Zika viruses encode multiple upstream open reading frames in the 5′ viral region with a role in neurotropism

Charlotte Lefèvre[1], Georgia M. Cook[1], Adam M. Dinan[1,5], Shiho Torii[2], Hazel Stewart[1], George Gibbons[3], Alex S. Nicholson[4], Liliana Echavarría-Consuegra[1], Luke W. Meredith[1], Valeria Lulla[1], Julia C. Kenyon[1], Ian Goodfellow[1], Janet E. Deane[4], Stephen C. Graham[1], Andras Lakatos[3], Louis Lambrechts[2], Ian Brierley[1] and Nerea Irigoyen[1]*.

[1]Division of Virology, Department of Pathology, University of Cambridge, Tennis Court Road, Cambridge, United Kingdom.
[2]Institut Pasteur, Université Paris Cité, CNRS UMR2000, Insect-Virus Interactions Unit, Paris, France.
[3]John van Geest Centre for Brain Repair, University of Cambridge, Cambridge, United Kingdom.
[4]Cambridge Institute for Medical Research, University of Cambridge, Cambridge, United Kingdom.
[5]Current address: Department of Medicine, University of Cambridge, MRC Laboratory of Molecular Biology, Cambridge, United Kingdom.

Running Title: Zika translation in the 5′ untranslated region influences neurotropism.

*Correspondence: ni236@cam.ac.uk

Keywords: Zika virus, upstream open reading frames, ribosome profiling, RNASeq, neurovirulence

Abstract

Zika virus (ZIKV) is an emerging mosquito-borne flavivirus recently associated with congenital diseases and neurological complications. As for all flaviviruses, the ZIKV RNA genome is expected to encode a single polyprotein with all the enzymatic activities required for viral replication. Here, we report the discovery of multiple non-canonical open reading frames (ORFs) identified by ribosome profiling. In both mammalian and insect cells infected with Asian/American and African ZIKV strains, we observed translation of previously unrecognised upstream ORFs (uORFs) in the 5′ region. In the Asian/American ZIKV lineage, ribosomes translated uORF1 and uORF2 that initiated from non-AUG start codons, whereas in the African ZIKV lineage, these two uORFs were fused into a single uORF (African uORF). Using a reverse genetics system, we examined the impact on ZIKV fitness of the expression of single or dual uORFs by analysing a panel of mutant viruses. We found that expression of the African uORF, and more significantly, the Asian/American uORF1, modulated virus growth and tropism in human cortical neurons and 3D organoid tissue, indicating that these novel uORFs contribute to ZIKV neurotropism. Although ZIKV uORFs are expressed in mosquito cells, they did not have a detectable effect on transmission by the mosquito vector *in vivo*. Our discovery of ZIKV uORFs sheds new light on ZIKV-induced neuropathogenesis and raises the question of their existence in other neurotropic flaviviruses.

42 **Introduction**

43 Zika virus (ZIKV) is an emerging *Aedes* mosquito-borne *Flavivirus*, a genus which also
44 includes dengue, yellow fever, Japanese encephalitis, tick-borne encephalitis and West Nile
45 viruses. ZIKV was isolated initially from a febrile monkey in Uganda in 1947[1] but was
46 considered of low importance as most human infections appeared to be asymptomatic.
47 However, it gained prominence from 2007, when the first large epidemic was reported from
48 Yap Island in Micronesia and spread to French Polynesia in 2013[2]. In 2015, the virus reached
49 Brazil and spread across the Americas. Typically, symptoms associated with ZIKV are mild
50 and might include fever, maculopapular rash, conjunctivitis, and myalgia. In adults, the virus
51 can cause severe neurological symptoms, including Guillain-Barré syndrome[2]; and infection
52 of pregnant women can lead to vertical transmission of the virus to cells of the developing brain
53 of the foetus[3]. A key pathological aspect of such infection is impaired development of the
54 unborn brain, leading to gestational abnormalities, including congenital Zika syndrome (CZS).
55 The most well-known feature of CZS is microcephaly[4], but the syndrome also includes general
56 neurological impairment, neurosensory alterations, delays in motor acquisition and an 11-fold
57 greater risk of death during the first three years of life[5,6]. The devastating impact of ZIKV on
58 newborns has far-reaching social consequences, including stigmatisation of affected mothers
59 and babies plus an increase in illegal abortions across Latin America[7]. Recent epidemiolocal
60 studies also indicate that immunity to ZIKV may not persist for as long as previously thought,
61 with neutralizing antibody levels declining over time in adults[8].

62

63 The relative pathogenesis of African and Asian/American ZIKV isolates has been the subject
64 of debate. The American strain is generally considered more pathogenic, although more serious
65 disease manifestations of the African strain may have been missed in the past because,
66 paradoxically, they were more severe (i.e., early abortion instead of birth defects) and thus, less

67 visible[9,10]. Here, we investigate potential pathological correlates of virus gene expression in
68 representative strains of the Asian/American and African ZIKV lineages in mammalian and
69 mosquito cells through RNA sequencing (RNA-Seq) and ribosome profiling (Ribo-Seq). Like
70 other flaviviruses, ZIKV has a positive-sense, single-stranded RNA genome (gRNA) of ~11
71 kb, which contains a single long open reading frame (ORF) flanked by 5′ and 3′ untranslated
72 regions (UTRs) of approximately 100 and 400 nucleotides, respectively. The ORF encodes a
73 large polyprotein that is cleaved by host and viral proteases to yield three structural proteins
74 derived from the N-terminal region (capsid - C, precursor/membrane - pr/M, and envelope - E)
75 and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5)[11]. We show
76 here that in addition to the polyprotein, ZIKV encodes previously unrecognised upstream ORFs
77 (uORFs) in the 5′ UTR, initiating from non-AUG start codons. In African ZIKV isolates, a
78 single uORF (African uORF) is present, whereas, in Asian/American strains, this uORF is split
79 into two uORFs (uORF1 and uORF2) by insertion of a single nucleotide.
80 The impact on ZIKV fitness of the expression of single or dual uORFs is explored by analysing
81 a panel of mutant viruses. We find that expression of the African uORF, and the American
82 uORF1, can modulate virus growth, virulence and tropism in the human brain. We also show
83 that expression of uORF1 is essential for ZIKV neurotropism in the adult and developing brain
84 but has a limited role in the mosquito vector. These experiments reveal novel players in ZIKV
85 pathogenesis and identify new targets for antiviral intervention.
86
87 **Results**
88 **Ribosome profiling reveals the presence of novel uORFs in the 5′ UTR of different ZIKV**
89 **strains.**
90 We utilized Ribo-Seq, in combination with whole transcriptome sequencing (RNA-Seq), to
91 investigate translation of the ZIKV genome at sub-codon resolution. Ribo-Seq exploits the

92 capacity of elongating ribosomes to protect ~30 nucleotides of mRNA from digestion during
93 nuclease incubation of cell extracts[12]. Such ribosome-protected fragments (RPFs) are purified
94 and deep-sequenced, revealing the location of translating ribosomes in a cell at the time of
95 harvesting with single-nucleotide precision[13]. African green monkey (Vero) cells and human
96 glioblastoma-astrocytoma (U251) cells were infected at a multiplicity of infection (MOI) of
97 three with the ZIKV American isolate PE243, representative of the Asian/American lineage[14],
98 and the ZIKV African isolate Dak84 as a model for the original East African strain[15]. To
99 preserve the positions of translating ribosomes upon cell lysis, infected cells at 24 hours post-
100 infection (h p.i.) were flash-frozen or pre-treated for three minutes with the translation inhibitor
101 cycloheximide (CHX) before harvesting. CHX treatment is widely used in Ribo-Seq studies
102 but can lead to the accumulation of 80S ribosomes on start codons and, in stressed cells, can
103 induce the accumulation of RPFs in the 5′ region of coding sequences[16]. For this reason, cells
104 were harvested by flash-freezing to avoid these potential biases (unless stated). Ribo-Seq and
105 RNA-Seq libraries were prepared and deep sequenced as previously described[17].
106
107 At 24 h p.i., the viral envelope (E) protein of each strain is robustly expressed in infected Vero
108 and U251 cells (**Fig 1A**). The Ribo-Seq (red) and RNA-Seq (green) read densities on the virus
109 genome for infected Vero and U251 cells are illustrated in **Fig 1B and Supp Fig 1A-C**.
110 Consistent with the translation of a single polyprotein from the genomic mRNA, read coverage
111 across the main ORF was even, although localized variations in RPF density appear, that may
112 arise from technical biases (ligation, PCR and nuclease biases[18]) or, potentially, from ribosome
113 pausing during translation[19]. Few RPFs were present in the 3′ UTR, consistent with the absence
114 of translation in this region. The 3′ UTR RNA-Seq density was noticeably higher (65-70%)
115 than across the upstream part of the genome in each cell line (**Fig 1B**). Structured flavivirus 3′
116 UTRs resist degradation by the 5′-3′ Xrn1 host exonuclease, giving rise to noncoding

117 subgenomic flavivirus RNAs (sfRNAs) that accumulate during infection[20] and are linked to
118 cytopathic and pathologic effects[14]. A sharp spike of RNA-Seq density was observed in the
119 PE243 (nt 10,478) and Dak84 (nt 10,477) libraries (**Supp Fig. 1D and E**), consistent with the
120 presence of a nuclease-resistant RNA structure at this position. Indeed, this location is two
121 nucleotides upstream of the predicted 5′ end of RNA "stem-loop 2" (SL2)[20].
122
123 In Ribo-Seq samples, the number of negative-sense reads (dark blue) was negligible (<0.005%
124 of mapped reads), indicating that they are unlikely to be genuine RPFs[13,19]. In comparison, in
125 RNA-Seq samples (**Fig 1B and Supp Fig 1A-C**), a low, uniform coverage of negative-sense
126 reads was observed (at ~1% c.f. positive sense), corresponding to negative-sense intermediates
127 that act as templates for genome replication. The translational efficiency (TE) of each virus
128 genome was calculated by applying a 15-nt running mean filter and dividing the number of
129 Ribo-Seq reads by the number of RNA-Seq reads (light blue), revealing relatively even
130 coverage across the genome. Strikingly, we found substantial TE within the 5′ UTR in both
131 CHX-treated and flash-frozen cells (**Fig 1C and 1D**). In PE243-infected cells, a prominent
132 peak of RPF density was seen at nucleotide 25 of the 5′ UTR in Vero and U251 cells, coinciding
133 with a non-canonical (CUG) initiation codon **(Fig 1C and 1D)**. RPFs mapped in the –2 phase
134 along the length of the associated 29-codon upstream ORF (uORF1, in blue), which ends at
135 nucleotide 111 (the 4th nucleotide of the polyprotein ORF). Additionally, RPFs mapped in the
136 –1 phase to a second uORF (uORF2, in orange), which appears to be translated *via* a non-
137 canonical (UUG) initiation codon at nucleotide 80 (**Fig 1C and 1D**). This uORF2 is 77 codons
138 in length and extends 202 nucleotides into the polyprotein ORF. Analysis of the mapping of
139 Ribo-Seq reads to the ZIKV Dak84 5′ UTR (**Fig 1E**) revealed a single uORF (African uORF),
140 translated in the –1 frame. This African uORF initiates at nucleotide 25 (the same initiation
141 codon used by uORF1) and terminates at nucleotide 309, at the stop codon of uORF2 encoding

