553 **Determination of the specific infectivity**
554 The specific infectivity was determined as the ratio of the extracellular infectivity titer value
555 (determined by ffu counting) relative to the extracellular ZIKV RNA value (as absolute number
556 genome equivalent quantified by RT-qPCR) for each individual experiment, and expressed as
557 the mean values ± standard deviation (SD).
558
559 **Foci size analysis**
560 Foci sizes were determined by infection using Huh7.5.1 cells and Vero cells with a
561 carboxymethylcellulose (CMC) overlay in 12-well plate. Forty-eight hours later, cells were
562 washed with PBS and fixed by 4% PFA. Foci were detected by anti-E protein (4G2 clone), as
563 above-described. Imaging by Olympus CKX53 microscope of randomly picked foci were
564 processed using Image J software. Size index was determined as the ratio of E staining intensity
565 relative to Hoechst staining with normalization of each intensity on the mean intensity of three
566 randomly selected cell unit. More than 80 foci were acquired for parental and passaged virus
567 comparison, and more than 50 foci for the molecular clone analysis.
568
569 **Immunostaining and FACS analysis of monocyte-derived macrophages**
570 Surface immunostainings were performed at day 6 post-isolation of the CD14+ cells from
571 PBMCs. After a 15 minute-incubation step with human Fc blocking reagent (MACS Miltenyi
572 Biotec), cell surface markers were detected by a 30 minute-incubation at 4°C with 4 µg/mL
573 Pacific Blue-conjugated mouse anti-CD14 (clone 63D3, Biolegend), a 1:20 dilution of APC-

574 conjugated mouse anti-CD163 (clone REA812, MACS Miltenyi Biotec), 0.14 µg/mL PE-
575 conjugated mouse anti-HLA-DR (clone LN3, Invitrogen) and a 1:20 dilution of PE-Cy7-
576 conjugated mouse anti-CD11b (clone ICRF44, BD Biosciences), diluted in staining buffer
577 (PBS - 1% FBS). Flow cytometric analysis was performed using a BD FACS Canto II and the
578 data were analyzed with Flow Jo software (Tree Star).
579
580 **Western blot analysis**
581 Cell lysates of the indicated cells were extracted using lysis buffer (150 mM NaCl 50 mM Tris
582 HCl pH 8, 1% NP40, 0.5% Deoxycholate, 0.1% Sodium dodecyl sulfate) and analyzed by
583 Western blotting using anti-ISG15 (sc-50366, H150; Santa cruz; 1 µg/mL), anti-Flag (clone M2
584 Sigma-Aldrich; 1 µg/mL), and actin (clone AC74; Sigma-Aldrich; 1 µg/mL), followed by
585 secondary HRP-coupled antibodies and chemiluminescence.
586
587 **Serial passaging of viral populations, deep sequencing and bioinformatic analysis for the**
588 **selection of variants.**
589 A clinical isolate of ZIKV from Brazil collected from a patient during the epidemics
590 (PE243_KX197192) was amplified using Vero E6 cells[39]. Supernatants were harvested and
591 filtrated (0.45 µm) before storage at -80°C. The serial passaging of viral populations was
592 performed by inoculation of Huh7.5.1 cells expressing ΔTIR TLR3. The target cells were
593 seeded at $2x10^6$ cells in two T175 cell culture flasks (Corning) the day prior infection (*i.e.*, at
594 MOI 0.1 for the first infection, and at MOI 0.01 for subsequent viral passaging because of the
595 limited viral yield from early passages, and MOI 0.1 for D18-to-D54 of round #1). At 3 days
596 post-infection, the supernatants were harvested and their infectivity was determined by foci
597 counting. Next, harvested supernatants were used to infect naïve cells, which were similarly
598 modified. The serial passaging of viral populations was repeated up to 18 times (for a total up

599 to 54 days). To avoid contamination of the libraries by cellular RNA released from dying cells,
600 an additional amplification step was performed by passaging of the supernatant with the
601 corresponding cells for 2 days, using MOI of 0.1. Supernatants (SN, 10ml, approx. 10E7
602 secreted infectious particles per sample were used) were concentrated using Vivaspin units with
603 100 MWA cut-of (Sartorius) by centrifugation at 3 000 g for 20 min at room temperature. Viral
604 RNAs were extracted from the concentrated SN by phenol-chloroform extraction procedure as
605 previously described[58]. Next, RNAs were fragmented by sonication using Covaris M220 (peak
606 incident 50 Watts; duty cycle 20%; 200 cycles per burst, time 200 seconds at 4°C) using Covaris
607 M220. Fragmented RNAs were concentrated with isopropanol precipitation, followed by
608 analysis with Bioanalyzer (Agilent Technology) using Pico Total RNA Chip. The median size
609 of the RNA fragments was around 150 nucleotides. We then generated tandem repeats of the
610 fragment to reduce the error rate of next-generation sequencing[25]. The steps of circularization of
611 RNA fragments and retrotranscription were adapted from a previously described protocol[25].
612 Briefly, circular RNAs were obtained by ligation using T4 RNA ligase (New England Biolabs),
613 followed by phenol-chloroform extraction. The retrotranscription was performed using
614 Superscript III (Life Technology). Then, libraries were prepared using the Ultra DNA Library
615 Prep kit for Illumina (New England Biolabs). Additional steps of clean-up were performed
616 using AMPure XP beads, ratio 1:0.8 (Agencourt) to remove free adaptors (done before the step
617 of PCR enrichment of adaptor-ligated DNA) and to remove the free index adaptors (after the
618 PCR enrichment of adaptor-ligated DNA). The quality of the libraries was assessed by
619 Bioanalyzer (Agilent Technology) using High Sensitivity DNA Chip. The libraries were
620 quantified using NEB Next Libraries Quantification kit for Illumina (New England Biolabs)
621 and multiplexed at equimolarity. Multiplexed libraries were sequenced using HiSeq 2500XL
622 (Illumina), using a 200-PE run at the EMBL Genecore Facility (Heidelberg, Germany).
623

624 **Bioinformatic analysis of the sequencing libraries**

625 Reads were quality-checked and trimmed of sequencing adapters and then mapped using
626 PEAR[59] and in-house software (https://github.com/Kobert/viroMapper). PEAR was used to
627 merge the repeats in the forward and reverse reads, to improve quality scores and minimize
628 sequencing errors. Viromapper automatically returns a table with statistics per position such as
629 quality quantiles, coverage, raw numbers of A, C, G, T bases, and quality-weighted counts and
630 proportions of A, C, G, T bases. These tables were then read and analyzed using Python scripts
631 developed in house. To evaluate the efficiency of our protocol to generate tandem repeat, we
632 measured repeat sizes within forward and reverse reads separately using a script developed in
633 house, based on samples of 5000 reads in 22 read files. All libraries contained repeats. The
634 worst library (HV5GLBCXY_ZIKV_17s006139-1-1_DREUX_lane1CirseqD3) had repeats in 59% of its
635 forward reads and 67% of its reverse reads, with sizes ranging between 33 and 97 bases, with
636 a mean at 50. In all other libraries, repeat sizes had similar distributions, but repeats were
637 present in up to 94% of the reads. Datasets for the sequence data file available in SRA with
638 accession numbers SRX9704326-SRX9704344; bioproject PRJNA686429.

639

640 **Bioinformatics analysis for the selection of variants of interest**

641 We were interested in variants absent in the starting clone and that reached high frequencies in
642 the experiments. We reasoned that variants of interest should vary in frequency in the course
643 of several independent experiments. To detect variants of interest, for each site and in each
644 experiment, we computed the standard deviation of the frequency of the major variant. The
645 distribution of these standard deviations is represented in **Fig S3B**. Based on these standard
646 deviations, we counted the number of experiments in which a given site had a standard
647 deviation above some threshold. Two threshold values were used: 0.015, and 0.1 **(Fig S3)**. Sites
648 that passed the threshold in several experiments were examined further (**Fig S3**). A Jupyter

649 notebook reproducing these analyses is available
650 (https://github.com/Boussau/Notebooks/blob/master/Notebooks/Viromics/FiguresAndAnalys
651 es_NoTLR3.ipynb).
652
653 **Introduction of selected mutation and analysis of adapted mutants in the ZIKV infectious**
654 **clone.**
655 Mutations were introduced in the genomic length ZIKV infectious clone cDNA, pCCI-SP6-
656 ZIKV BeH819015 plasmid (kindly provided by Dr A. Merits)[40]. For adequate comparison with
657 clinical isolate PE243[39], the R99G and Y122H substitution in NS1 were also introduced in
658 pCCI-SP6-ZIKV BeH819015 plasmid.
659 In a first step, introduction of Y87H in NS4B were performed by overlap-PCR (OL-PCR) using
660 mutagenic primers, as described previously[56]. The OL-PCR fragments were purified and
661 transferred using In-Fusion HD Cloning Plus CE technology (638917, Ozyme) into the pCCI-
662 SP6-ZIKV BeH819015 plasmid cleaved using BamHI restriction enzyme.
663 In a second step, the S455L substitution in E protein and R98G and Y121H substitution in NS1
664 were obtained by generating a synthetic gene *i.e.,* fragment from nucleotides 1507-to-3375 from
665 5'UTR of BeH819015 sequence, which containing the NS1 mutations combined or not with E
666 S455L substitution. Next, the different fragments were amplified by PCR and transferred by
667 using In-Fusion HD Cloning Plus CE technology into the pCCI-SP6-ZIKV BeH819015
668 plasmid which contained or not the Y87H substitution in NS4B that were cleaved/opened using
669 Avr2 and Pml1 restriction enzyme.
670 All the sequences of the primers used for the mutagenesis and In-Fusion HD Cloning Plus CE
671 technology are described in the **Table S1**. The expected sequences for the PCR derived regions
672 and the presence of mutations were validated by Sanger sequencing.
673

674 **Analysis of adapted mutants in the ZIKV infectious clone.**

675 *In vitro* RNA transcripts were prepared from the parental and mutated pCCI-SP6-ZIKV
676 BeH819015 plasmids. Briefly, these plasmids were linearized with AgeI. After DNA
677 extraction, *in vitro* RNA transcripts were generated using mMESSAGE mMACHINE SP6 Kit
678 (Ambion) followed by Lithium precipitation, as previously described[56]. *In vitro* RNA
679 transcripts were transfected into Huh7.5.1 cells using Lipofectamine 3000 transfection reagent
680 (Life Technologies), following the manufacturer instruction. At 24, 48 and 72 hours post-
681 transfection, the supernatants were collected for the quantification of viral RNA and infectious
682 titer. At 6 hours post-transfection, the cells were washed with PBS and harvested to determine
683 RNA levels.

684

685 **Statistical analysis of viral parameters**

686 Data are presented as the mean values ± standard deviation (SD). The Figure legend section
687 reports the number of independent experiments. Statistical analysis was performed using R
688 software environment for statistical computing and graphics, version 3.3.2. For levels of viral
689 RNA and specific infectivity, the values were considered relative to reference/parental for each
690 independent experiment and analysis using an one-way ANOVA on ranks (Kruskal−Wallis
691 test), as we previously performed[31]. When the test was significant, we used the Nemenyi *post*
692 *hoc* test for multiple comparisons of mean rank sums to determine which contrast(s) between
693 individual experimental condition pairs was significant. Statistical analysis of foci size was
694 performed by pairwise comparisons using Wilcoxon rank sum test. Statistical analysis of the
695 kinetics of viral replication of the different viral populations was done using a mixed linear
696 model, which takes into account the levels at the different time-point before those indicated in
697 the table. Data considered significant demonstrated adjusted p-value by False Discovery Rate
698 (FDR) less than 0.05, as previously[31].