142 a putative protein of 94 amino acids in length. In all African ZIKV isolates with available
143 sequence data, uORF1 and uORF2 are present as a single ORF, which appears to have been
144 split in two by the insertion of a uracil residue at position 81 in the 1996 Malaysian lineage that
145 gave rise to the Asian/American strain[21].
146
147 To provide further support for the presence of ZIKV uORFs, we examined the phasing of RPFs
148 in viral 5′ UTRs. For RPFs, mapping of the 5′ end positions to coding sequences (CDSs)
149 characteristically reflects the triplet periodicity (herein referred to as 'phasing') of translational
150 decoding[13,17]. We summarized phase relative to the polyprotein (main) ORF by plotting bar
151 charts of the percentage of reads in each phase (**Fig 1F-H**). To avoid the potential confounding
152 effects of overlapping ORFs on these calculations, regions with the least overlap were selected.
153 For uORF1, ribosomal phasing was measured over a 55 nucleotide region (position 25-79),
154 which does not overlap another ORF. Here, a clear dominance of the –2 phase was seen (light
155 blue), consistent with translation in the uORF1 frame (**Fig 1F and 1G**). For uORF2, a short
156 region of 27 nucleotides (position 80-107) was selected, to avoid the high ribosomal occupancy
157 in the 0 phase corresponding to the main ORF that starts at position 108. Here, the majority of
158 reads map to the –1 phase, consistent with uORF2 translation; however, the overlap with
159 uORF1 is still evident from the increased read density in the –2 phase (**Fig 1F and 1G**). For
160 the African uORF (**Fig 1H**), within the chosen region (position 25-106), the majority of reads
161 are attributed to the –1 phase, supporting the expression of this fused uORF. As a positive
162 control, phasing within the polyprotein ORF was assessed, in a region with no known
163 overlapping ORFs (position 310/311-480), revealing a clear dominance of reads attributed to
164 the 0 phase (**Fig 1F-H**). Additionally, as shown in **Supp Fig 2**, the length distribution of Ribo-
165 Seq reads mapping to the ZIKV uORFs mirrored that of polyprotein-mapping RPFs, indicating
166 that they are *bona fide* ribosome footprints.

167 **The translation of ZIKV uORFs can modulate main ORF expression.**

168 The presence of uORFs in mRNAs is often associated with the regulation of downstream gene
169 expression[22,23]. To investigate the potential modulation of ZIKV main ORF expression by the
170 5′ uORFs, T7 RNA polymerase-capped-derived synthetic reporter mRNAs were prepared in
171 which PE243 uORF1, uORF2 or a 5′ portion of the main ORF was placed upstream of, and in
172 frame with, the firefly luciferase (FF-Luc) reporter gene. The ZIKV and FF-Luc sequences are
173 separated by the short, foot and mouth disease virus 2A autoprotease-encoding sequence that
174 liberates the FF-Luc enzyme following expression in cells (**Fig 2A**). RNAs were reverse
175 transfected into Vero cells alongside a T7-derived RNA expressing Renilla luciferase (Ren-
176 Luc) as transfection control. FF-Luc and Ren-Luc activity was measured at 30 hours post-
177 transfection (h p.t.), and translation efficiencies were determined after normalisation with main
178 ORF (wild-type, WT) translation levels (**Fig 2B**, right panel in red). Under these conditions,
179 uORF1 and uORF2 expression levels were 0.80% and 4.13%, respectively (**Fig 2B**). To assess
180 whether translation of the uORFs could affect main ORF translation, mutations were
181 introduced into uORF1 and uORF2 that were predicted to reduce or increase their expression
182 (**Fig 2B-2D**). As shown in **Fig 2B**, mutation of the uORF1 start codon from CUG to CUA
183 (uORF1-KO, blue) led to a modest reduction in uORF1 expression (left panel), no change in
184 uORF2 expression (middle panel) and a small increase in main ORF translation (right panel).
185 Reducing uORF2 expression by changing the initiation codon from UUG to UUA (uORF2-
186 KO, pink), or introducing a premature stop codon within uORF2 at residue number 6 (uORF2-
187 PTC1, orange), reduced uORF2 translation by 50% (**Fig 2B**, middle panel) and led to a slight
188 decrease in main ORF expression (**Fig 2B**, right panel). Replacing the uORF2 initiation codon
189 with a canonical AUG codon (in purple) led to a substantial increase in uORF2 translation and
190 prevented main ORF expression (**Fig 2B**, middle and right panels). A fusion of uORF1 and
191 uORF2 that recapitulated the African uORF (African-like; in green) did not significantly

192 change uORF expression (**Fig 2B**, middle panel) and led to a modest, albeit significant, increase
193 in main ORF translation (**Fig 2B**, right panel). Similar results were obtained in U251 cells
194 transfected with main ORF-2A-FFLuc mutants (**Supp Fig 3A**).
195
196 We went on to ask whether viral infection could influence the relative utilisation of main and
197 uORFs. In these experiments, reporter mRNA-transfected cells were infected at 6 h p.t with
198 PE243 (MOI:3) and harvested 24 h later. In the context of infection, the expression of the
199 uORFs and main ORF relative to each other remained similar, although the total expression of
200 each ORF significantly increased compared to uninfected cells (**Fig 2E**). Notably, the raw
201 luciferase values (**Supp Table 1**) were slightly lower in the presence of the virus, which may
202 reflect some impairment of translation initiation as a result of the phosphorylation of the alpha
203 subunit of the initiation factor 2 (p-eIF2α) during infection (**Supp Fig. 3B**). The effect of uORF
204 mutations on main ORF expression in infected cells was also tested (**Fig 2F**). In all cases,
205 except uORF2-KO, the relative expression of the main ORF was increased modestly. In
206 conclusion, virus infection modestly and uniformly increased expression from upstream and
207 main ORFs.
208
209 **uORF translation modulates virus replication**
210 To investigate the potential role of ZIKV uORFs in virus replication, a panel of viruses
211 containing the mutations tested above was generated using reverse genetics of an American
212 ZIKV infectious clone[24], detailed in **Fig 3A**. Given that structured RNA elements and long-
213 range interactions in the 5′ and 3′ terminal regions of the ZIKV genome are essential for virus
214 translation and replication[25,26], we began by confirming that the 5′ end structures were retained
215 in full-length RNA transcripts of the mutant viruses. Using selective 2′-hydroxyl acylation
216 analysed by primer extension (SHAPE), we found that the structure of the 5′ UTR and the start

217 of the main ORF of the mutant viruses generally very closely matched that of the American
218 WT infectious clone (**Fig 3B and Supp Fig 4A**). Two differences were observed; modelling
219 of the African-like mutant virus indicated a slightly shorter 3rd stem-loop (cHP), with loss of
220 two base pairs at the bottom of the helix (**Fig 3B**), and in the uORF2-AUG mutant, the internal
221 loop in the centre of the 2nd stem-loop (SLB, **Supp Fig 4A**) is base-paired. Mutant viruses were
222 also analysed for the stability of the introduced mutations. After five passages, RT-PCR
223 analysis of intracellular viral RNA (i.e., initially infected with MOI 0.01 PFU/cell) revealed
224 that all mutations were stable, with two exceptions. ZIKV uORF2-KO reverted to WT after
225 passage 1, and we could not recover any virus following electroporation of the ZIKV uORF2-
226 AUG mutant (**Supp Fig 4B**). These observations suggest that the translation efficiency of
227 uORF2 might be critical to ZIKV replication.
228
229 To assess the growth and infectivity of the stable mutants, U251 cells were infected with
230 sequence-verified American WT, African-like, uORF1-KO or uORF2-PTC1 viruses at low
231 multiplicity (0.01 PFU/cell) in a multi-step growth experiment from 0 to 96 h p.i. As shown in
232 **Fig 3C**, from 48 h p.i. onwards, African-like and uORF1-KO mutant viruses reached
233 significantly higher titres (~6-fold) than the American WT, whereas no difference was
234 observed with the uORF2-PTC1 mutant. This phenotype was confirmed in competition assays
235 in which cells were simultaneously infected with a set ratio (50:50 and 90:10) of American WT
236 virus to the corresponding mutant virus to a final MOI of 0.01 PFU/cell (**Fig 3D**). Note that the
237 uORF2-PTC1:American WT comparison was 90:10 instead. Supernatants containing virus
238 particles were harvested 72 h later and used to infect fresh cells at a dilution of 1:200. This
239 regimen was repeated six times[27], and intracellular RNA was harvested, reverse-transcribed,
240 and Sanger sequenced. Passage 0 corresponds to 16 h post-viral RNA input and represents the
241 starting point. From passage 2, the African-like and uORF1-KO mutant viruses began to

242 dominate the cultures (**Fig 3D**), indicating that both mutant viruses can outcompete American
243 WT. In contrast, the proportion of uORF2-PTC1 viral RNA was relatively unchanged
244 throughout the course of the experiment (**Fig 3D, right panel**), indicating no competition with
245 the American WT virus. The phenotypic differences amongst viruses were also evident, albeit
246 subtler, in Vero-infected cells (**Supp Fig 4C and Supp Fig 4D**).
247
248 In an effort to detect expression of uORF products in infected cells, antibodies to expressed
249 proteins and peptides were prepared, but they lacked sensitivity and specificity. Therefore, we
250 returned to ribosome profiling to examine translation of the uORFs in ZIKV mutant viruses
251 (**Supp Fig 4E and Supp Fig 4F**). As seen in **Fig 3E**, reads corresponding to translation of
252 uORF1 (–2 frame) and uORF2 (–1 frame) were visible in the American WT virus, similar to
253 PE243 (**Fig 1C** and **Fig 1D**). The uORF1-KO mutant, as expected, blocked translation of
254 uORF1, with no reads observed at the initiation codon (blue asterisk); and the African-like
255 virus with fused uORFs recapitulated the pattern of uORF translation of the African virus (as
256 in **Fig 1E**). These observations were further confirmed in phasing bar charts, with the highest
257 proportion of reads in the predicted reading frames (**Fig 3E, right panels**). Unexpectedly, reads
258 corresponding to the uORF2 phase were present after the premature termination codon in
259 uORF2-PTC1 (**Fig 3E,** red asterisk), supported by the high proportion of reads corresponding
260 to the –1 phase in phasing plots (**Fig 3E, right panels**). Thus, the translation of uORF2 might
261 also be achieved through an additional mechanism.
262
263 **ZIKV uORF2 gets translated by alternative mechanisms**
264 As described earlier (**Fig 2B, middle panel**), knockout of the uORF2 initiation codon or
265 introduction of a premature termination codon (uORF2-PTC1) depressed uORF2 translation
266 only two-fold in comparison to the American WT. In a further attempt to block the synthesis

267 of the uORF2 product, a premature termination codon was placed within the middle of the
268 uORF2 sequence (uORF2-PTC34, **Fig 3F** in bright green) in the context of the uORF2-2A-
269 FFLuc mRNA, and luciferase assays performed in transfected cells. As shown in **Fig 3G**,
270 introduction of PTC34 reduced uORF2 translation ~10-fold, suggesting that a signal for
271 alternative uORF2 expression is present between residue 6 (PTC1) and residue 34 (PTC34).
272 To investigate this, we looked for potential initiation sites between codons 6 and 34. No AUG
273 codons are present, but four potential sites of non-AUG translation initiation could be accessed
274 through leaky scanning of 40S initiation complexes on the mRNA. These amino acids were
275 sequentially mutated (Mut1-AUC, Mut2-AGG, Mut3-AUU and Mut4-ACG; **Supp Fig 4G**)
276 within the uORF2-2A-FFLuc reporter mRNA and tested in transfected cells (**Supp Fig 4H**).
277 In all cases, luciferase levels were unchanged, arguing against a role for leaky scanning in
278 uORF2 expression. We noticed that the sequence immediately downstream of the main ORF
279 AUG has two A-rich stretches upstream of a stem-loop (cHP) that could potentially act as
280 signals for programmed ribosomal frameshifting (PRF[28]), diverting a proportion of ribosomes
281 into the overlapping –1 (uORF2) reading frame (denoted FS2 and FS3 in **Fig 3F**). Reducing
282 the homopolymeric nature of each of these stretches by the introduction of two G residues
283 within the uORF2-2A-FFLuc reporter reduced uORF2 translation at the FS3 site by 50%
284 suggesting that at least half of uORF2 expression is derived from ribosomes that undergo –1
285 PRF very early during main ORF translation **(Fig 3G)**. Disruption of the main ORF initiation
286 codon (FS1) also significantly reduced uORF2 expression, consistent with a requirement for
287 translation of the main ORF in uORF2 expression and frameshifting into the –1 frame at the
288 'blue' slippery sequence (FS3 region). Furthermore, within the context of the uORF2-PTC1-
289 2A-FFLuc reporter mRNA, both FS1 and FS3 mutations reduced uORF2 expression further,
290 presumably as ribosomes can no longer access uORF2 following initiation at the main ORF
291 AUG codon (FS1) and frameshifting into uORF2 at the FS3 signal (**Fig 3H**).

292 **ZIKV uORF1-encoded protein helps in the formation of the cytoskeletal cage during**
293 **infection**

294 Bioinformatic analyses suggest no homology of ZIKV uORFs to known proteins. Given their
295 predicted sizes (uORF1, 3.1 kDa; uORF2, 8.4 kDa; and African uORF, 10.3 kDa), however, it
296 is likely that some, or all, will have intrinsic biological activity. To examine their cellular
297 localization, we transiently expressed mCherry-tagged and TAP (Strep-Strep-FLAG)-/FLAG-
298 tagged uORF variants in uninfected and infected mammalian cells, with PE243 and Dak84,
299 and performed subcellular fractionation (**Fig 4A and Supp Fig 5**). As observed in **Supp Fig**
300 **5C and Supp Fig 5D**, uORF2 and African uORF were mainly present in the cytoplasm, and
301 there was no relocalization upon infection. However, uORF1, tagged with mCherry or TAP,
302 and in two different cell lines (Vero and U251 cells), appeared not only in the cytoplasm but
303 also in the cytoskeletal fraction (**Fig 4A and Supp Fig 5E-5H**). Note that uORF1-TAP also
304 appeared in the nuclear fraction, but this was probably due to passive diffusion through the
305 nuclear membrane of a small protein. To discern the specific cytoskeletal target of uORF1, we
306 performed confocal analysis with different proteins marking the three types of cytoskeletal
307 filaments: actin for microfilaments, tubulin for microtubules and vimentin for intermediate
308 filaments (**Fig 4B**). It was found that uORF1-TAP could form denser granular structures that
309 co-aligned with an abnormal 'collapsed' vimentin (**Fig 4B**, **orange square**), and this was not
310 observed in mock-transfected cells. ZIKV infection reorganises microtubules and intermediate
311 filaments to form a cytoskeletal cage surrounding viral factories[29,30]. These factories are the
312 sites of viral RNA replication and virion assembly, and cytoskeletal remodelling might partly
313 contribute to localizing these processes in a closed environment to prevent access to innate
314 immune defence components. To test specificity, the cytoskeletal phenotype of the uORF1-KO
315 mutant was compared to the American WT. Previously synchronised Vero cells in the G0/G1
316 phase were infected (at MOI:3) with either virus, the cells fixed at 18 and 24 h p.i. and stained

317 for viral E protein and vimentin (**Fig 4C**). At 18 h p.i., vimentin was perinuclear in all cells,
318 although a denser arrangement was observed in infected cells (**Fig 4C**, upper panel). At 24 h
319 p.i., the cytoskeletal cage, as defined by the collapse of vimentin, was visible in American WT
320 infected cells (**Fig 4C, lower panel orange squares**) but vimentin remained in a perinuclear
321 distribution in uORF1-KO infected cells. We quantified this phenotype by scoring infected
322 cells in four categories; no change in vimentin reorganization (1), increased density of vimentin
323 in the perinuclear region (2), evidence of vimentin collapse (3), and complete collapse of
324 vimentin (4) (see **Supp. Fig 6**). The data derived from scoring between 110-170 cells per
325 condition in two independent experiments are shown in **Fig 4D**. These data indicate that the
326 expression of the uORF1 polypeptide likely helps in the formation of the cytoskeletal cage in
327 infected cells.
328
329 **ZIKV uORFs are involved in neurotropism in human cortical neurons and cerebral**
330 **organoids.**
331 To assess whether expression of the different uORFs might influence the neurotropism and
332 neurovirulence of ZIKV, 2D human induced pluripotent stem cell (hiPSC)-derived cortical
333 neuronal mono-cultures and human 3D cortical organoid slice cultures were infected with
334 either American WT, uORF1-KO or the African-like ZIKV.
335 Differentiated hiPSC-derived glutamatergic cortical neurons (i^3Neurons[31]) were infected at
336 high MOI (10 PFU/cell based on titre in Vero cells) for 4 days. Released virions in the
337 supernatant were quantified and from 48 h p.i. onwards, mutant virus titres were significantly
338 lower, decreasing by 1.5- to 2-log10 at 96 h p.i. (**Fig 5A**), in sharp contrast to their increased
339 replication in U251 cells (**Fig 3C**). This was further confirmed by immunolabelling for ZIKV
340 E protein and confocal microscopy (**Fig 5B** and **Supp Fig 7**), where spread of infection was

341 less obvious with the African-like virus and infection with the uORF1-KO mutant was limited
342 to single infected cells, suggesting little to no viral spread.
343 Next, we used cortical organoids to test the extent of neurotropism in a human brain-like 3D
344 tissue environment and to explore whether uORF1 is also required for the infectivity of other
345 cell-types. To do so, we infected cortical organoid slices grown at the air-liquid interface (ALI-
346 COs), which recapitulate cortical cell type-diversity, layering and neurodevelopmental
347 milestones[32,33]. At 82 days in vitro (DIV), reflecting the first trimester of gestation, ALI-COs
348 were infected with the American WT, the uORF1-KO and the African-like viruses at MOI 5.
349 Virus inoculum was removed after 24 h and the ALI-COs were grown for a further 7 days
350 before being fixed. 6-µm tissue sections were stained for viral E protein and positive cells were
351 quantified in relation to the total number of nuclei. Approximately, 50% of cells were positive
352 for the E protein in the American WT virus infection experiment. This was significantly
353 reduced in the African-like and uORF1-KO virus infection experiments, to ~35% and ~7.5%,
354 respectively (**Fig 5C, Supp Table 2**). This was further confirmed by IF (**Fig 5D-5F, Supp Fig**
355 **7**), corroborating our findings on the involvement of the ZIKV 5′ UTR in neurovirulence.
356 To examine the differential neural cell tropism of these mutant viruses, cryostat sections of
357 ALI-COs were immunostained for a panel of cellular markers, including nestin (as a marker
358 for neural progenitor cells, NPCs), GFAP (a marker for the astroglial lineage such as radial
359 glial cells, glial progenitors and astrocytes), and MAP2 (for mature neurons). As observed in
360 **Fig 5D-G**, ZIKV viruses preferentially infected nestin- or GFAP-positive cells, whereas very
361 few MAP2-positive cells were also positive for E protein. In general, a similar infectivity
362 pattern was seen with all viruses tested (~40% nestin^{+}, ~31% GFAP^{+} and ~1.5% MAP2^{+}, **Supp**
363 **Table 3**), suggesting that the modifications to the ZIKV 5′ UTR do not affect neural cell
364 tropism.
365

366 Together, these data indicate that the expression of ZIKV uORF1 is an essential neurotrophic
367 factor required for the infection of cortical neurons and NPCs in human tissue. These results
368 also demonstrate that the African 5′ UTR arrangement impairs the infection of NPCs and
369 precursors of the astroglial lineage in the developing brain, although to a lesser extent.
370
371 **No detectable effect of ZIKV uORFs in the mosquito vector.**
372 ZIKV typically cycles between humans and *Aedes* mosquitoes; hence we also tested whether
373 the uORFs were expressed during the infection of cells derived from the mosquito vector. Ribo-
374 Seq analysis of *Aedes albopictus* (C6/36) cells infected with PE243 for 24 h suggested that
375 both uORFs were occupied by ribosomes during infection (**Fig 6A**). We compared the
376 transmission dynamics of American WT, African-like and uORF1-KO viruses in female *Aedes*
377 *aegypti*. Mosquitoes were exposed to an artificial, infectious blood meal containing an expected
378 virus titre of 2 x 10^6 PFU/mL of blood. Actual titres ranged from 0.8 x 10^6 to 3.6 x 10^6 PFU/mL
379 across the different experiments, and this uncontrolled variation was accounted for in the
380 statistical analysis. At several time points after the infectious blood meal, we determined the
381 rates of mosquito infection and systemic viral dissemination by RT-PCR and detected the
382 presence of ZIKV in saliva by infectious assay. We calculated infection prevalence as the
383 proportion of blood-fed mosquitoes with a body infection (determined by RT-PCR, **Fig 6B left**
384 **panel**), dissemination prevalence as the proportion of infected mosquitoes with a virus-positive
385 head (determined by RT-PCR, **Fig 6B middle panel**) and transmission prevalence as the
386 proportion of mosquitoes with a disseminated infection that had infectious saliva (determined
387 by focus-forming assay, **Fig 6B right panel**), as previously described[34,35]. A total of 99, 128
388 and 234 individual mosquitoes were analysed in the first, second and third experiments,
389 respectively.