References

1 Zimmerman, M. G. *et al.* STAT5: a Target of Antagonism by Neurotropic Flaviviruses. *Journal of virology* **93** (2019).

2 Serman, T. M. & Gack, M. U. Evasion of Innate and Intrinsic Antiviral Pathways by the Zika Virus. *Viruses* **11** (2019).

3 Riedl, W. *et al.* Zika Virus NS3 Mimics a Cellular 14-3-3-Binding Motif to Antagonize RIG-I- and MDA5-Mediated Innate Immunity. *Cell host & microbe* **26**, 493-503 e496 (2019).

4 Wang, B. *et al.* Structural basis for STAT2 suppression by flavivirus NS5. *Nat Struct Mol Biol* **27**, 875-885 (2020).

5 Brasil, P. *et al.* Zika Virus Outbreak in Rio de Janeiro, Brazil: Clinical Characterization, Epidemiological and Virological Aspects. *PLoS Negl Trop Dis* **10**, e0004636 (2016).

6 Cao-Lormeau, V. M. *et al.* Guillain-Barre Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. *Lancet* (2016).

7 de Oliveira, W. K. *et al.* Infection-related microcephaly after the 2015 and 2016 Zika virus outbreaks in Brazil: a surveillance-based analysis. *Lancet* **390**, 861-870 (2017).

8 Robbiani, D. F. *et al.* Risk of Zika microcephaly correlates with features of maternal antibodies. *The Journal of experimental medicine* (2019).

9 Cugola, F. R. *et al.* The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* **534**, 267-271 (2016).

10 Yockey, L. J. *et al.* Type I interferons instigate fetal demise after Zika virus infection. *Sci Immunol* **3** (2018).

11 Li, C. *et al.* Zika Virus Disrupts Neural Progenitor Development and Leads to Microcephaly in Mice. *Cell Stem Cell* (2016).

12 Caine, E. A. *et al.* Interferon lambda protects the female reproductive tract against Zika virus infection. *Nat Commun* **10**, 280 (2019).

13 Jagger, B. W. *et al.* Gestational Stage and IFN-lambda Signaling Regulate ZIKV Infection In Utero. *Cell host & microbe* **22**, 366-376 e363 (2017).

14 Miner, J. J. *et al.* Zika Virus Infection during Pregnancy in Mice Causes Placental Damage and Fetal Demise. *Cell* **165**, 1081-1091 (2016).

15 Dang, J. *et al.* Zika Virus Depletes Neural Progenitors in Human Cerebral Organoids through Activation of the Innate Immune Receptor TLR3. *Cell Stem Cell* (2016).

16 Hamel, R. *et al.* Biology of Zika Virus Infection in Human Skin Cells. *Journal of virology* **89**, 8880-8896 (2015).

17 Bollati, M. *et al.* Structure and functionality in flavivirus NS-proteins: perspectives for drug design. *Antiviral Res* **87**, 125-148 (2010).

18 Moser, L. A. *et al.* Growth and adaptation of Zika virus in mammalian and mosquito cells. *PLoS Negl Trop Dis* **12**, e0006880 (2018).

19 Barzon, L. *et al.* Virus and Antibody Dynamics in Travelers With Acute Zika Virus Infection. *Clin Infect Dis* **66**, 1173-1180 (2018).

20 Stone, M. *et al.* Zika virus RNA and IgM persistence in blood compartments and body fluids: a prospective observational study. *Lancet Infect Dis* (2020).

21 Sanchez-Montalva, A., Salvador, F. & Molina, I. Persistence of Zika Virus in Body Fluids - Final Report. *N Engl J Med* **380**, 198 (2019).

22 Paz-Bailey, G., Rosenberg, E. S. & Sharp, T. M. Persistence of Zika Virus in Body Fluids - Final Report. *N Engl J Med* **380**, 198-199 (2019).

23 Hoarau, J. J. *et al.* Persistent chronic inflammation and infection by Chikungunya arthritogenic alphavirus in spite of a robust host immune response. *J Immunol* **184**, 5914-5927 (2010).
24 Ozden, S. *et al.* Human muscle satellite cells as targets of Chikungunya virus infection. *PLoS ONE* **2**, e527 (2007).
25 Acevedo, A., Brodsky, L. & Andino, R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* **505**, 686-690 (2014).
26 Dolan, P. T. *et al.* Principles of dengue virus evolvability derived from genotype-fitness maps in human and mosquito cells. *Elife* **10** (2021).
27 Gout, J. F. *et al.* The landscape of transcription errors in eukaryotic cells. *Sci Adv* **3**, e1701484 (2017).
28 Sumpter, R. *et al.* Regulating intracellular antiviral defense and permissiveness to hepatitis C virus RNA replication through a cellular RNA helicase, RIG-I. *Journal of virology* **79**, 2689-2699 (2005).
29 Wang, N. *et al.* Toll-like receptor 3 mediates establishment of an antiviral state against hepatitis C virus in hepatoma cells. *Journal of virology* **83**, 9824-9834 (2009).
30 Vicenti, I. *et al.* Comparative analysis of different cell systems for Zika virus (ZIKV) propagation and evaluation of anti-ZIKV compounds in vitro. *Virus research* **244**, 64-70 (2018).
31 Assil, S. *et al.* Plasmacytoid Dendritic Cells and Infected Cells Form an Interferogenic Synapse Required for Antiviral Responses. *Cell host & microbe* **25**, 730-745 e736 (2019).
32 Grandvaux, N. *et al.* Transcriptional profiling of interferon regulatory factor 3 target genes: direct involvement in the regulation of interferon-stimulated genes. *Journal of virology* **76**, 5532-5539 (2002).
33 Alexopoulou, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* **413**, 732-738 (2001).
34 Sexton, N. R. *et al.* Genome Number and Size Polymorphism in Zika Virus Infectious Units. *Journal of virology* **95** (2021).
35 Carlin, A. F. *et al.* Deconvolution of pro- and antiviral genomic responses in Zika virus-infected and bystander macrophages. *Proceedings of the National Academy of Sciences of the United States of America* **115**, E9172-E9181 (2018).
36 Jurado, K. A. *et al.* Zika virus productively infects primary human placenta-specific macrophages. *JCI Insight* **1** (2016).
37 Lang, J. *et al.* An hPSC-Derived Tissue-Resident Macrophage Model Reveals Differential Responses of Macrophages to ZIKV and DENV Infection. *Stem Cell Reports* **11**, 348-362 (2018).
38 Quicke, K. M. *et al.* Zika Virus Infects Human Placental Macrophages. *Cell host & microbe* **20**, 83-90 (2016).
39 Donald, C. L. *et al.* Full Genome Sequence and sfRNA Interferon Antagonist Activity of Zika Virus from Recife, Brazil. *PLoS Negl Trop Dis* **10**, e0005048 (2016).
40 Mutso, M. *et al.* Reverse genetic system, genetically stable reporter viruses and packaged subgenomic replicon based on a Brazilian Zika virus isolate. *The Journal of general virology* **98**, 2712-2724 (2017).
41 Gallego, I. *et al.* Broad and Dynamic Diversification of Infectious Hepatitis C Virus in a Cell Culture Environment. *Journal of virology* **94**, e01856-01819 (2020).
42 Perales, C. *et al.* Response of Hepatitis C Virus to Long-Term Passage in the Presence of Alpha Interferon: Multiple Mutations and a Common Phenotype. *Journal of virology* **87**, 7593-7607 (2013).

43 Chen, J. *et al.* AXL promotes Zika virus infection in astrocytes by antagonizing type I interferon signalling. *Nat Microbiol* **3**, 302-309 (2018).
44 Acosta, E. G., Castilla, V. & Damonte, E. B. Alternative infectious entry pathways for dengue virus serotypes into mammalian cells. *Cell Microbiol* **11**, 1533-1549 (2009).
45 Wells, M. F. *et al.* Genetic Ablation of AXL Does Not Protect Human Neural Progenitor Cells and Cerebral Organoids from Zika Virus Infection. *Cell Stem Cell* **19**, 703-708 (2016).
46 Arai, Y. *et al.* Novel Polymerase Gene Mutations for Human Adaptation in Clinical Isolates of Avian H5N1 Influenza Viruses. *PLoS Pathog* **12**, e1005583 (2016).
47 Chambers, B. S., Li, Y., Hodinka, R. L. & Hensley, S. E. Recent H3N2 influenza virus clinical isolates rapidly acquire hemagglutinin or neuraminidase mutations when propagated for antigenic analyses. *Journal of virology* **88**, 10986-10989 (2014).
48 Sourisseau, M. *et al.* Deep Mutational Scanning Comprehensively Maps How Zika Envelope Protein Mutations Affect Viral Growth and Antibody Escape. *Journal of virology* **93** (2019).
49 Duggal, N. K. *et al.* Mutations present in a low-passage Zika virus isolate result in attenuated pathogenesis in mice. *Virology* **530**, 19-26 (2019).
50 Kostyuchenko, V. A. *et al.* Structure of the thermally stable Zika virus. *Nature* **533**, 425-428 (2016).
51 Prasad, V. M. *et al.* Structure of the immature Zika virus at 9 A resolution. *Nat Struct Mol Biol* **24**, 184-186 (2017).
52 Sirohi, D. *et al.* The 3.8 A resolution cryo-EM structure of Zika virus. *Science (New York, N.Y* **352**, 467-470 (2016).
53 Klein, D. E., Choi, J. L. & Harrison, S. C. Structure of a dengue virus envelope protein late-stage fusion intermediate. *Journal of virology* **87**, 2287-2293 (2013).
54 Zhang, X. *et al.* Cryo-EM structure of the mature dengue virus at 3.5-A resolution. *Nat Struct Mol Biol* **20**, 105-110 (2013).
55 Stiasny, K., Kiermayr, S., Bernhart, A. & Heinz, F. X. The membrane-proximal "stem" region increases the stability of the flavivirus E protein postfusion trimer and modulates its structure. *Journal of virology* **87**, 9933-9938 (2013).
56 Decembre, E. *et al.* Sensing of immature particles produced by dengue virus infected cells induces an antiviral response by plasmacytoid dendritic cells. *PLoS Pathog* **10**, e1004434 (2014).
57 Frahm, T., Hauser, H. & Koster, M. IFN-type-I-mediated signaling is regulated by modulation of STAT2 nuclear export. *Journal of cell science* **119**, 1092-1104 (2006).
58 Dreux, M. *et al.* Short-range exosomal transfer of viral RNA from infected cells to plasmacytoid dendritic cells triggers innate immunity. *Cell host & microbe* **12**, 558-570 (2012).
59 Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614-620 (2014).
60 Guedj, J., Dahari, H., Uprichard, S. L. & Perelson, A. S. The hepatitis C virus NS5A inhibitor daclatasvir has a dual mode of action and leads to a new virus half-life estimate. *Expert Rev Gastroenterol Hepatol* **7**, 397-399 (2013).
61 Rong, L. & Perelson, A. S. Mathematical analysis of multiscale models for hepatitis C virus dynamics under therapy with direct-acting antiviral agents. *Math Biosci* **245**, 22-30 (2013).
62 Goo, L. *et al.* Zika Virus Is Not Uniquely Stable at Physiological Temperatures Compared to Other Flaviviruses. *mBio* **7** (2016).