390 Most mosquitoes became infected across experiments, different viruses and time points (mean:
391 90.5%; median: 90.9%; range: 66.7%–100%), and none of these variables had a statistically
392 significant effect on infection prevalence (**Supp Table 4**). Both dissemination and transmission
393 prevalence significantly increased over time (**Fig. 6B-C**), but there was no detectable
394 difference between the different ZIKV mutant viruses (**Supp Table 4**). We conclude that the
395 different ZIKV mutant viruses have similar transmission dynamics in mosquitoes regardless of
396 their 5′-UTR arrangement.
397

398 **Discussion**

399 In this study, we provide the first high-resolution maps of ZIKV translation in both mammalian
400 and mosquito cells and describe two previously overlooked uORFs (uORF1 and uORF2) in
401 contemporary ZIKV strains, which exist as a single uORF in African isolates. By generating
402 mutant viruses, we show that the uORFs can modulate virus growth, virulence and tropism in
403 human brain cells. How the uORFs exert their effects is uncertain, but is likely through
404 modulation of expression of the main polyprotein ORF and *via* properties of their encoded
405 proteins.
406

407 The translational activity of the ZIKV uORFs is noticeably lower than that of the main
408 polyprotein ORF, consistent with their initiation at non-AUG codons. Whilst the effect of
409 flanking bases on initiator AUG codons is well established (Kozak context), it is less well
410 understood for non-AUG codons, making the strength of non-AUG initiation codons harder to
411 predict[36]. Based on our present understanding, the ZIKV uORFs rank as modest start
412 codons[36,37]. However, their utilisation may also be modulated by RNA structure and,
413 potentially, by virus and host proteins. Recent toeprinting analysis of ZIKV initiation *in vitro*
414 has also detected the utilisation of non-AUG initiation sites at nucleotide positions UUG_{41},

415 UUG_{80} and UUG_{86} (c.f. CUG_{25} and UUG_{80} here). The magnitude of the toeprints observed was
416 influenced by the availability of particular translation factors in the reconstituted assay system
417 and the RNA circularisation status of the mRNAs used[26]. The absence of a toeprint at CUG_{25}
418 in the latter study might reflect the different experimental milieu (*in vitro* reconstitution versus
419 infected cells).
420
421 In luciferase reporter gene assays, expression of uORF1 had an inhibitory effect on main ORF
422 translation, whereas expression of uORF2 or the African uORF enhanced translation of the
423 reporter gene. This might explain why mutant viruses bearing these changes (uORF1-KO and
424 African-like) in the 5′ UTR grew to higher titres and were fitter, in comparison to American
425 WT, in tissue culture cells. The inhibitory effect of the uORF1 start codon mutation might be
426 related to the fact that the uORF1 stop codon (UGA) overlaps with the initiation codon of the
427 main ORF (AUGA). We speculate that ribosomes terminating at the stop codon may decrease
428 access of ribosomes to the main ORF AUG, which would be obviated in uORF1-KO, as fewer
429 ribosomes are delivered to the stop codon. In a similar vein, ribosomes translating uORF2 or
430 the African uORF would unwind SLB without terminating within it, increasing access of
431 ribosomes to the main ORF AUG. Viruses with mutations intended to increase (uORF2-AUG)
432 or decrease (uORF2-KO) translation of uORF2 were non-viable or reverted to the wild-type
433 sequence. The dramatic effect of the uORF2-AUG mutation in the virus context could be
434 mediated in several ways. Upregulation of uORF2 expression may drastically reduce the
435 number of scanning ribosomes reaching the main ORF AUG and diminish polyprotein
436 expression, an idea consistent with the significant reduction of main ORF expression in reporter
437 gene assays (**Fig. 2B**). Alternatively, overexpression of the uORF2 protein product in infected
438 cells may be detrimental to virus growth. We also note from SHAPE analysis that this RNA
439 adopts a slightly different fold in the SLB region, which could also contribute to improper

440 replication and translation. The uORF2-KO virus revertant phenotype might be attributable to
441 a reduced amount of uORF2 protein product in infected cells, or altered translational dynamics
442 in the complex 5′ UTR of ZIKV.
443
444 Our results, and those of others[38,39], reveal that eIF2 becomes phosphorylated on its α-subunit
445 upon ZIKV infection (**Supp Fig. 3B**). eIF2α phosphorylation leads to a general shutdown of
446 protein synthesis by inhibiting translation initiation, but in some cases, such as the translation
447 of the activation transcription factor 4 (ATF4[40]), uORFs act as enhancers of main ORF
448 translation as a response to cellular stress. The finding that the ZIKV ORFs are translationally
449 enhanced during infection suggests that the ZIKV 5′ arrangement, and the likely interplay of
450 the two uORFs and main ORF, may offer some resistance to the translation inhibition
451 engendered by phosphorylated eIF2α. This is most evident with the African-like arrangement
452 (**Fig. 2F**) and might explain why the African-like virus outgrows American WT in tissue
453 culture (**Fig. 3D**).
454
455 An unexpected finding is the dual mode of uORF2 expression, via initiation from the UUG
456 initiation codon, and also expression of an N-terminally truncated version by –1 ribosomal
457 frameshifting of ribosomes initiating at the main polyprotein (0) frame. Based on previous
458 work[28], a likely site of frameshifting in this region is the C_CCA_AAA putative "slippery"
459 sequence located four codons into the polyprotein and three nucleotides upstream of the cHP
460 stem-loop. Consistent with this hypothesis, mutation of this sequence to non-slippery
461 C_CCG_AAG reduced uORF2 expression in reporter gene assays (**Fig. 3H**). Frameshifting to
462 CCC_AAA (uORF2 frame) generates an N-terminal variant of uORF2, in which five amino
463 acids derived from the start of the polyprotein replace the 14 N-terminal residues from initiation
464 at the UUG codon. Conceivably, 'non-canonical' expression of uORF2 could also arise as a

465 result of transcriptional slippage of T7 RNA-polymerase on the A-rich tract in this region, and
466 indeed, transcriptional slippage has been reported to account for the proposed frameshifting
467 that takes place at a similar position in the hepatitis C virus (HCV) genome[41]. However,
468 analysis of the ZIKV RNA-Seq data does not reveal any heterogeneity in this A-rich region, in
469 contrast to what was observed with HCV[41].
470
471 Bioinformatics analyses do not reveal any apparent homology of ZIKV uORFs to known
472 proteins, although AlphaFold 2 analysis[42] suggests that uORF1 adopts an alpha-helical
473 conformation, whereas uORF2 is likely to have no well-defined secondary structure (**Supp Fig.**
474 **8A**). Related uORFs can be identified in other neurotropic flaviviruses, such as Japanese
475 encephalitis virus (JEV)[43] and West Nile virus (WNV) (**Supp Fig. 8B**), and they are also
476 predicted to have adopted alpha-helical conformations. The expression in *trans* of uORF2 and
477 the African uORF resulted in a cytoplasmic localization that was not changed upon infection.
478 However, transfected uORF1 was predominantly cytoskeletal and led to the collapse of the
479 vimentin network, with no change observed in microfilaments or microtubules. Previous
480 electron tomography experiments have revealed that ZIKV infection leads to a reorganization
481 of microtubules and intermediate filaments, forming a cytoskeletal cage that surrounds viral
482 factories[29,30]. The remodelling of the cytoskeleton around these sites of viral RNA replication
483 and virion assembly confines these processes to a closed environment where components of
484 the innate immune response cannot sense them. Our results reveal that uORF1-KO viruses are
485 delayed in the formation of the cytoskeletal cage in infected cells and suggest a potential role
486 of uORF1 in avoiding innate immune recognition.
487
488 Vimentin is a type III intermediate filament and displays typical features of such proteins, with
489 an N-terminal domain (~100 aa), a central α-helical rod (310 aa) and a C-terminus of variable

490 length (15-1400 aa). The central rod domain allows the formation of coiled-coil dimers that
491 can assemble into higher-order filaments. The bundling of peripheral vimentin into thicker
492 assemblies in the perinuclear area of the cells, as observed upon uORF1 transfection, has been
493 previously observed in normal cells[44]. It has been associated with post-translational
494 modifications of vimentin, such as phosphorylation of the head domain[44,45] or deacetylation
495 (by histone deacetylase 6) upon expression of certain oncogenes[46]. It is unknown whether
496 vimentin is subjected to any post-translational modifications upon ZIKV infection but, based
497 on the α-helical prediction of AlphaFold2, the uORF1 protein could potentially interact with
498 the central rod domain of vimentin, intercalate during dimer formation, and antagonise correct
499 polymerization.
500
501 In addition, vimentin, in conjunction with nestin, are the first intermediate filaments expressed
502 during early brain development[47]. During neural development, the cytoskeleton is involved in
503 different processes (e.g., orientation of the mitotic spindle, interkinetic nuclear migration,
504 microtubule organising centre centrosome activity and assembly of primary cilia) that are
505 fundamental for the correct generation and maturation of neurons[48]. However, the disruption
506 of any of these processes impairs neurogenesis (differentiation of NPCs into mature neurons)
507 in the developing brain, provoking neurological defects such as microcephaly, the key
508 pathological aspect of ZIKV infection. It has been previously described that ZIKV infection
509 can trigger premature differentiation by increasing vertically oriented NPCs[49] and shortened
510 cilia[50]. Our results suggest that the interaction of uORF1 with the cytoskeleton might also play
511 a role in premature neurogenesis, although this needs further characterisation in a relevant
512 human model that can mimic the 3D organisational structure of the developing brain, such as
513 cerebral organoids.
514

515 In this work, we have used ALI-COs, a highly accessible slice culture version of whole-cortical
516 organoids for viral infections, with improved oxygenation, nutrition and longevity[32,33]. This
517 allowed optimal characterisation of the neurotropism exerted by the different ZIKV 5′-UTR
518 mutant viruses. Our finding that the uORF1 knock-out virus has a defect in infecting ALI-COs
519 suggests that the protein encoded by uORF1 - only present in the Asian/American strains - is
520 an essential neurovirulence factor required for ZIKV neuropathogenesis in the developing
521 brain. Interestingly, a virus with the African uORF in the American virus background was also
522 impaired, although to a lesser extent, in brain organoids and iPSC-derived cortical neurons.
523 This may explain, at least in part, how a mutation in the 5′ region in a rarely detected pathogen
524 such as ZIKV drove a major epidemic in the Americas in 2015/2016. Until now, the only
525 mechanism associated with an increase in neuropathogenesis in epidemic Asian/American
526 strains compared to pre-epidemic Asian strains (e.g., Cambodia 2010) was the S139N mutation
527 in the membrane precursor (prM) protein[51]. However, the different mutant viruses had similar
528 infection, dissemination, and transmission dynamics in *Aedes* mosquitoes, regardless of their
529 5′-UTR arrangement, although these uORFs are also translated in the mosquito vector,
530 suggesting that they may uniquely function in mammalian hosts.
531
532 It is, therefore, necessary to further investigate and characterize the role of these 5′ uORFs in
533 the neurovirulence and neuropathogenesis of flaviviruses. To date, they have only been
534 described or predicted in neurotropic flaviviruses (ZIKV, JEV[43] and WNV, **Supp. Fig. 8B**) and
535 they could drive phenotypic divergence with visceral flaviviruses such as dengue or yellow
536 fever viruses.
537
538
539

540 **Materials and Methods**

541 **Cells and viruses:** Vero (ATCC, CCL81) and U-251 (ATCC, CRL-1620) cells were maintained
542 in Dulbecco's modification of Eagle's medium (DMEM) supplemented with 5% and 10%
543 (vol/vol) foetal calf serum (FCS), respectively. Culture medium also contained 100 U/ml
544 penicillin, 100 µg/ml streptomycin and 2 mM L-glutamine. Cells were incubated at 37 °C in
545 the presence of 5% CO_2. *Aedes albopictus* C6/36 cells (ATCC, CLR 1660) were maintained in
546 L-15 medium supplemented with 10% FCS at 28 ºC without CO_2. The American strain
547 ZIKV/Brazil/PE243/2015 (Genbank accession number KX197192.1) was kindly provided by
548 Prof. Alain Kohl (Centre for Virus Research, University of Glasgow, United Kingdom), and
549 the African strain ZIKV/*A.taylori*-tc/SEN/1984/41662-DAK (Genbank accession number
550 KU955592.1) was obtained from the European Virus Archive. ZIKV mutant viruses were
551 rescued by electroporation of Vero cells with SP6-transcribed RNAs of pCC1-SP6-ZIKV
552 derivatives as previously described[24]. Virus stocks were amplified in Vero cells and
553 concentrated by ultracentrifugation (110,500 × g for 3h at 4ºC). All mutant viruses were
554 passaged five times at a low MOI (0.01), and viral RNA was subjected to RT-PCR analysis to
555 confirm the presence or reversion of the introduced mutations.

556 Upon reaching 70-80% confluence, Vero and U251 cells were infected with parental or mutant
557 viruses in serum-free DMEM supplemented with 100 U/ml penicillin, 100 µg/ml streptomycin
558 and 2 mM L-glutamine at the indicated MOI. After adsorption for 1h at 37 °C, the inoculum
559 was removed and replaced with DMEM containing 2% FCS and 20 mM HEPES pH 7.4. In
560 cases where the infection was preceded by transfection, the transfection medium was washed
561 three times with PBS prior to infection. C6/36 cells were infected as above but using L-15
562 medium. Mosquito cells were incubated at 28 ºC.

563

564 **Plasmids:** A pCC1BAC vector containing the open reading frame of the ZIKV BeH819015
565 isolate flanked by the 5′ and 3′ UTRs for ZIKV PE243 (pCC1-SP6-Am ZIKV)[24] was kindly
566 provided by Prof. Andres Merits (University of Tartu, Tartu, Estonia). Mutations in the 5′ UTR
567 were generated by site-directed mutagenesis with the indicated oligonucleotides **(Supp Table**
568 **7)** in a subcloned gene fragment of the first 2,000 nucleotides of ZIKV PE243 maintained in
569 pcDNA.3. Following sequencing to confirm no inadvertent changes had arisen, mutated
570 fragments were cloned back into pCC1-SP6-Am ZIKV at the *EcoR1* and *AvrII* sites.
571 The previously described pSGDLuc vector[52] was used as a template to design the Renilla and
572 the Firefly luciferase reporters (pREN-Luc and pFF-Luc). The pREN-Luc cassette was
573 amplified using the following oligonucleotides 5′- TCCGCCCAGTTCCGC CCATTCTCCGC
574 and 5′-GCGCTCTAGATTATCTCGAGGTGTAGAAATAC and cloned back into pSGDLuc
575 previously digested with *AvrII* and *XbaI*. This created a plasmid expressing only the Renilla
576 luciferase gene under the T7 promoter. For the FF-Luc reporters, the first 194 nucleotides of
577 the ZIKV genome were cloned in the pSGD plasmid previously digested with *XhoI* and *BglII*
578 using a different set of oligonucleotides **(Supp Table 7)** to introduce the uORF1, uORF2 and
579 Main ORF in frame with the 2A-Firefly luciferase gene. In addition, a T7 promoter was
580 introduced upstream of the ZIKV coding region. Mutations within the uORFs were introduced
581 by site-directed mutagenesis.
582 The uORF1, uORF2 and African uORF coding sequences were amplified using specific
583 oligonucleotides **(Supp Table 7)** and cloned in frame at the C-terminus with a FLAG-, TAP-
584 or mCherry tag. *PacI*- and *AflII*-digested PCR fragments were ligated into pCAG previously
585 digested with these restriction enzymes. All sequences were confirmed by Sanger sequencing.
586
587 **RNA transcript preparation and luciferase reporter assays:** Transcription reactions were
588 performed using the mMESSAGE mMACHINE™ T7 transcription kit (Thermo Fisher

589 Scientific) for 2 h at 37 ºC and terminated by treatment with TURBO™ DNase (Thermo Fisher
590 Scientific) for 30 min at 37 ºC. Prior to transfection, RNA transcripts were polyadenylated for
591 30 min at 37 ºC. Vero and U251 cells were transfected in triplicate with Lipofectamine 2000
592 reagent (Invitrogen) by RNA reverse-transfection, where cells in suspension were added
593 directly to the RNA complexes in 96-well plates. For each transfection, 100 ng of purified T7
594 ZIKV-uORFs FF-Luc RNA and 5 ng of purified T7 Ren-Luc RNA plus 0.3 µl Lipofectamine
595 2000 in 20 µl Opti-MEM (Gibco-BRL) supplemented with RNaseOUT (Invitrogen; diluted
596 1:1,000 in Opti-MEM) was added to each well containing 10^5 cells. RNA to be transfected was
597 purified using an RNA Clean and Concentrator kit (Zymo Research). Transfected cells in
598 DMEM supplemented with 2% FCS and 20 mM HEPES pH 7.4 were incubated at 37 ºC for
599 the indicated times. At 6 hours post-transfection, cells were infected with PE243 and Dak84 at
600 MOI:3 by adding 50 µl of DMEM-2% FCS medium-containing virions. Firefly and Renilla
601 luciferase activities were determined using the Dual-Luciferase Stop & Glo Reporter Assay
602 System (Promega). Translation efficiency was calculated as the ratio of FF-Luc units (uORFs-
603 /Main ORF-translation) to Ren-Luc units (transfection control), normalized by the ratio of
604 Main ORF WT-translation.
605
606 **DNA transfection:** Vero and U251 cells were transiently transfected with pCAG-uORF
607 plasmids using a commercial liposome method (TransIT-LT1, Mirus). Transfection mixtures
608 containing plasmid DNA, serum-free medium (Opti-MEM; Gibco-BRL) and liposomes were
609 set up as recommended by the manufacturer and added dropwise to the tissue culture growth
610 medium.
611
612 **Plaque assays:** Viral titres were determined by plaque assay. Viral supernatant was 10-fold
613 serially diluted in serum-free DMEM. Vero cells in 6-well plates were infected with 400 µL of

614 diluted supernatant. After 1h at 37°C with regular rocking, the inoculum was removed and
615 replaced with a 1:1 mixture of 3% low-melting agarose and MEM 2X (20% 10X MEM, 8 mM
616 L-glutamine, 0.45% Na_2CO_3, 0.4% non-essential amino acids, 4% HEPES pH 7.4 and 4%
617 FCS). Plates were incubated at 37°C for 5 days prior to fixing with formal saline (4%
618 formaldehyde and 0.9% NaCl) at room temperature. Cell monolayers were stained with 0.1%
619 toluidine blue to visualise plaques. The number of plaque-forming units per mL (PFU/mL) was
620 calculated as nº plaques * 2.5. The final number of PFU/mL per time point was calculated by
621 averaging the results of three biological repeats.
622

623 **$TCID_{50}$ assays:** ZIKV replication was assessed using a 50% tissue culture infective dose
624 ($TCID_{50}$) assay in Vero cells. Supernatant derived from infected i^3Neurons cells was subjected
625 to 10-fold serial dilutions. Cells were fixed and stained at 96 h p.i. as indicated above. Wells
626 showing any cytopathic effect (CPE) were scored positive.

627

628 **Virus competition assays:** Dual infection/competition assays were performed in duplicate on
629 Vero and U251 cells using the American WT and different mutant viruses at an equal or 9:1
630 ratio at a total MOI of 0.1. Media collected from infected plates was used for eight blind
631 passages using 1:10,000 volume of obtained virus stock (corresponding to MOI 0.05–0.2).
632 Total RNA from passages 0 (16 h post-infection to determine viral input), 2, 4 and 6 was
633 isolated as previously described[17]. cDNA was synthesised from 1,000 ng total RNA using M-
634 MLV Reverse Transcriptase (Promega) and 1 µL random primers (1:25 dilution of Random
635 Primers, Invitrogen™). The ZIKV 5′ UTR region was amplified using specific primers **(Supp**
636 **Table 7)**. Following PCR reactions, the resulting amplicons were subjected to electrophoresis
637 in 3% agarose gels, and DNA bands were purified using the Zymoclean Gel DNA recovery kit
638 (Zymo Research) and subjected to Sanger sequencing. Quantification was performed by