63 Zmurko, J. *et al.* The Viral Polymerase Inhibitor 7-Deaza-2'-C-Methyladenosine Is a Potent Inhibitor of In Vitro Zika Virus Replication and Delays Disease Progression in a Robust Mouse Infection Model. *PLoS Negl Trop Dis* **10**, e0004695 (2016).
64 Shakeri, F. & Dehghani, M. The method of lines for solution of the one-dimensional wave equation subject to an integral conservation condition. *Computers & Mathematics with Applications* **56**, 2175-2188 (2008).
65 Raue, A. *et al.* Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, 1923-1929 (2009).

Acknowledgements

We thank P. Despres (PIMIT, Université de La Réunion-INSERM France) for the anti-E 4G2 antibody; F.V. Chisari (Scripps Research Institute, La Jolla, CA) for Huh-7.5.1 cells; Dr M Bouloy (Institut Pasteur, Paris, France) for Vero E6 cells and by Dr M. Köster (Helmholtz-Zentrum für Infektionsforschung) for U6A cells. We are grateful to Y. Jaillais for critical reading of the manuscript. We acknowledge the contribution of SFR Biosciences (UMS3444/CNRS, US8/Inserm, ENS de Lyon, UCBL) facilities. This work was supported by grants from the "Agence Nationale pour la Recherche" (ANR-JCJC-EXAMIN), the ''Agence Nationale pour la Recherche contre le SIDA et les Hépatites Virales'' (ANRS-AO 2017-01), the European Union's Horizon 2020 Research and innovation program under "ZIKALLIANCE" (Grant Agreement no. 734548) and FINOVI foundation (AO11) to MD. The grants from "Fondation pour la recherche médicale'' (contract Bioinformatic analysis for research in biology, DBI20141231313 and from the "Agence Nationale pour la Recherche" LabEx Ecofect (Grant ANR-11-LABX-0048) to MD, BB and AB. AK is supported by the UK Medical Research Council (MC_UU_12014/8, MR/N017552/1).

Author Contributions: V.G., T.H., B.B. and M.D. designed research; V.G., E.H., K.K., S.R.T., E.D., C.G., P.V.M., M.P., S.M.-G. and L.S. performed research; V.G., E.H., K.K., S.R.T., E.D., C.G., P.V.M., A.K., A.B., M.P., S.M.-G. and L.S. contributed new

reagents/analytic tools; V.G., E.H., K.K., S.R.T., E.D., C.G., P.V.M., A.K., M.P., A.B., S.M.-G., L.S., T.H., B.B. and M.D. analyzed data; and V.G., B.B., and M.D. wrote the paper.

Competing Interest Statement: The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

877 **Appendix 1, Mathematical model**

878

879 **The multiscale model of ZIKV infection and replication.** To have a better understanding of
880 the impact of the delay of the virus replication on ZIKV dynamics we used both intracellular–
881 ZIKV GE·mL^{-1} and extracellular ZIKV measurements–FFU·mL^{-1}. The intracellular dynamics
882 of ZIKV RNA measured in bulk are the sum of replication dynamics in individual cells each of
883 which may have their own time point of infection. To capture this, we construct a multi-scale
884 model of infection, where the relative time course of viral replication is the same in each cell,
885 but the times of infection and, with a delay, of the start of replication, vary. Similar to previous
886 studies on HCV[60,61], we developed a system of coupled partial and ordinary differential
887 equations as follows. Susceptible cells (S) get infected by ZIKV with rate k_{inf}:

$$\frac{dS}{dt} = -k_{inf}\, S\, V \tag{1}$$

888

889 where V denotes the concentration of virions. Infected cells I, have an age of infection, a,
890 measuring how much time elapsed since infection. Their number is given by the age-structured
891 balance equation k_{icd}:

$$\frac{\partial I}{\partial t} + \frac{\partial I}{\partial a} = -k_{icd} \cdot I(a,t) \tag{2}$$

892 which also accounts for cell death with rate k_{icd}. For simplicity, cell proliferation is neglected.
893 We have the initial condition $I(a,0) = 0$ and the boundary condition $I(0,t) = k_{inf} \cdot S \cdot V$
894 representing the new infections. The dynamics of the virions is obtained by integrating over
895 infected cells of all ages:

$$\frac{dV}{dt} = k_{vp} \cdot \int_{0}^{72\,hrs} I(a,t) \cdot R(a)\, da - k_{\text{vd}} \cdot V \tag{3}$$

896

where infectious virions are produced with rate k_{vp} and decay with rate k_{vd}; the latter was set to 0.06 (h^{-1})[62].

Eq. 3 depends on the number of intracellular ZIKV RNA produced, which we describe by a logistic growth function:

$$R(a) = R_0 + \frac{R_{max}}{1 + R_{max} \cdot exp\left(-k_{rep} \cdot (a - \tau)\right)} \tag{4}$$

with R_0 being the initial number of ZIKV RNA in the cell upon infection, R_{max} the maximum number achieved by replication, k_{rep} the ZIKV replication rate and τ is the delay from infection to the onset of virus replication. The delay for the parental virus τ_p was set to 8 hours according to a previous report[63]. Finally, the number of ZIKV RNA per cells was calculated as:

$$R_{pc}(t) = \frac{\int_0^{72\ hrs} I(a,t) \cdot R(a)\ da}{Total\ cells(t)} \tag{5}$$

The model was simulated using the method of lines, discretizing first derivatives according to the upwind rule[64]. For this purpose, we used CVODES, a module from the SUNDIALS numerical simulation package, in the MATLAB environment.

To simulate ZIKV transfection using lipofectamine–as in the **Fig 5B**– we defined as a new state variable infectious agent (*IA*):

$$IA(t) = V(t) + Lipo(t) \cdot L2V \tag{6}$$

where the amount of the infectious agent is the sum of ZIKV virions and the RNA supplied by lipofectamine pulse. The scale factor *L2V* is estimated. The lipofectamine pulse is modeled using a combination of two logistic growth functions with different delays and signs,

$$Lipo(t) = 1/(1 + \exp(-5 \cdot (t - 0.1))) - 1/(1 + \exp(-5 \cdot (t - 7))), \quad (7)$$

with time in hours. Parameter estimation was conducted by minimizing the weighted least-squares of the simulated values versus the experimental data

$$wSSR = \sum_{i=1}^{N} \sum_{j=1}^{M} \left[\frac{y_{simulation_{i,j}} - y_{experiment_{i,j}}}{\sigma_{i,j}} \right]^2 \quad (8)$$

where j is the experiment number and i is the data point of the jth experiment. The sum of squared residuals is normalized with the standard deviation for each experimental data point $\sigma_{i,j}$. The logarithm of ZIKV RNA and ffu/ml^{-1} values with base 10 was used to estimate model parameters. We calculated the 95% confidence interval for the estimated parameter values using the profile-likelihood method[65] (**Fig S6**). The maximum-likelihood for the remaining parameters is calculated. Finally, the change in the maximum-likelihood ($\Delta\chi^2$) versus different values of the parameter of interest is plotted as profile-likelihood.

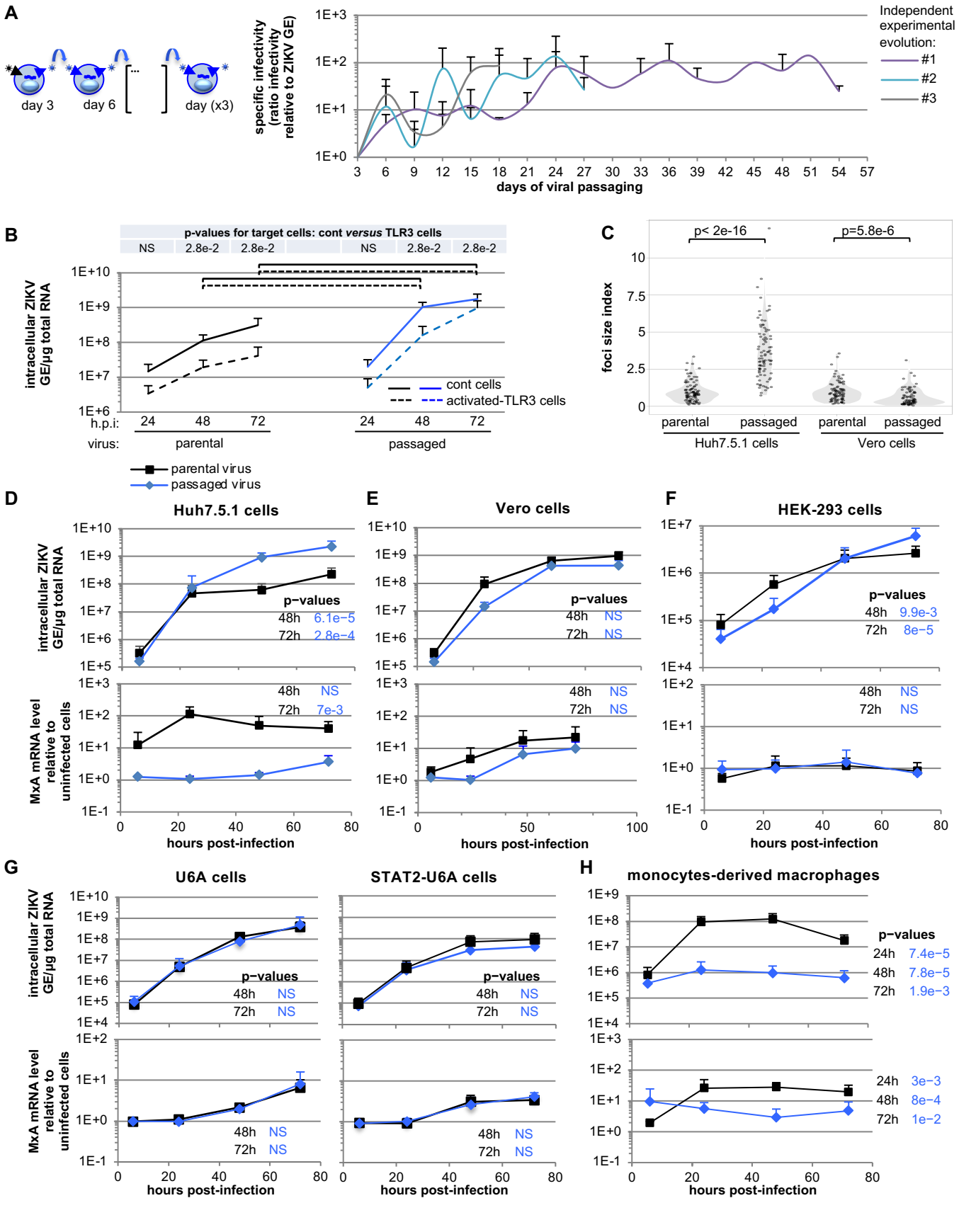

A
day 3
day 6
day (x3)
specific infectivity (ratio infectivity relative to ZIKV GE)
days of viral passaging
Independent experimental evolution:
#1
#2
#3
B
p-values for target cells: cont versus TLR3 cells
NS 2.8e-2 2.8e-2 NS 2.8e-2 2.8e-2
intracellular ZIKV GE/µg total RNA
h.p.i:
virus:
parental
passaged
cont cells
activated-TLR3 cells
parental virus
passaged virus
C
p< 2e-16
p=5.8e-6
foci size index
parental
passaged
Huh7.5.1 cells
Vero cells
D
Huh7.5.1 cells
E
Vero cells
F
HEK-293 cells
p-values
48h 6.1e-5
72h 2.8e-4
48h NS
72h 7e-3
48h NS
72h NS
48h 9.9e-3
72h 8e-5
MxA mRNA level relative to uninfected cells
hours post-infection
G
U6A cells
STAT2-U6A cells
H
monocytes-derived macrophages
24h 7.4e-5
48h 7.8e-5
72h 1.9e-3
24h 3e-3
48h 8e-4
72h 1e-2

Figure 1

Figure 1. Increased specific infectivity and resistance to TLR3-mediated inhibition of serially passaged viral populations. A. As shown on the schematic representation of the experimental procedure for serial passaging of ZIKV viral populations: every 3 days, viral supernatants were harvested, infectivity levels determined and, accordingly, used to infect naïve cells. Quantification of the specific infectivity (*i.e.*, defined as the probability for one physical virion to initiate infection) of the viral populations harvested over serially passaging in the course of 3 independent run of experimental evolution (referred to as #1, #2, #3 and stop at days 54, 21 and 18 respectively). Results are expressed as the ratio of the extracellular infectivity levels relative to extracellular ZIKV RNA levels and relative to day 3 for each independent passaging in control cells set to 1; 2-to-3 independent determination by RT-qPCR and infectious titrations for each evolution experiment; mean ± SD. **B.** Quantification of the intracellular ZIKV genome levels in kinetic analysis of serially passaged viral populations obtained at day 51 of the viral passaging *versus* parental virus assessed in activated-TLR3 cells (dotted lines) as compared to control cells (solid lines) ; at the indicated times post-infection at MOI 0.05; 4 independent experiments; mean ± SD. The statistical comparison of intracellular ZIKV GE levels for the same viral population at the same time point post-infection between activated-TLR3 and control cells is indicated in the table at the top of the graphs with indicated p-values and NS; $p>0.05$. The intracellular ZIKV GE levels at a same time post-infection between parental and passaged virus determined as significantly different ($p<0.05$) are indicated by brackets: dotted lines for the comparison of levels in activated-TLR3 cells and solid lines for the comparison of levels in control cells. **C.** Violin plot representation of the focus size in Huh7.5.1 cells and Vero cells at 48 hours post-infection by parental virus *versus* serially passaged virus harvested at day 51 of the viral passaging. The size index for each infectious focus is displayed by an individual dot; 5 independent measurements; statistical analysis as indicated by p-values. **D-H.** Kinetic quantification in Huh7.5.1 (**D**), Vero cells (**E**), HEK-293 cells (**F**), U6A cells and STAT2 expressing U6A cells (**G**) and macrophages derived from monocytes (**H**) post-infection by parental versus serially passaged viral population obtained at day 51 of the viral passaging. Results present the levels of intracellular ZIKV GE (upper panels) and MxA (lower panels) mRNA levels relative to the levels in non-infected cells, at the indicated time post-infection at MOI 0.1; 3-to-7 independent experiments; mean ± SD. The p-values of the statistical analysis of the kinetics performed using mixed linear model are indicated on the right side of the graphs, p-value are for the comparison of passaged viral population versus parental virus, and NS; $p>0.05$.

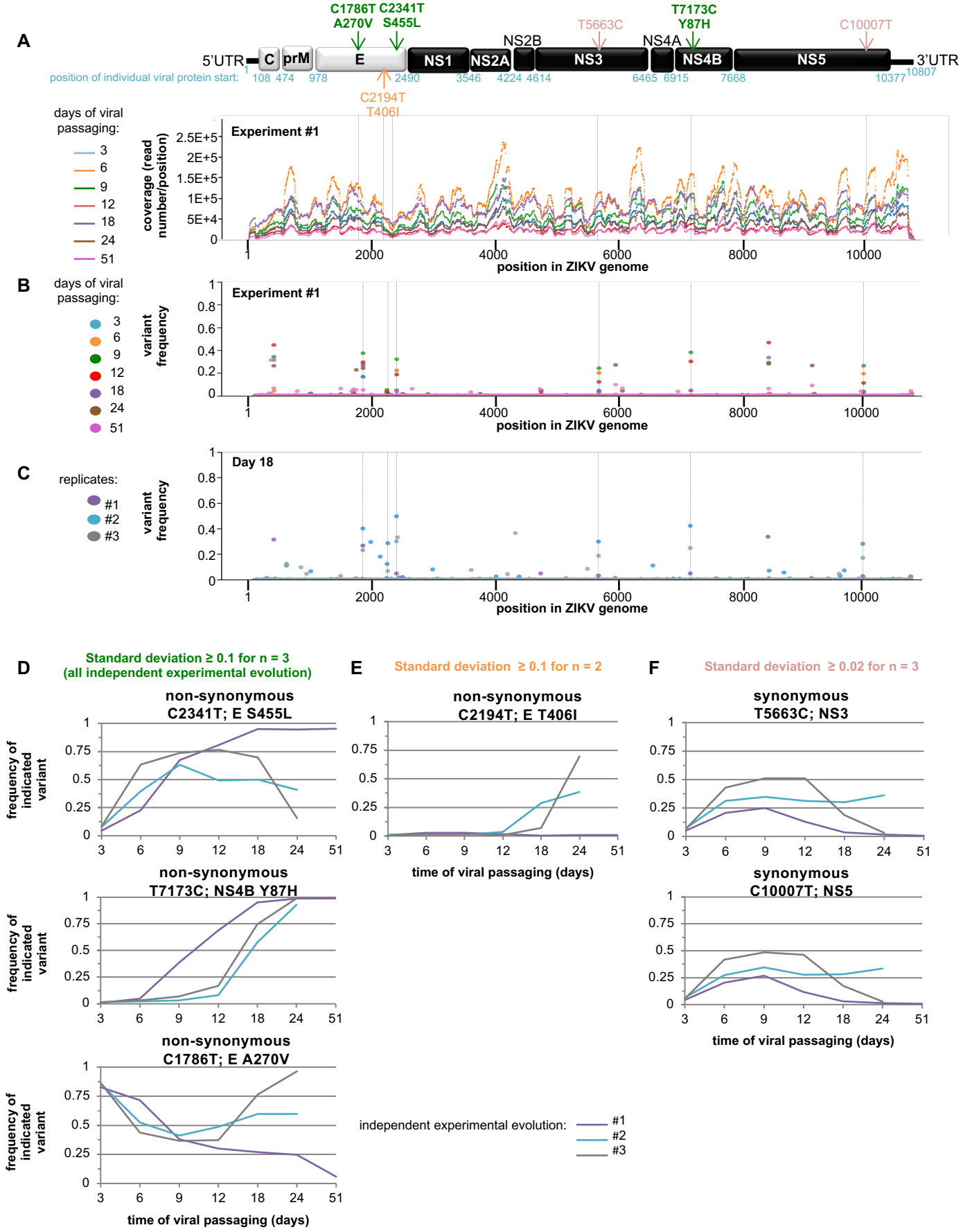

A
C1786T A270V
C2341T S455L
T5663C
T7173C Y87H
C10007T
5'UTR
C
prM
E
NS1
NS2A
NS2B
NS3
NS4A
NS4B
NS5
3'UTR
position of individual viral protein start:
108 474 978 2490 3546 4224 4614 6465 6915 7668 10377 10807
C2194T T406I
days of viral passaging:
3
6
9
12
18
24
51
Experiment #1
coverage (read number/position)
position in ZIKV genome
B
variant frequency
C
replicates:
#1
#2
#3
Day 18
D
Standard deviation ≥ 0.1 for n = 3 (all independent experimental evolution)
non-synonymous C2341T; E S455L
non-synonymous T7173C; NS4B Y87H
non-synonymous C1786T; E A270V
frequency of indicated variant
time of viral passaging (days)
E
Standard deviation ≥ 0.1 for n = 2
non-synonymous C2194T; E T406I
F
Standard deviation ≥ 0.02 for n = 3
synonymous T5663C; NS3
synonymous C10007T; NS5
independent experimental evolution:
#1
#2
#3

Figure 2

Figure 2. Bioinformatic analysis of the genetic evolution of viral populations obtained by next-generation sequencing. A. Coverage of the next-generation sequencing analysis along the ZIKV genome sequence of viral populations harvested at the indicated time points of the serial passaging of one representative independent run of the evolution experiment. Results are expressed as number of reads per position; schematic representation of ZIKV genome at the top. **B-C.** Time-course quantification of the frequency of the second most frequent variants at each position along ZIKV genome in the viral populations harvested in one representative independent run of evolution experiment (**B**) and in the viral populations harvested at day 18 in 3 independent runs of evolution experiments (**C**). Dotted lines indicate the positions in the viral genome with high standard deviations in several runs of experimental evolution, as defined in **Fig S4. C. D-F.** Time-course quantification of the frequency of variants determined by next-generation sequencing. The variants were selected when the standard deviations of their frequencies were: ≥ 0.1 for all the 3 independent runs of experimental evolution (n =3; n referred to one replicate of one condition at given time of harvest) (**D**); ≥ 0.1 for a minimum of 2 samples (**E**), and ≥ 0.02 for a minimum of 3 samples (**F**), with thresholds defined according to the density of variants relative to their frequency for the pool of all analyzed samples, as presented in Fig S4. The variants are indicated as nucleotide position (*e.g.,* C2340T), the corresponding viral protein (*e.g.*, E) and amino acid change for non-synonymous substitution (e.g., S455L); as also shown on the schematic representation of ZIKV genome organization (**A**).