639 integrating area under the curve for each nucleotide and expressing it as percentage of the total
640 signal at that base.
641
642 **Subcellular fractionation:** Vero or U251 cells seeded in 4-cm dishes were transfected with
643 plasmids of interest. At 40h post-transfection, samples were collected and immediately
644 subjected to subcellular fractionation using the 'Subcellular Protein Fractionation Kit for
645 Cultured Cells' (ThermoFisher Scientific). Purified fractions were stored in Laemmli's buffer
646 at -20°C for downstream biochemical characterisation.
647
648 **Immunoblotting:** Cells were lysed in 1X Laemmli's sample buffer. After denaturation at 98
649 ºC for 5 minutes, proteins were separated by 12 % SDS-PAGE and transferred to a 0.45 µm
650 nitrocellulose membrane. For proteins below 16 kDa, 20% SDS-PAGE and 0.22 µm
651 nitrocellulose membrane was used instead. Membranes were blocked (5% non-fat milk powder
652 or bovine serum albumin in PBST [137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.5 mM
653 KH_2PO_4, pH 6.7, and 0.1% Tween 20]) for 30 min at room temperature and probed with
654 specific primary antibodies **(Supp Table 8)** at 4ºC overnight. Membranes were incubated in
655 the dark with IRDye-conjugated secondary antibodies diluted to the recommended
656 concentrations in PBST for 1 h at room temperature. Blots were scanned using an Odyssey
657 Infrared Imaging System (Licor).
658
659 **Selective 2′ hydroxyl acylation analysis by primer extension (SHAPE):** SHAPE reactions
660 were conducted as previously described[53]. 2 pmol of purified *in vitro* transcribed RNA template
661 (full-length WT ZIKV RNA and mutant derivatives) was denatured at 95 °C for 3 minutes, then
662 incubated on ice for 3 minutes before the addition of folding buffer (final concentration 100
663 mM HEPES pH 8, 6 mM $MgCl_2$, 100 mM NaCl). Refolding was conducted at 37 °C for 30

664 minutes. Samples were divided into two equal volumes and treated with either N-methylisatoic
665 anhydride (NMIA) dissolved in DMSO (10 mM final concentration) or DMSO only, for 45 min
666 at 37 °C. Samples were ethanol precipitated and resuspended in 10 µL ddH_20 before being
667 reverse transcribed. NMIA or DMSO treated RNA samples were mixed with 5′ NED labelled
668 primer (final concentration 1 µM) (5′ CTTCCTAGCATTGATTATTCTCAGCA TG 3′, Applied
669 Biosystems). Annealing was conducted at 85 °C for 3 minutes, 60 °C for 10 minutes, and 35 °C
670 for 10 minutes. Reverse transcription with Superscript III (Invitrogen) was conducted
671 according to the manufacturer's instructions at 53 °C for 40 minutes. Sequencing ladders were
672 prepared by reverse transcription of 1 pmol of non-treated RNA in the presence of ddCTP,
673 using 5′ VIC labelled primer. All primer extensions were terminated by alkaline hydrolysis,
674 neutralised with HCl and ethanol precipitated before being resuspended in ddH_20. The
675 sequencing ladders were divided equally between cDNAs from NMIA- and DMSO-treated
676 RNA. The final samples were dried onto sequencing plates before being resuspended in
677 formamide and fractionated by capillary electrophoresis (Applied Biosystems 3130xl analyser,
678 University of Cambridge Department of Biochemistry DNA Sequencing Facility). SHAPE was
679 performed in triplicate for each genome from separate RNA preparations. Each SHAPE dataset
680 was analysed with the QuSHAPE software[54]. The average normalised NMIA reactivity was
681 calculated for each individual nucleotide and used to model RNA secondary structure using the
682 RNA structure software[55] using the SHAPE reactivity profile as soft pseudo-free energy
683 constraints. RNA secondary structures were drawn using XRNA
684 (http://rna.ucsc.edu/rnacenter/xrna/xrna.html).
685
686 **Immunofluorescence microscopy:** Vero or U251 cells were seeded in a 24-well plate with
687 coverslips to be transfected and/or infected. At the indicated times, cells were fixed with 50:50
688 methanol: acetone for 10 minutes. Coverslips were blocked for 30 minutes in 5% bovine serum

689 albumin (BSA) in 1X PBS at room temperature. Coverslips were incubated in 1% BSA with
690 specific primary antibodies **(Supp Table 8)**. Secondary antibodies were diluted in 1X PBS and
691 applied for 1.5 h at room temperature. Cells were washed with 1X PBS three times prior to
692 incubation with DAPI (1:1,000) for 5 minutes and mounted with Prolong Gold antifade
693 (Thermo Fisher Scientific). All confocal images were captured with a Zeiss LSM 700 laser
694 scanning microscope with the ZEN microscope software. A 63X oil immersion objective with
695 the pinhole set to 1 airy unit was utilized, as well as the same laser power, gain, and zoom for
696 all sets of images and Z-stacks obtained. Z-stacks were obtained with 1024x1024 pixels within
697 a 16-bit range. Image J was used to make representative Z-stack montages and merged Z-stack
698 images[56]. General 20X and 40X immunofluorescence imaging for quantification were
699 performed using an EVOS FL AUTO 2 microscope and the associated-imaging system
700 software.
701
702 **Growth and infection of iPSC-derived i^{3}Neurons:** i^{3}Neuron stem cells were maintained at
703 37 °C in a complete E8 medium (Gibco) on plates coated with Matrigel (Corning) diluted 1:50
704 in DMEM. Initial three-day differentiation was induced with DMEM/F-12, HEPES (Gibco)
705 supplemented with 1X N2 supplement (Thermo Fisher Scientific), 1X NEAA, 1X Glutamax
706 and 2 μg/ml doxycycline. 10 μM rock inhibitor (Y-27632, Tocris) was added during initial
707 plating of cells to be differentiated and cells were plated onto Matrigel-coated plates.
708 Differentiation medium was replaced daily and after three days of differentiation, partially
709 differentiated neurons were re-plated into cortical neuron (CN) media in coverslips in 24-wells
710 coated with 100 μg/ml poly-L-ornithine (PLO, Sigma). Cortical neuron media consisted of:
711 neurobasal plus medium (Gibco) supplemented with 1X B27 supplement (Gibco), 10 ng/ml
712 BDNF (Peprotech), 10 ng/ml NT-3 (Peptrotech), 1 μg/ml laminin (Gibco) and 1 μg/ml
713 doxycycline. After initial replating, neurons were then maintained in CN media without

714 doxycycline for a further 11 days until mature. Fully differentiated neurons were infected with
715 ZIKV viruses at an MOI 10 in CN media. After 2 hours, the virus inoculum was removed and
716 replaced with 50% fresh – 50% conditioned media. After 96 h p.i., infected cells were fixed
717 1:1 using 20 mM PIPES pH 6.8, 300 mM NaCl, 10 mM EGTA, 10 mM glucose, 10 mM $MgCl_2$,
718 2% (w/v) sucrose and 5% (v/v) PFA for 10 minutes; and permeabilised for 15 min with 0.3%
719 Triton X-100. Cells were blocked, incubated and imaged as described above.
720
721 **Generation, sampling, preparation and infection of human cortical organoids:**
722 ALI-COs were prepared as previously described[32,33]. At 82 days *in vitro*, ALI-CO slices were
723 infected with ZIKV at MOI 5 in brain organoid slice media, consisting of: neurobasal medium
724 (Thermo Fisher Scientific) supplemented with 1x B27 supplement (Thermo Fisher Scientific),
725 0.45% (w/v) glucose (Sigma-Aldrich), 1x Glutamax (Thermo Fisher Scientific) and 1%
726 antibiotic-antimycotic (Thermo Fisher Scientific). After 24 hours, the virus inoculum was
727 removed and replaced with fresh media. Culture medium was replaced by fresh medium every
728 24 h, and at 7 days post-infection, ALI-COs were fixed with 4% PFA for 3 h. Before cryo-
729 sectioning, ALI-CO tissues were sequentially incubated overnight in 5% and 20% sucrose for
730 cryoprotection and later embedded in OCT. ALI-COs were sectioned at 6 μm thickness. All
731 staining steps were done in permeabilization/blocking buffer (5% BSA, 0.3% Triton X-100 in
732 PBS) at room temperature and their duration was as follows: permeabilization/blocking (1 h);
733 primary antibodies incubation (overnight); wash steps (x3) 10 min each with just PBS 1X;
734 secondary antibodies incubation (1 h); wash steps (x3) 10 min each with just PBS 1X; and
735 DAPI-counterstaining (1:1,000) for 5 min. Sections mounted and cured overnight with Pro-
736 long Gold antifade (Thermo Fisher Scientific) and imaged as described above.
737
738

739 **Mosquitoes**

740 All *in vivo* mosquito experiments used the 15th and 16th laboratory generations of an *Aedes*
741 (*Ae.*) *aegypti* colony originally established from wild specimens caught in Barranquilla,
742 Colombia, in 2017. Mosquitoes were maintained under controlled insectary conditions
743 (28° ± 1 °C, 12 h:12 h light: dark cycle and 70% relative humidity) as previously described[34].
744 Larvae were reared in dechlorinated tap water supplemented with a standard diet of TetraMin
745 fish food (Tetra). Adults were kept in insect cages (BugDorm) with permanent access to a 10%
746 sucrose solution.

747

748 **Mosquito exposure to ZIKV**

749 Mosquitoes were orally challenged with ZIKV by membrane feeding as previously
750 described[34,35]. Briefly, 7-day-old female mosquitoes were starved for 24 h prior to the
751 infectious blood meal. The artificial blood meal consisted of a 2:1 mix of washed rabbit
752 erythrocytes (BCL) and ZIKV suspension containing 6.0 x 10^6 PFU/mL, supplemented with
753 10 mM adenosine triphosphate (Merck). The infectious blood meal was offered to mosquitoes
754 for 15 minutes via a membrane-feeding apparatus (Hemotek Ltd.) with porcine intestine as the
755 membrane. Mosquitoes were allowed blood feed for 15 minutes, and fully engorged females
756 were sorted on ice, transferred into 1-pint cardboard containers and maintained under
757 controlled conditions (28° ± 1 °C, 12 h:12 h light: dark cycle and 70% relative humidity) in a
758 climatic chamber with permanent access to 10% sucrose solution for 21 days in the first
759 experiment and 17 days in the second and third experiments.

760 At 7, 14 and 21 days post-blood feeding in the first experiment and 7, 10, 14 and 17 days post-
761 blood feeding in the second and third experiments, mosquitoes were paralyzed using
762 triethylamine (Sigma) for 5 minutes to collect saliva *in vitro* as previously described[35]. Briefly,
763 the legs of each mosquito were removed, and the proboscis was inserted into a 20-µL pipet tip

764 containing 10 μL of foetal bovine serum (FBS). After 30 minutes of salivation, the saliva-
765 containing FBS was collected and mixed with 40 μL of DMEM (Sigma) containing 2%
766 PenStrep (Gibco) and then stored at −80 °C until use. After the salivation, the head and body
767 remainder of each mosquito were separated with forceps and scalpels and stored individually
768 in separate tubes at −80 °C.
769
770 **ZIKV detection in mosquitoes**
771 To assess transmission potential, infectious ZIKV was detected in saliva samples by focus-
772 forming assay in Vero E6 cells as previously described[34]. Briefly, the saliva samples were
773 inoculated onto Vero E6 cells in 96-well plates after removing the cell culture supernatant and
774 incubated at 37 °C for 1 hour. The inoculum was removed and replaced by DMEM containing
775 1.6% carboxymethylcellulose, 1% FBS, 4% Anti-Anti 100X (Gibco) and 1% PenStrep
776 (Gibco). After 5 days of incubation at 37 °C, cells were fixed with 4% formaldehyde solution
777 (Sigma) for 30 minutes. After fixation, cells were permeabilized with 1X PBS (Gibco)
778 containing 0.1% Triton X-100 (Sigma) for 10 minutes, blocked with 1% BSA (Sigma) in PBS,
779 and then reacted with mouse anti-flavivirus group antigen monoclonal antibody clone D1-4G2-
780 4-15 (Merck) in PBS for 1 h at room temperature (20-25°C). After washing with PBS three
781 times, cells were incubated with a 1:1,000 dilution of Alexa Fluor 488-conjugated goat anti-
782 mouse IgG (Invitrogen) in PBS for 1 h at room temperature. Immunopositive signals were
783 confirmed under a fluorescence microscope EVOS FL (Thermo Fisher Scientific) with
784 appropriate barrier and excitation filters. Transmission potential was determined qualitatively
785 (i.e., positive or negative).
786 To determine the rates of ZIKV infection and systemic dissemination, mosquito heads and
787 bodies were placed individually in 300 μL of squash buffer containing 10 mM of Tris pH8.0
788 (Thermo Fisher Scientific), 50 mM of NaCl (Thermo Fisher Scientific) and 1.27 mM of EDTA

789 pH8.0 (Thermo Fisher Scientific) supplemented with proteinase K (Eurobio Scientific) at a
790 final concentration of 0.35 mg/mL and homogenized. A 100-uL sample of the homogenate was
791 transferred to a PCR plate and incubated for 5 minutes at 56 °C, followed by 10 minutes at 98
792 °C to extract total RNA. Detection of ZIKV RNA was performed by two-step RT-PCR.
793 According to the manufacturer's protocol, the cDNA was synthesized using M-MLV reverse
794 transcriptase (Thermo Fisher Scientific) with the extracted RNA and random hexameric
795 primers. The synthesized cDNA was subsequently amplified by PCR using DreamTaq DNA
796 polymerase (Thermo Fisher Scientific) with the following primer pair: ZIKV-PCR-Fw (5′-
797 GTATGGAATGGAGATAAGGCCCA-3′) and ZIKV-PCR-Rv (5′-ACCAGCACTGCCATT
798 GATGTGC-3′), according to the manufacturer's protocol. The cycling conditions were as
799 follows: 2 minutes at 95 °C, 35 cycles of 30 seconds at 95 °C, 30 seconds at 55 °C, and
800 90 seconds at 72 °C with a final extension step of 7 minutes at 72 °C. Amplicons were
801 visualized by electrophoresis on a 2% agarose gel.
802 The 5′-UTR mutations of the ZIKV strains were verified by Sanger sequencing in 14-15 head
803 samples for each strain in mosquitoes collected on day 14 post-blood feeding of the second
804 experiment. The synthesized cDNA was amplified using PrimeStar GXL (Takara Bio) with the
805 following primer pair: ZIKV-uORF-PCR-Fw (5′-GTATGGAATGGAGATAAGGCCCA-3′)
806 and ZIKV-uORF-PCR-Rv (5′-TATTGATGAGACCCAGTGATGGC-3′). The cycling
807 conditions were as follows: 2 minutes at 98 °C, 30 cycles of 10 seconds at 98 °C, 15 seconds
808 at 55 °C, and 60 seconds at 68 °C with a final extension step of 2 minutes at 68 °C. Amplified
809 products were directly sequenced in both directions with the following primer pair: ZIKV-
810 uORF-PCR-Fw (same as above) and ZIKV-uORF-SEQ-Rv (5′-GACCCAGCAGAAGTCC
811 GGCTGGC-3′). The expected 5′-UTR sequence was confirmed in all the tested mosquitoes.
812

813 **Statistical analysis of results:** Data were analysed in GraphPad Prism 9.0 (GraphPad
814 Software, San Diego, CA, USA). Values represent mean ± standard deviation. Statistical
815 significance was evaluated using two-tailed *t*-tests on luciferase data or $\log_{10}$(virus titre) data
816 and one-way ANOVA to calculate the percentage of infected cells in brain organoids. Both
817 methods did not assume equal variances for the two populations being compared to calculate
818 the *p*-values. Differences as compared to the control with *p-value* ≤ 0.05 were considered
819 statistically significant, with *$p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ and **** $p < 0.0001$.
820 The prevalence of ZIKV infection, dissemination and transmission in mosquitoes were
821 analyzed by logistic regression as a function of the experiment, ZIKV strain and time point.
822 The initial statistical model included all their interactions, which were removed from the final
823 model if their effect was non-significant ($p<0.05$). Time was considered a continuous variable.
824 To account for minor differences in the measured virus concentration, the $\log_{10}$-transformed
825 blood meal titer was also included in the model as a covariate.
826
827 **Ribosomal profiling and RNA-Seq data:** 10^7 Vero, U251 or C6/36 cells were grown on 10-
828 cm dishes and infected with parental or mutant viruses at MOI 3. At 24 h p.i., cells treated with
829 CHX (Sigma-Aldrich, 100 µg/ml) were incubated for 3 minutes and then rinsed with 5 ml of
830 ice-cold PBS before being flash-frozen. Cells harvested by 'flash-freezing' were directly rinsed
831 with 5 ml of ice-cold PBS, flash-frozen in a dry ice/ethanol bath, and lysed with 400 µl of lysis
832 buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 1% Triton X-100,
833 100 µg/ml cycloheximide and 25 U/ml TURBO DNase). The cells were scraped extensively to
834 ensure lysis, collected and triturated ten times with a 26-G needle. Cell lysates were clarified
835 by centrifugation at 13,000 *g* for 20 min at 4 °C. Lysates were subjected to Ribo-Seq and RNA-
836 Seq based on previously reported protocols[17]. Vero and U251 ribosomal RNA was removed
837 using Ribo-Zero Gold rRNA removal kit (Illumina); and C6/36 ribosomal RNA contamination

838 was removed by treatment with duplex-specific nuclease (DSN) as previously described[57].
839 Library amplicons were constructed using a small RNA cloning strategy adapted to Illumina
840 smallRNA v2 to allow multiplexing. Amplicon libraries were deep-sequenced using an
841 Illumina NextSeq500 platform. Ribo-Seq and RNA-Seq sequencing data have been deposited
842 in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress).
843

844 **Computational analyses of sequence data**

845 Reads were demultiplexed. Adaptors were trimmed off using the FASTX-Toolkit version
846 0.0.14 (http://hannonlab.cshl.edu/fastx_toolkit/), and sequences shorter than 25 nt after
847 trimming were discarded. Sequences not linked to an adaptor and sequences constituted of
848 adaptors ligated together were also removed. Reads remaining after this initial selection were
849 sequentially mapped to (i) ribosomal RNA (rRNA); (ii) the American or African ZIKV genomic
850 RNA (viral genome sequences were confirmed by *de novo* assembly using Trinity
851 version2.8.5); (iii) messenger mRNA (mRNA); (iv) non-coding RNA (ncRNA); (v) genomic
852 DNA (gDNA) and (vi) a contaminants database, using bowtie (version 1.2.3) with parameters
853 -v n_mismatches --best (i.e., maximum of n_mismatches mismatches permitted, report best
854 match) where n_mismatches was 1 for mapping in U251 cells and 2 for Vero and C6/36 cells.
855 The remaining reads were classified as unmapped. The rRNA databases included the Genbank
856 accession numbers listed in **Supp Table 9**. The mRNA sequences for each host were
857 downloaded from NCBI RefSeq. The ncRNA databases for *C. sabaeus* were retrieved from
858 Ensembl release 91 and compiled with 534 tRNA sequences of *M. mulatta* and Genbank
859 accession XM_007980346; for *H. sapiens,* the ncRNA sequences were downloaded from
860 Ensembl release 105 and for *A. albopictus* from Ensembl release 55. The gDNA sequences for
861 each host were retrieved from the same Ensembl releases as the ncRNA. The contaminants
862 database comprises Genbank accessions of potential contaminants that might be found in the

863 lab and are used as quality control to ensure that no/very few reads map to this database. The
864 database includes viruses (e.g., HSV-1, PRRSV) and bacteria (e.g., *E. coli*) that are regularly
865 used within the Division of Virology, University of Cambridge, and other common potential
866 contaminants (e.g., mycoplasma sequences). The composition of libraries (**Supp Fig. 9**)
867 indicated some contamination of Vero Dak84 RNA-Seq (replicate 1 and 2) with Toscana virus
868 (TOSV). These reads do not exhibit the same features as the rest of the reads in those libraries
869 (i.e., length distribution of the reads does not have the same shape), indicating that the
870 contamination occurred after lysates were harvested and thus, they do not affect our
871 conclusions.
872
873 Quality control analyses of the reads (**Supp Figs. 10-13**), including length distribution,
874 phasing, and phasing per read length, were carried out using Ribo-Seq and RNA-Seq reads
875 mapping to host mRNA and viral RNA as previously described[17]. To assess ribonucleoprotein
876 (RNP) contamination of Ribo-Seq samples, a comparison was made between the number of
877 reads mapping to CDSs and the 3′ UTR of host mRNAs only (**Supp Fig. 14**).
878
879 For the visualisation of read densities on the viral genome, the position of the P-site was
880 calculated by adding a fixed +12 nt offset to the 5′ end of each read. Read densities were
881 normalized by the sum of total viral RNA and total host mRNA and plotted in reads per million
882 mapped reads (RPM). The genomic regions with coordinates 15-125 or 18-315 for the parental
883 and mutant viruses, respectively, were chosen for the zoom plots of the 5′ UTR. The region
884 from genomic coordinate 10,370 to the end of the genome was used for the zoom plots of the
885 3′ UTR. Summarising bar charts show the percentage of reads in each phase, normalised by
886 the total number of reads in the designated region (**Supp Table 10**). The reads per kilobase per

887 million mapped reads (RPKM) values for Ribo-Seq were divided by the corresponding RNA-
888 Seq values to calculate the translational efficiency (TE).
889