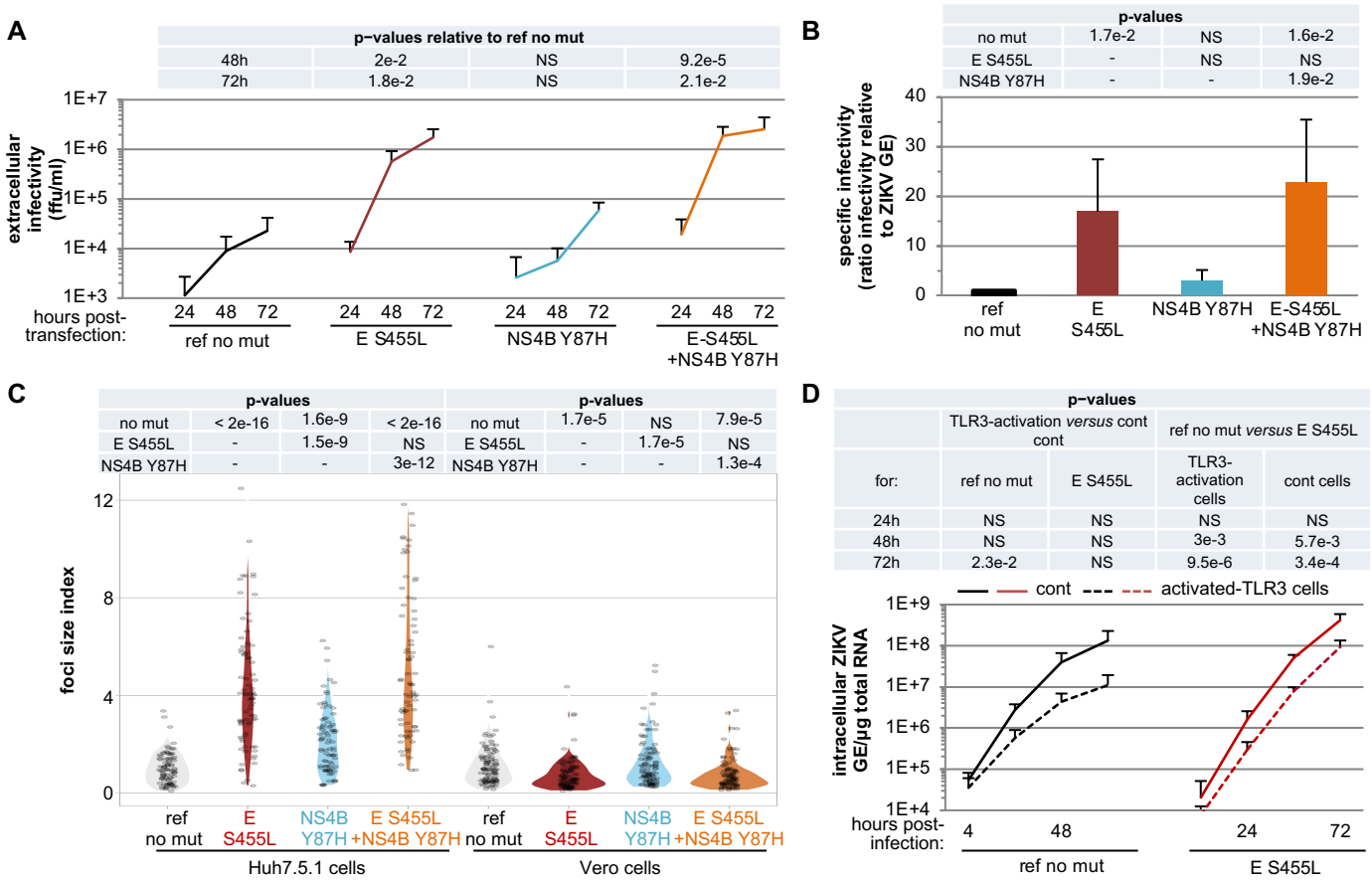


Figure 3. Introduction of the selected non-synonymous mutations into a ZIKV molecular clone. ZIKV genome bearing the selected mutations (*i.e.,* single S455L mutation in E, Y87H in NS4B and combined E S455L and NS4B Y87H mutations), and as a reference ZIKV genome without the mutation (*i.e.,* ref no mut), were transfected in Huh7.5.1 cells. **A.** Time-course quantification of infectious viral production at the indicated times post-transfection. Results of 4-to-6 independent experiments; mean ± SD; p-values as indicated in the table above the graph and relative to the reference (ref no mut). **B.** Quantification of the specific infectivity (as in **Fig. 1A**) in the viral supernatants harvested at 72 hours for ZIKV genome mutants or not. Results are the mean ± SD relative to the reference virus set to 1 for each independent experiment; 4-to-6 independent experiments. The p-values indicated in the table correspond to one-by-one comparisons of the condition/mutant displayed below in the graph with conditions/mutants indicated on the left side of the table. **C.** Analysis of the focus size index of the indicated mutated or reference ZIKV determined in Huh7.5.1 and Vero cells, as indicated, for supernatants harvested at 72 hours post-transfection. The quantifications are displayed by violin plots, determined as in **Fig. 1C**. Results of 4 independent experiments; p-values as indicated in the table above the graph. **D.** Time-course analysis of the replication of the E S455L mutant versus reference virus (ref no mut) assessed in activated-TLR3 Huh7.5.1 cells (dotted lines) as compared to control cells (solid lines). Quantification of the intracellular ZIKV genome levels at the indicated times post-infection at MOI 0.005; 4 independent experiments; mean ± SD. The statistical comparison of intracellular ZIKV GE levels for the same target cells and between mutant and reference virus is indicated in the table at the top of the graphs with indicated p-values and NS; p>0.05.

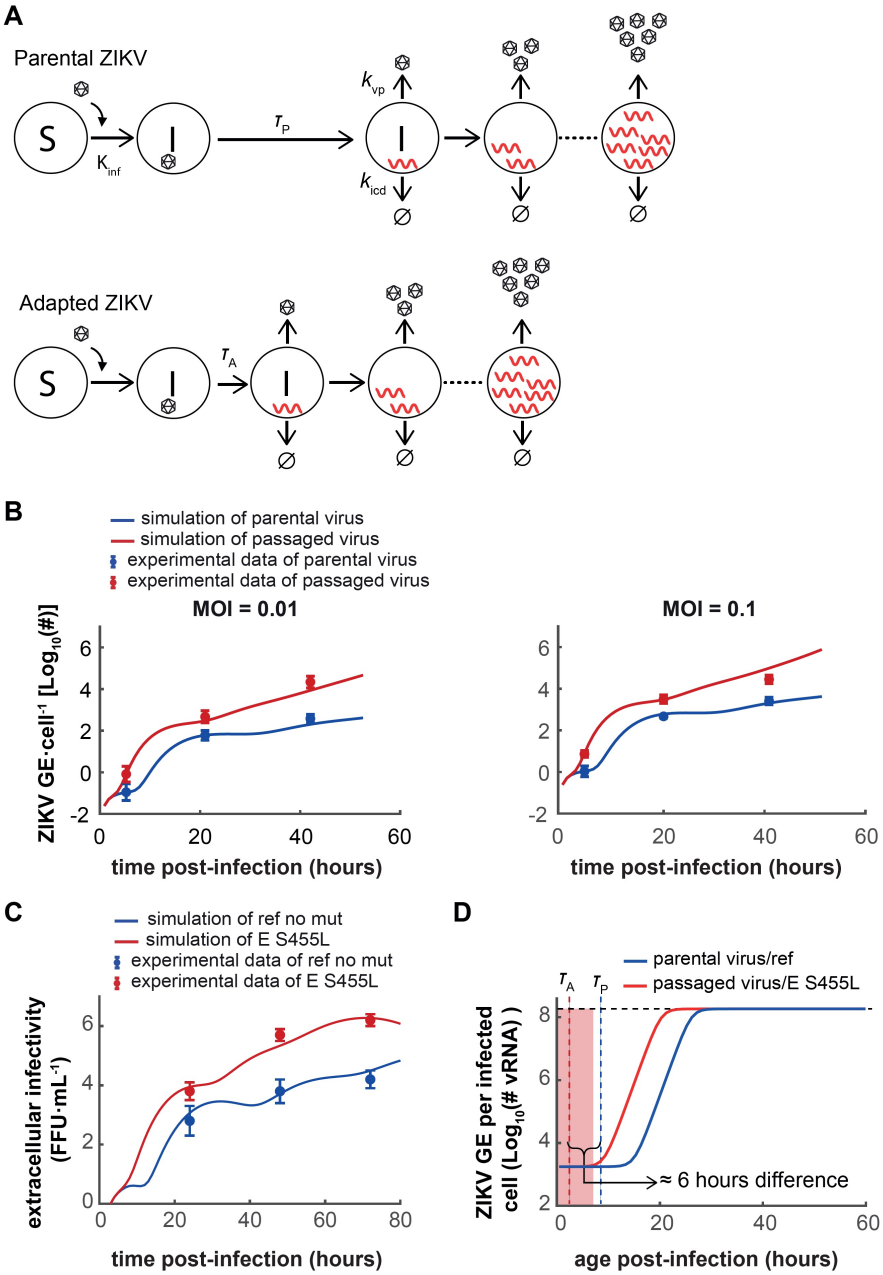
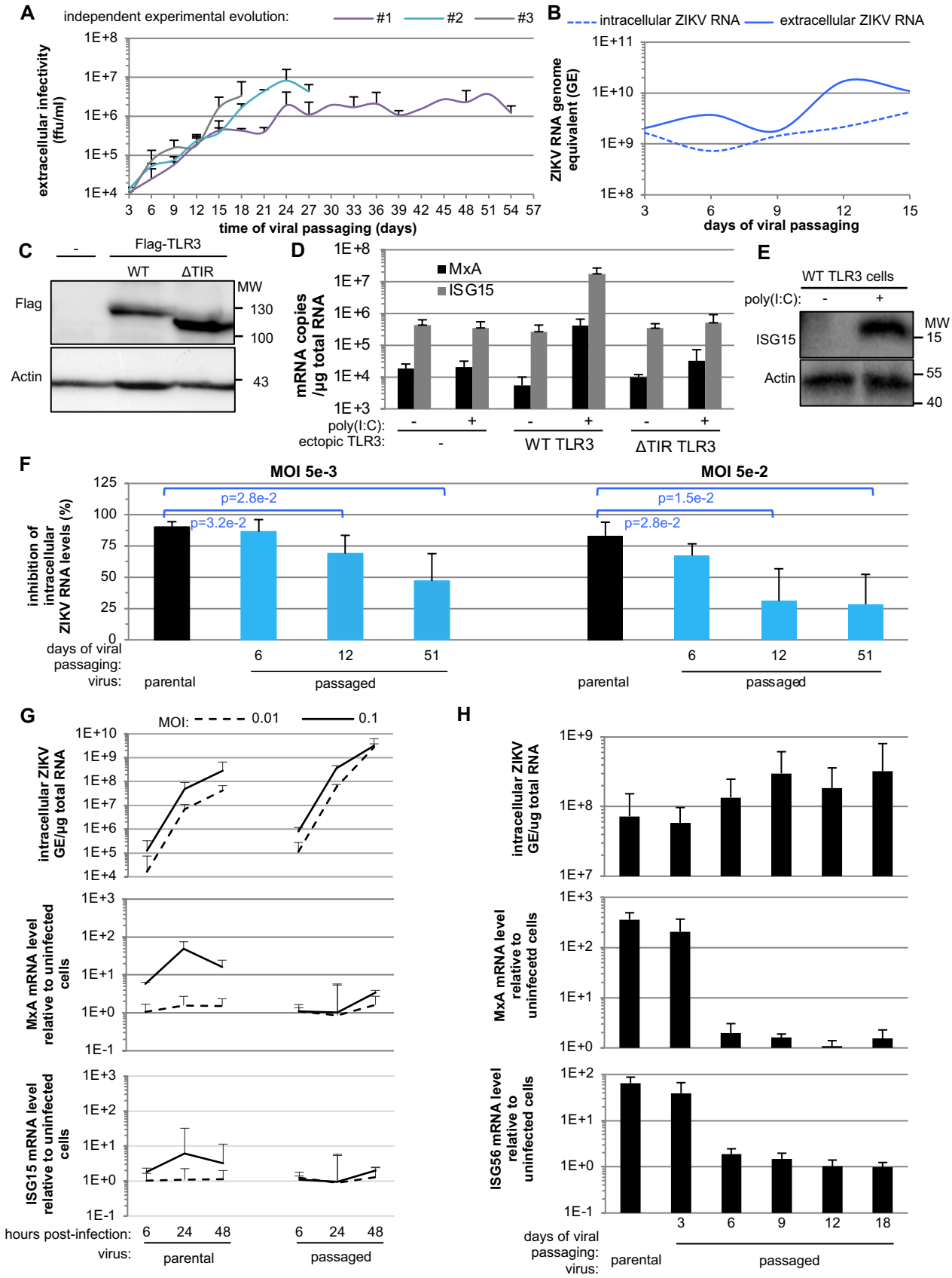
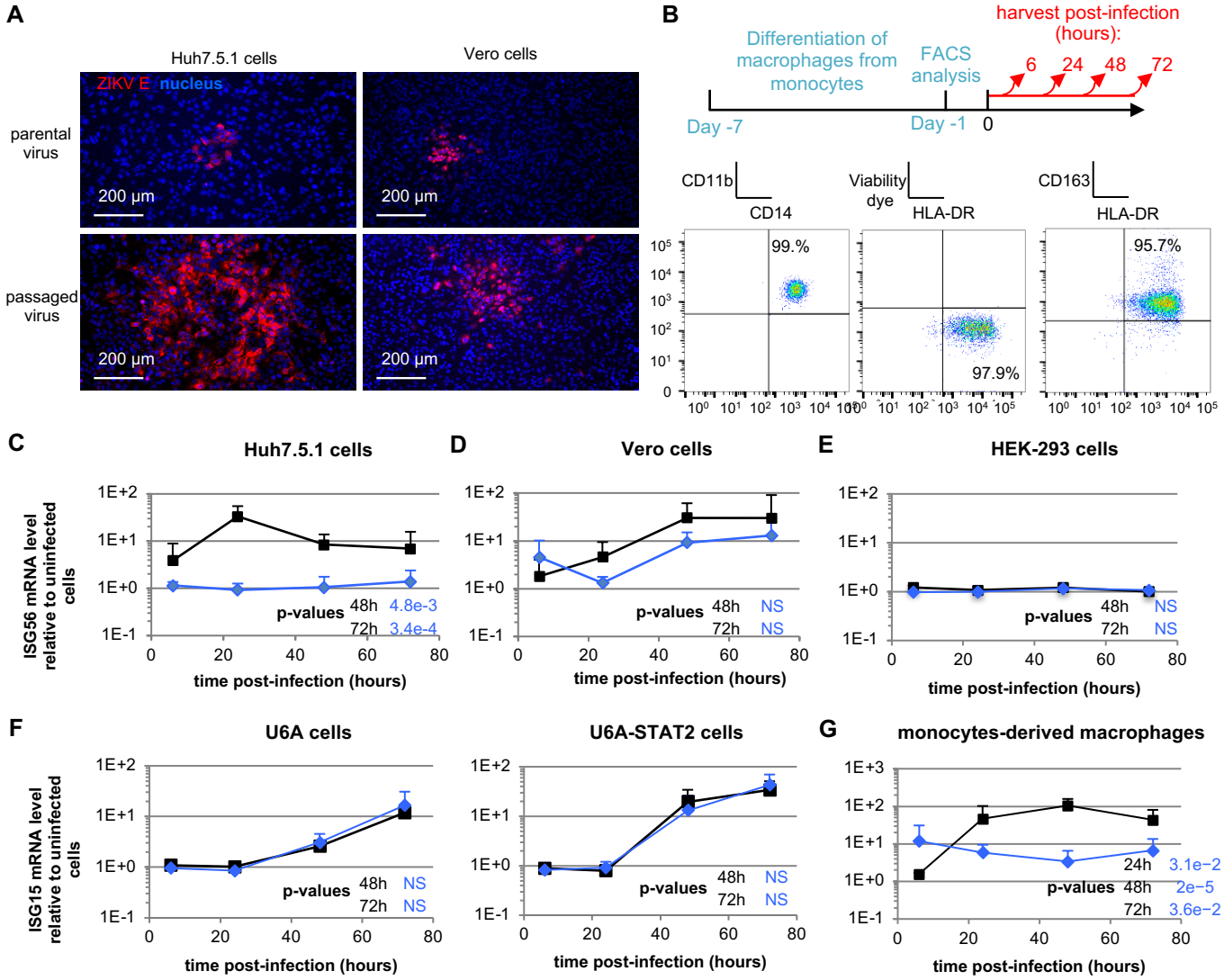


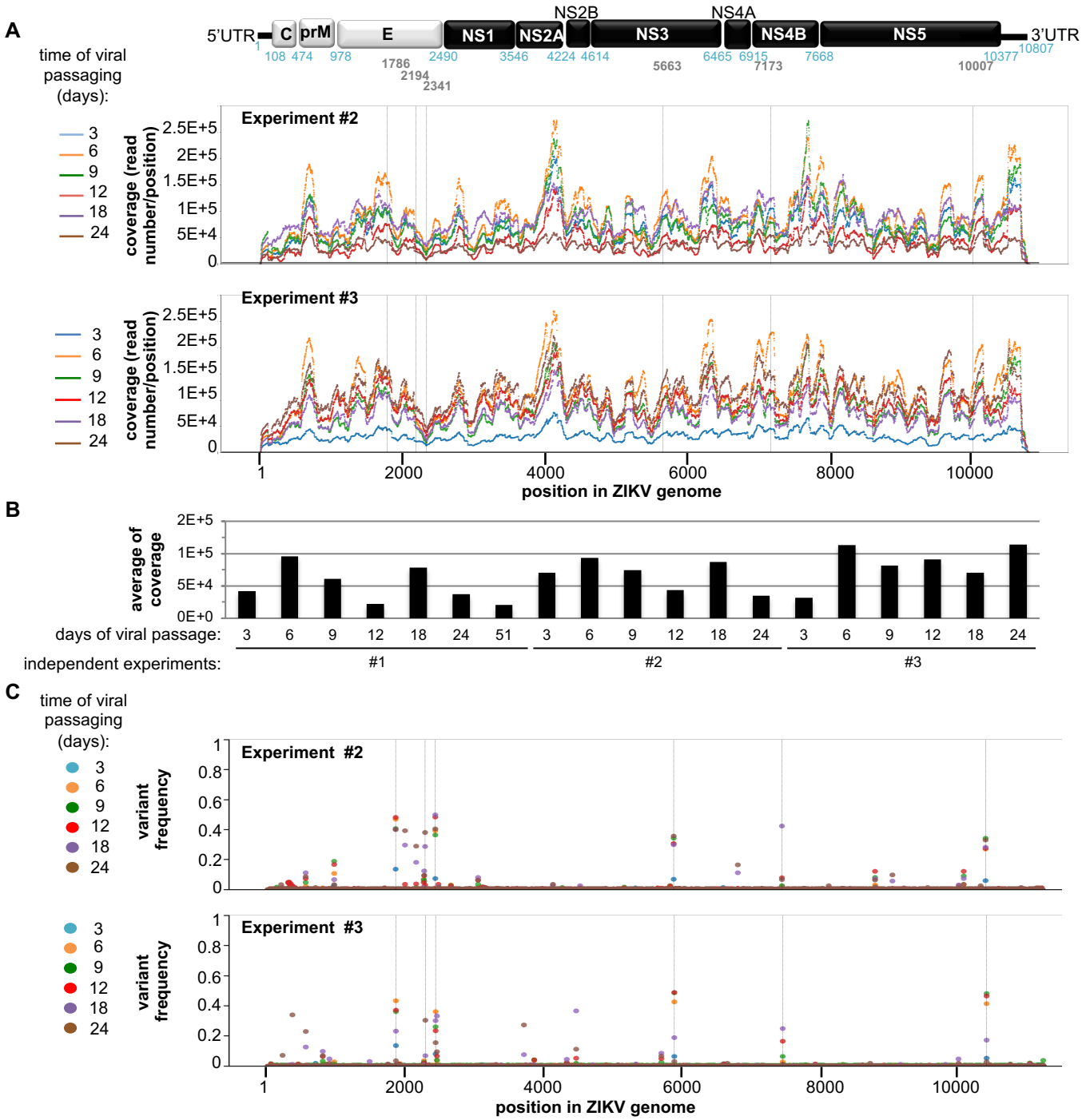
Figure 4. Multiscale model of ZIKV infection and replication. A. Schematic representation of the model. ZIKV infect susceptible cells with rate k_{inf}. Productive viral replication begins with a time delay τ_P for the parental virus and τ_A for the evolved or mutated viruses. Virus replication within infected cells is modeled as logistic growth (**Appendix 1,** mathematical model). Infected cells produce virions with rate k_{vp}, and die with rate k_{icd}. **B**. Model simulation with optimized parameters *versus* experimental measurements of the number of ZIKV genomes per cell upon infection with different doses of parental and adapted ZIKV. **C**. Model simulation *versus* experimental measurements of the number of foci-forming units (ffu) per ml of supernatant upon transfection with parental and the E S455L mutant ZIKV. **D**. Simulation of the ZIKV replication inside the infected cells for parental/ref and adapted/E S455L mutant. The shaded region is the 95% CI for the estimated τ_A value.


A
independent experimental evolution:
#1
#2
#3
extracellular infectivity (ffu/ml)
1E+8
1E+7
1E+6
1E+5
1E+4
3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57
time of viral passaging (days)
B
intracellular ZIKV RNA
extracellular ZIKV RNA
ZIKV RNA genome equivalent (GE)
1E+11
1E+10
1E+9
1E+8
3 6 9 12 15
days of viral passaging
C
-
Flag-TLR3
WT
ΔTIR
MW
Flag
130
100
Actin
43
D
MxA
ISG15
mRNA copies /µg total RNA
1E+8
1E+7
1E+6
1E+5
1E+4
1E+3
poly(I:C):
ectopic TLR3:
-
WT TLR3
ΔTIR TLR3
E
WT TLR3 cells
poly(I:C):
ISG15
Actin
MW
15
55
40
F
MOI 5e-3
MOI 5e-2
inhibition of intracellular ZIKV RNA levels (%)
125
100
75
50
25
0
p=2.8e-2
p=3.2e-2
p=1.5e-2
p=2.8e-2
days of viral passaging:
virus:
parental
passaged
6
12
51
G
MOI:
0.01
0.1
intracellular ZIKV GE/µg total RNA
1E+10
1E+9
1E+8
1E+7
1E+6
1E+5
1E+4
MxA mRNA level relative to uninfected cells
1E+3
1E+2
1E+1
1E+0
1E-1
ISG15 mRNA level relative to uninfected cells
hours post-infection:
6 24 48
virus:
parental
passaged
H
intracellular ZIKV GE/ug total RNA
1E+9
1E+8
1E+7
MxA mRNA level relative to uninfecetd cells
1E+3
1E+2
1E+1
1E+0
ISG56 mRNA level relative to uninfected cells
1E+2
1E+1
1E+0
1E-1
3 6 9 12 18
days of viral passaging:
virus:
parental
passaged

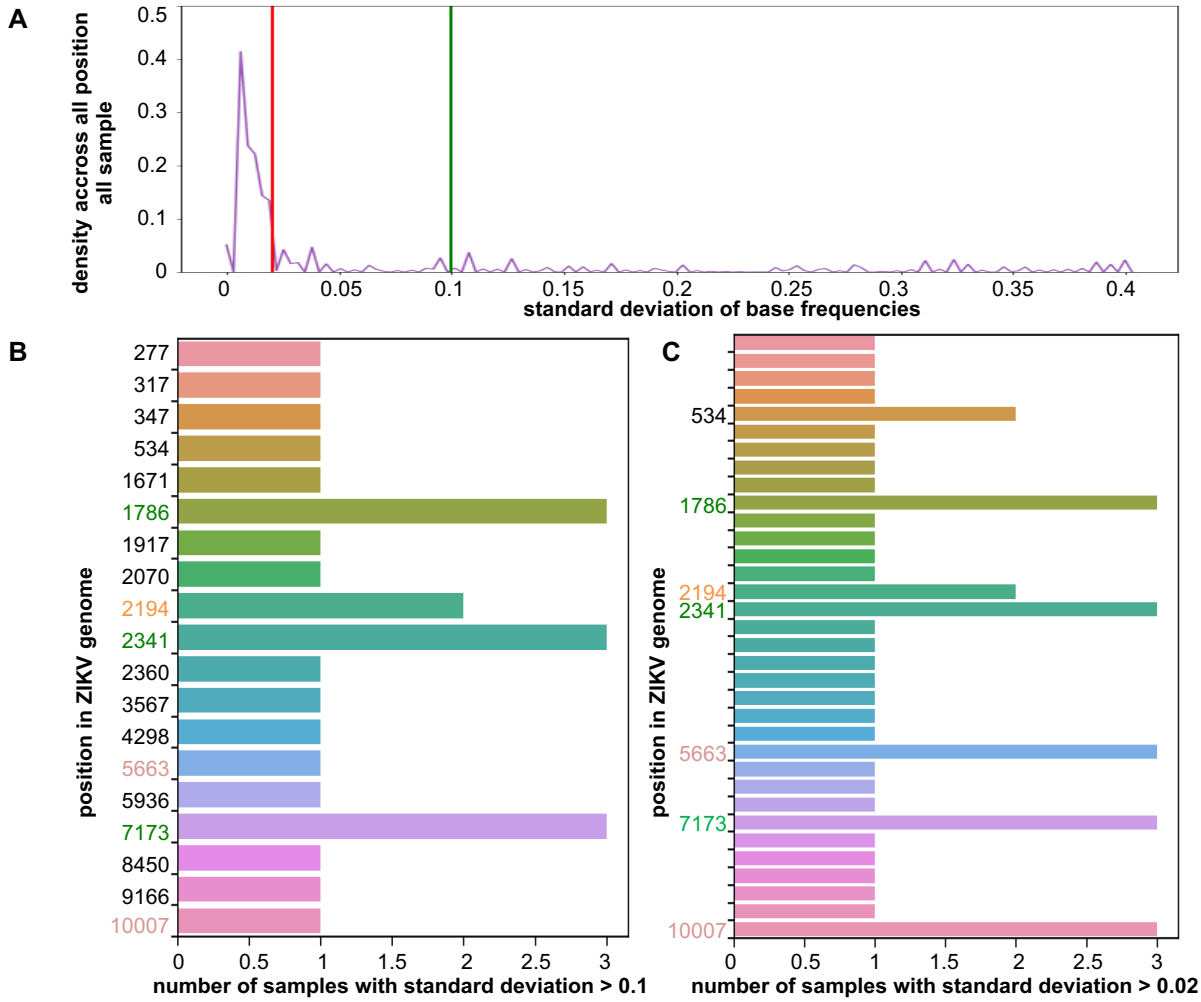
Supplementary Figure 1. Validation of cell study model and analysis of the replication and ISG response upon infection by parental virus and passaged viral populations. A. Quantification of the extracellular infectivity of the viral populations harvested over serially passaging in the course of 3 independent run of experimental evolution (referred to as #1, #2, #3 and stop at days 54, 21 and 18 respectively). Results are expressed as foci forming unit (ffu/ml); determined by 2-to-3 independent measurements for each independent run of experimental evolution; mean ± SD. **B.** Quantification of extracellular and intracellular ZIKV genome (genome equivalent; GE) levels in serial passaging of viral populations by RT-qPCR, results of replicate #1. **C.** Western blot analysis of ectopic WT and ΔTIR TLR3 expression in Huh7.5.1 cells, and as reference the parental Huh7.5.1 cells (-). WT and ΔTIR TLR3 are detected by their Flag-tag; actin detection used as loading control. **D-E**. Analysis of MxA and ISG15 induction upon treatment by poly(I:C) at 100 µg/ml by quantification of MxA and ISG15 mRNA levels by RT-qPCR (**D**) and western blot (**E**). **F.** Inhibition by TLR3-signaling of replication of viral populations harvested at the indicated times of the serial passaging (i.e., day 6, 12 and 51). Results represent the percentage of inhibition in activated-TLR3 cells relative to control cells of intracellular ZIKV genome levels determined at 3 days post-infection at MOI 5e-3 (left panel) and 5e-2 (right panel); 4 independent experiments; mean ± SD; significant differences ($p<0.05$) are indicated by the brackets the top of the graphs. **G.** Quantification of the intracellular ZIKV RNA (upper panel), MxA (middle panel) and ISG15 mRNA (lower panel) levels at the indicated time points post-infection of Huh7.5.1 cells by parental virus and viral populations harvested at day 36 of the viral passaging, as indicated. Infection were performed at MOI 0.1 and 0.01; 3 independent experiments; mean ± SD. **H.** Quantification of the intracellular ZIKV RNA (upper panel), MxA (middle panel) and ISG56 (lower panel) mRNA levels at 24 hours post-infection by viral populations harvested at the indicated times of the serial passaging viral population and parental virus. Infections performed at MOI 1; 3 independent experiments; mean ± SD.



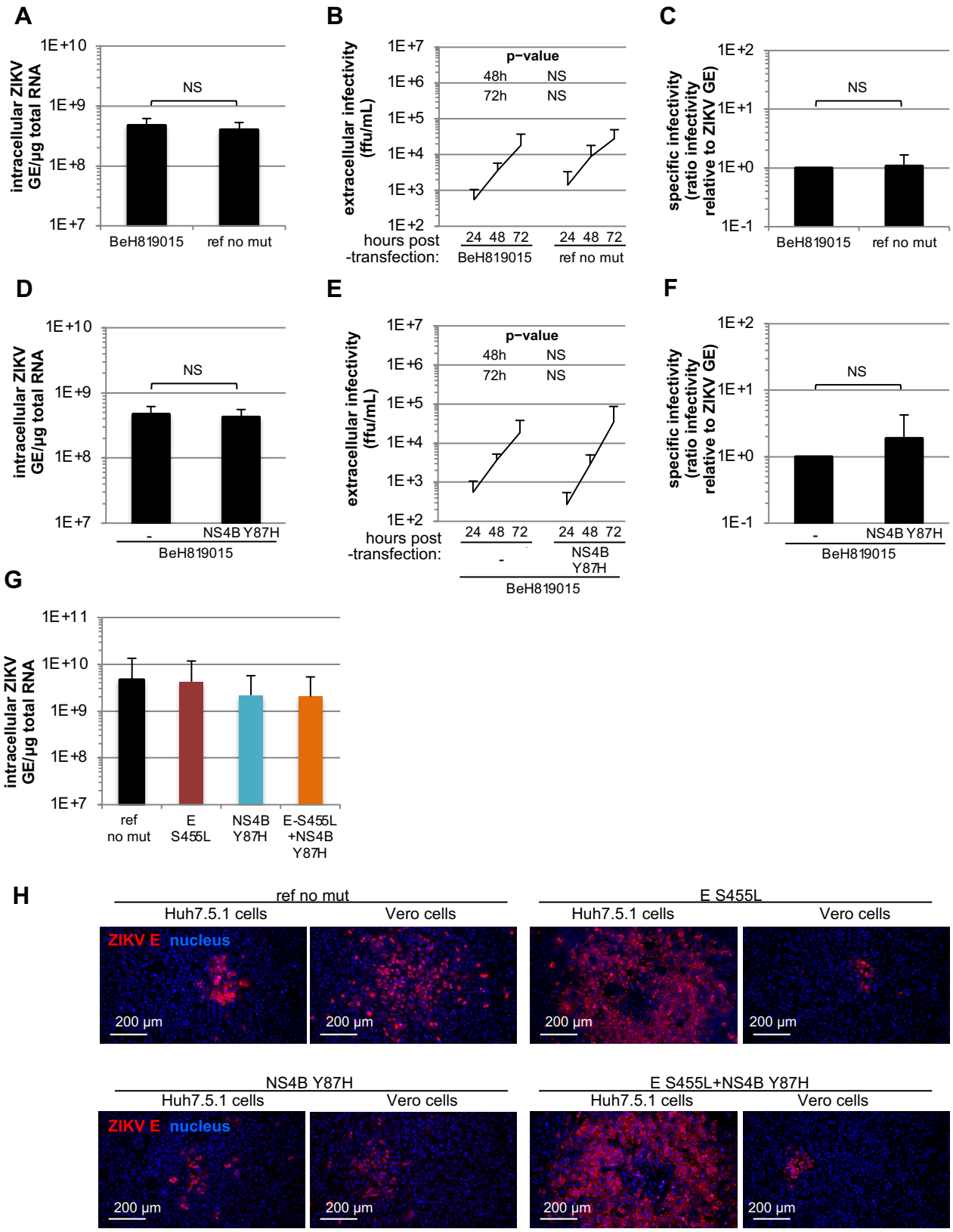
Supplementary Figure 2. Comparison of viral replication of the passaged viral population in various cell lines. A. Representative imaging of infectious foci at 48 hours post-infection by parental virus *versus* serially passaged virus (*i.e.,* day 51 of the viral passaging) in Huh7.5.1 cells (right panels) and Vero cells (left panels) by immunostaining of E enveloppe proteins (red); nucleus stained by Hoechst (blue); scale-bar indicated. **B.** Upper panel, schematic representation of the experimental timeline for macrophages derived from monocytes and infection. Lower panels, their differentiation levels were determined by FACS at the day prior infection and using CD11b, CD14, HLA-DR, CD163 differentiqtion marker and viability markers, as indicated; representative of 4 independent experiments. **C-G.** Kinetic quantification of ISG56 mRNA levels relative to the levels in non-infected cells, at the indicated time post-infection at MOI 0.1 by parental virus versus serially passaged viral population (*i.e.,* day 51 of the viral passaging) in Huh7.5.1 (**C**), Vero cells (**D**), HEK293 cells (**E**), U6A cells and STAT2 expressing U6A cells (**F**), and macrophages derived from monocytes (**G**). 3-to-7 independent experiments; mean ± SD. The p-values of the statistical analysis of the kinetics performed using mixed linear model are indicated on the right side of the graphs, p-value are for the comparison of passaged viral population versus parental virus, and NS; p>0.05.



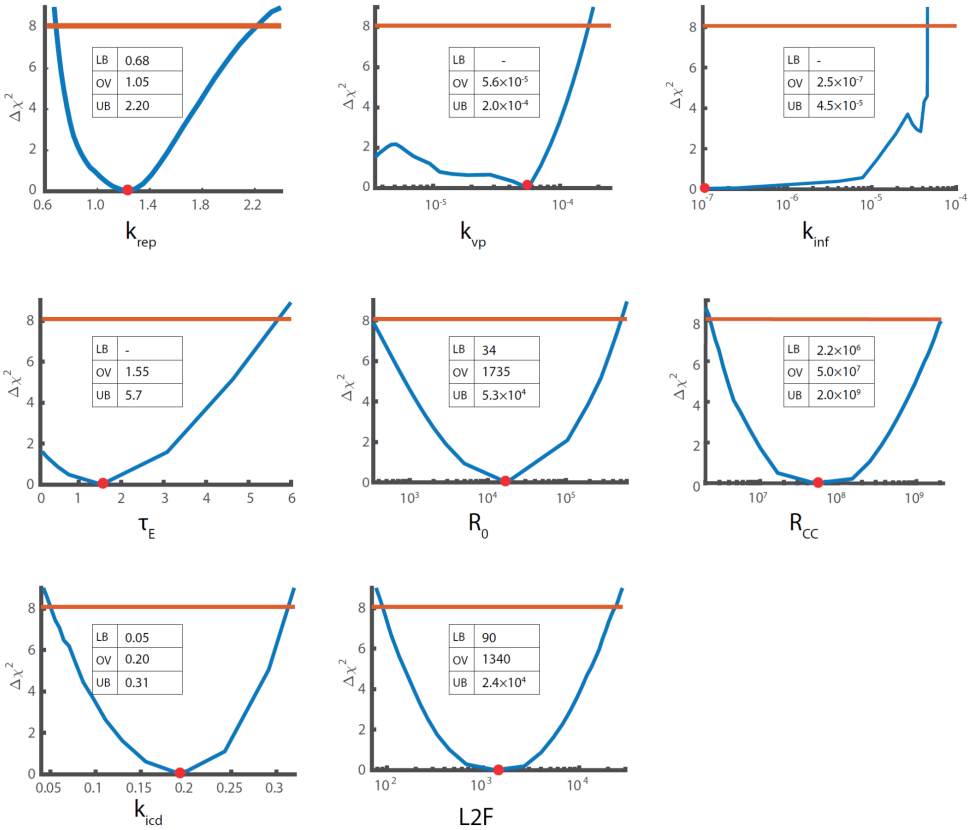
Supplementary Figure 3. Bioinformatic analysis of the genetic evolution of viral populations obtained by next-generation sequencing. A. Coverage of the next-generation sequencing analysis along the ZIKV genome sequence of viral populations harvested at the indicated time points of the serial passaging of independent run of evolution experiment #2 and #3. Schematic representation of ZIKV genome at the top. **B.** Average coverage of the sequenced ZIKV genome for the serially passaged viral population contained in supernatants harvested at the indicated time-points and for independent runs of evolution experiment. **C.** Time-course quantification of the frequency of the second most frequent variants at each position along ZIKV genome in the viral populations harvested in independent runs of evolution experiment #2 and #3. Dotted lines indicate the positions in the viral genome with high standard deviations in several runs of experimental evolution, as defined in **Fig S4A**.



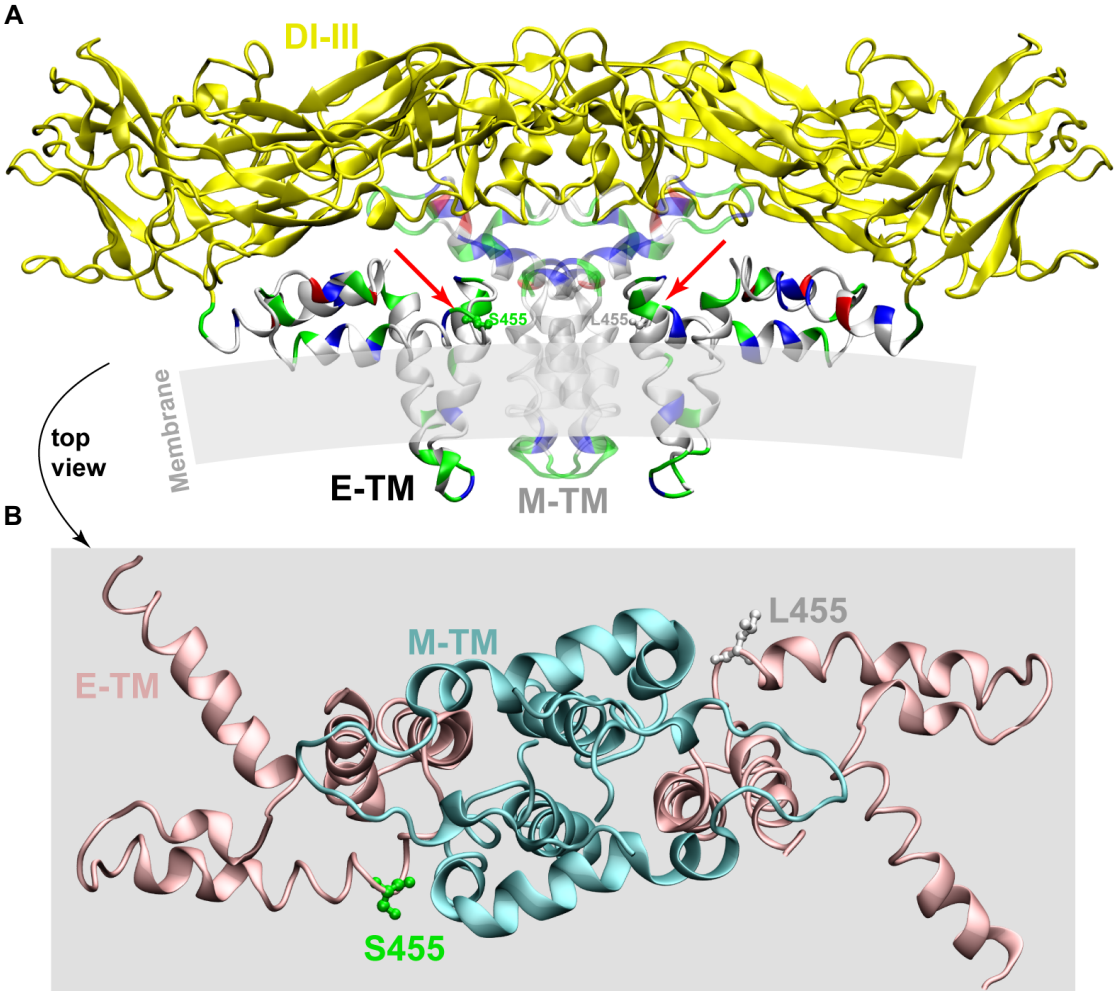
Supplementary Figure 4. Bioinformatics analysis leading to the selection of candidate adaptive mutations. A. Distribution of the standard deviations of variant frequencies across all analyzed samples. Thresholds at 0.02 (red line) and 0.1 (green line) are used for the selection of variants of interest. **B-C.** Representation of the number of samples where a given variant is above the standard deviation of 0.1 (**B**) and 0.02 (**C**) across ZIKV genome.


A
intracellular ZIKV GE/µg total RNA
1E+10
1E+9
1E+8
1E+7
NS
BeH819015
ref no mut
B
extracellular infectivity (ffu/mL)
1E+7
1E+6
1E+5
1E+4
1E+3
1E+2
p-value
48h NS
72h NS
hours post -transfection:
24 48 72
24 48 72
BeH819015
ref no mut
C
specific infectivity (ratio infectivity relative to ZIKV GE)
1E+2
1E+1
1E+0
1E-1
NS
BeH819015
ref no mut
D
intracellular ZIKV GE/µg total RNA
1E+10
1E+9
1E+8
1E+7
NS
-
NS4B Y87H
BeH819015
E
extracellular infectivity (ffu/mL)
1E+7
1E+6
1E+5
1E+4
1E+3
1E+2
p-value
48h NS
72h NS
hours post -transfection:
24 48 72
24 48 72
-
NS4B Y87H
BeH819015
F
specific infectivity (ratio infectivity relative to ZIKV GE)
1E+2
1E+1
1E+0
1E-1
NS
-
NS4B Y87H
BeH819015
G
intracellular ZIKV GE/µg total RNA
1E+11
1E+10
1E+9
1E+8
1E+7
ref no mut
E S455L
NS4B Y87H
E-S455L +NS4B Y87H
H
ref no mut
Huh7.5.1 cells
Vero cells
ZIKV E nucleus
200 µm
E S455L
NS4B Y87H
E S455L+NS4B Y87H

Supplementary Figure 5. Validation of ZIKV molecular clone used for the introduction of the selected non-synonymous mutations in ZIKV molecular clone. A-C. The BeH819015 ZIKV molecular clone (referred to as BeH819015) and derived clones with NS1 mutation, used as reference in **Fig 3** ("ref no mut") were transfected in Huh7.5.1 cells. **A.** Comparable transfection efficiency was assessed by quantification of the intracellular ZIKV GE levels at 6 hours post-transfection; 3 independent experiments; mean ± SD; no statistical difference. **B**. Time-course quantification of infectious viral production at the indicated times post-transfection; 4 independent experiments; mean ± SD; no significant difference as indicated in the table above the graph. **C.** Quantification of the specific infectivity in viral supernatants harvested at 72 hours post-transfection of the indicated molecular clones. Results represent the ratio of the extracellular infectivity relative to extracellular ZIKV RNA levels as in **Fig. 1A**. Results present mean ± SD of 4 independent experiments and set to 1 for BeH819015 for each independent analysis; no significant difference of viral spread overtime as indicated above the graph. **D-F.** The BeH819015 ZIKV molecular clone without mutation (-) versus NS4B Y87H mutation introduce in the same background (referred to as BeH819015) were transfected in Huh7.5.1 cells. Results displayed as in panels **A-C;** 3 independent experiments for panel D and 4 independent experiments for panels **E-F**; mean ± SD; no statistical difference. **G.** ZIKV genome bearing the selected mutations (*i.e.,* single S455L mutation in E, Y87H in NS4B and combined E S455L and NS4B Y87H mutations), and as a reference ZIKV genome without the mutation ("ref no mut"), were transfected in Huh7.5.1 cells. Similar transfection efficiencies were assessed by quantification of the intracellular ZIKV GE levels at 6 hours post-transfection (relative to **Fig. 3A**); 5 independent experiments; mean ± SD; no statistical differences. **H.** Representative imaging of infectious foci at 48 hours post-infection of ZIKV bearing the single S455L mutation in E, Y87H in NS4B and combined E S455L and NS4B Y87H mutations, and as a reference the ZIKV genome without these mutations ("ref no mut") in Huh7.5.1 cells, performed as in **Fig. 3C**; immunostained E surface proteins (red); nucleus stained by Hoechst (blue); scale-bar indicated.



Supplementary Figure 6. Parameter identification. The 95% confidence intervals for the estimated parameter values were calculated using the profile-likelihood method (**Appendix 1**, mathematical model). The blue curve is the maximum-likelihood profile, the solid red line shows the 95% threshold and the red circle represents the estimated parameter value. The estimated parameter values are reported together with their 95% confidence bounds when the parameter is bounded; LB; lower bound, OV; optimal value and UB; upper bound of the respective parameters.



Supplementary Figure 7. Positioning of the S455L mutation in E structure. A. Structure of ZIKA E (in solid) and M (transparent) proteins (5IRE.pdb) (44), with EI-III domains forming the envelope shell (yellow). The E and M protein transmembrane (TM) domains are shown with residue types indicated on the structure (white, hydrophobic; green, polar; bleu, basic; red, acidic) in order to locate membrane insertion and to point out (red arrows) the positions of S455 (left monomer) and the mutant form L455 (right monomer) at the membrane interface. **B.** Zoom of E-TM (cyan) and M-TM (pink), without EI-III, as view from the top of the structure shown in (**A**). The positions of the wild-type S455 and mutant L455 residues at the interface between E and M TM domains are shown on one monomer each (the orientation of the L455 side chain is arbitrary), and has been created using the mutation tool in Swiss-PdbViewer.