890 **Acknowledgements**

891 The authors are indebted to Andrew Firth (Department of Pathology, University of Cambridge)
892 for all the intellectual input and technical support in the bioinformatic analysis, and for the
893 critical reading of this manuscript. The authors would like to thank Alain Kohl (Centre
894 for Virus Research, University of Glasgow) and Lindomar J. Pena and Rafael
895 Oliveira de Freitas França (Fiocruz Recife, Pernambuco, Brazil) for the provision of PE243
896 ZIKV RNA used to generate the virus stock; and Andres Merits (University of Tartu, Tartu,
897 Estonia) for providing the pCC1-SP6-Am ZIKV plasmid. The African ZIKV isolate (Dak84)
898 was provided by the 'European Virus Archive goes Global (EVAg) project' funded by the
899 European Union's Horizon 2020 research and innovation programme under grant agreement
900 No 653316. We thank Catherine Lallemand for assistance with mosquito rearing. We are
901 grateful to Claudia Romero-Vivas, who initially provided the mosquito colony from Colombia.
902 NI would like to thank David Carpentier (Confocal microscopy facility, Department of
903 Pathology, University of Cambridge) and Louise Howard (Histology Facility, Department of
904 Pathology, University of Cambridge) for their helpful assistance.
905

906 C.L. was supported by the 'Wolfson-Pathology PhD studentship' (University of Cambridge).
907 G.M.C. was supported by a Wellcome Trust grant [203864/Z/16/Z]. A.M.D. and H.S. were
908 supported by the Wellcome Trust Senior Research Fellowships [106207/Z/14/Z,
909 220814/Z/20/Z] and a European Research Council grant [646891] awarded to Andrew Firth
910 (Department of Pathology, University of Cambridge). J.E.D. and A.S.N. were supported by a
911 Wellcome Trust Senior Research Fellowship [219447/Z/19/Z] awarded to J.E.D. L.W.M. and

I.G. were supported by a Wellcome Trust grant [207498/Z/17/Z] to I.G. Sir Henry Dale Fellowship [220620/Z/20/Z] from the Wellcome Trust and the Royal Society, MRC project grant [MR/T000376/1] and an Isaac Newton Trust/Wellcome Trust ISSF/University of Cambridge Joint Research Grant to V.L. Research in the A.L. laboratory was funded by the Medical Research Council UK (UKRI: MR/P008658/1; MR/X006867/1 awarded to A.L.). This work was also supported by the French Government's Investissement d'Avenir program Laboratoire d'Excellence Integrative Biology of Emerging Infectious Diseases (grant ANR-10-LABX-62-IBEID to L.L.) and MSDAVENIR (grant INTRANZIGEANT to L.L.). S.T. was supported by an MSCA fellowship from the European Union (grant MSCA-2021-PF-01 ZIKVMosTransmit). I.B. was supported by grants from the Wellcome Trust [202797/Z/16/Z] and the Biotechnology and Biological Sciences Research Council UK [BB/V000306/1]. N.I. was supported by an Isaac Newton Trust Grant [18.40r], a Royal Society Research Grant [RGS\R1\191137], an Isaac Newton Trust/Wellcome Trust ISSF/University of Cambridge Joint Research Grant and partially by the NIH Grant [5R21AI147172].

References:

1. Lazear, H.M., and Diamond, M.S. (2016). Zika Virus: New Clinical Syndromes and Its Emergence in the Western Hemisphere. J Virol *90*, 4864–4875. 10.1128/JVI.00252-16.
2. Cao-Lormeau, V.-M., Blake, A., Mons, S., Lastère, S., Roche, C., Vanhomwegen, J., Dub, T., Baudouin, L., Teissier, A., Larre, P., et al. (2016). Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study. Lancet *387*, 1531–1539. 10.1016/S0140-6736(16)00562-6.
3. Noronha, L. de, Zanluca, C., Azevedo, M.L.V., Luz, K.G., and Santos, C.N.D. Dos (2016). Zika virus damages the human placental barrier and presents marked fetal neurotropism. Mem Inst Oswaldo Cruz *111*, 287–293. 10.1590/0074-02760160085.
4. Mlakar, J., Korva, M., Tul, N., Popović, M., Poljšak-Prijatelj, M., Mraz, J., Kolenc, M., Resman Rus, K., Vesnaver Vipotnik, T., Fabjan Vodušek, V., et al. (2016). Zika Virus Associated with Microcephaly. N Engl J Med *374*, 951–958. 10.1056/NEJMoa1600651.
5. Moore, C.A., Staples, J.E., Dobyns, W.B., Pessoa, A., Ventura, C. V., Fonseca, E.B. da, Ribeiro, E.M., Ventura, L.O., Neto, N.N., Arena, J.F., et al. (2017). Characterizing the Pattern of Anomalies in Congenital Zika Syndrome for Pediatric Clinicians. JAMA Pediatr *171*, 288. 10.1001/jamapediatrics.2016.3982.
6. Paixao, E.S., Cardim, L.L., Costa, M.C.N., Brickley, E.B., de Carvalho-Sauer, R.C.O., Carmo, E.H., Andrade, R.F.S., Rodrigues, M.S., Veiga, R. V, Costa, L.C., et al. (2022). Mortality from Congenital Zika Syndrome - Nationwide Cohort Study in Brazil. N Engl J Med *386*, 757–767. 10.1056/NEJMoa2101195.

7. Valente, P.K. (2017). Zika and Reproductive Rights in Brazil: Challenge to the Right to Health. Am J Public Health *107*. 10.2105/AJPH.2017.303924.
8. Henderson, A.D., Aubry, M., Kama, M., Vanhomwegen, J., Teissier, A., Mariteragi-Helle, T., Paoaafaite, T., Teissier, Y., Manuguerra, J.-C., Edmunds, J., et al. (2020). Zika seroprevalence declines and neutralizing antibodies wane in adults following outbreaks in French Polynesia and Fiji. Elife *9*. 10.7554/eLife.48460.
9. Gilbert, R.K., Petersen, L.R., Honein, M.A., Moore, C.A., and Rasmussen, S.A. (2023). Zika virus as a cause of birth defects: Were the teratogenic effects of Zika virus missed for decades? Birth Defects Res *115*, 265–274. 10.1002/bdr2.2134.
10. Lambrechts, L. (2021). Did Zika virus attenuation or increased virulence lead to the emergence of congenital Zika syndrome? J Travel Med *28*. 10.1093/jtm/taab041.
11. Rice, C.M., Lenches, E.M., Eddy, S.R., Shin, S.J., Sheets, R.L., and Strauss, J.H. (1985). Nucleotide Sequence of Yellow Fever Virus: Implications for Flavivirus Gene Expression and Evolution. Science (1979) *229*, 726–733. 10.1126/science.4023707.
12. Wolin, S.L., and Walter, P. (1988). Ribosome pausing and stacking during translation of a eukaryotic mRNA. EMBO J *7*, 3559–3569. 10.1002/j.1460-2075.1988.tb03233.x.
13. Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-Wide Analysis in Vivo of Translation with Nucleotide Resolution Using Ribosome Profiling. Science (1979) *324*, 218–223. 10.1126/science.1168978.
14. Donald, C.L., Brennan, B., Cumberworth, S.L., Rezelj, V. V., Clark, J.J., Cordeiro, M.T., Freitas de Oliveira França, R., Pena, L.J., Wilkie, G.S., Da Silva Filipe, A., et al. (2016). Full Genome Sequence and sfRNA Interferon Antagonist Activity of Zika Virus from Recife, Brazil. PLoS Negl Trop Dis *10*, e0005048. 10.1371/journal.pntd.0005048.
15. McDonald, E.M., Duggal, N.K., and Brault, A.C. (2017). Pathogenesis and sexual transmission of Spondweni and Zika viruses. PLoS Negl Trop Dis *11*, e0005990. 10.1371/journal.pntd.0005990.
16. Echavarría-Consuegra, L., Cook, G.M., Busnadiego, I., Lefèvre, C., Keep, S., Brown, K., Doyle, N., Dowgier, G., Franaszek, K., Moore, N.A., et al. (2021). Manipulation of the unfolded protein response: A pharmacological strategy against coronavirus infection. PLoS Pathog *17*, e1009644. 10.1371/journal.ppat.1009644.
17. Irigoyen, N., Firth, A.E., Jones, J.D., Chung, B.Y.-W., Siddell, S.G., and Brierley, I. (2016). High-Resolution Analysis of Coronavirus Gene Expression by RNA Sequencing and Ribosome Profiling. PLoS Pathog *12*. 10.1371/journal.ppat.1005473.
18. Artieri, C.G., and Fraser, H.B. (2014). Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation. Genome Res *24*, 2011–2021. 10.1101/gr.175893.114.
19. Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome Profiling of Mouse Embryonic Stem Cells Reveals the Complexity and Dynamics of Mammalian Proteomes. Cell *147*, 789–802. 10.1016/j.cell.2011.10.002.
20. Akiyama, B.M., Laurence, H.M., Massey, A.R., Costantino, D.A., Xie, X., Yang, Y., Shi, P.-Y., Nix, J.C., Beckham, J.D., and Kieft, J.S. (2016). Zika virus produces noncoding RNAs using a multi-pseudoknot structure that confounds a cellular exonuclease. Science (1979) *354*, 1148–1152. 10.1126/science.aah3963.
21. Pettersson, J.H.-O., Eldholm, V., Seligman, S.J., Lundkvist, Å., Falconar, A.K., Gaunt, M.W., Musso, D., Nougairède, A., Charrel, R., Gould, E.A., et al. (2016). How Did Zika Virus Emerge in the Pacific Islands and Latin America? mBio *7*. 10.1128/mBio.01239-16.
22. Hinnebusch, A.G., and Lorsch, J.R. (2012). The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges. Cold Spring Harb Perspect Biol *4*, a011544–a011544. 10.1101/cshperspect.a011544.
23. Hinnebusch, A.G., Ivanov, I.P., and Sonenberg, N. (2016). Translational control by 5′-untranslated regions of eukaryotic mRNAs. Science (1979) *352*, 1413–1416. 10.1126/science.aad9868.
24. Mutso, M., Saul, S., Rausalu, K., Susova, O., Žusinaite, E., Mahalingam, S., and Merits, A. (2017). Reverse genetic system, genetically stable reporter viruses and packaged subgenomic

1002 replicon based on a Brazilian Zika virus isolate. Journal of General Virology *98*, 2712–2724.
1003 10.1099/jgv.0.000938.
1004 25. de Borba, L., Villordo, S.M., Iglesias, N.G., Filomatori, C. V., Gebhard, L.G., and Gamarnik,
1005 A. V. (2015). Overlapping Local and Long-Range RNA-RNA Interactions Modulate Dengue
1006 Virus Genome Cyclization and Replication. J Virol *89*, 3430–3437. 10.1128/JVI.02677-14.
1007 26. Sanford, T.J., Mears, H. V, Fajardo, T., Locker, N., and Sweeney, T.R. (2019). Circularization
1008 of flavivirus genomic RNA inhibits de novo translation initiation. Nucleic Acids Res *47*, 9789–
1009 9802. 10.1093/nar/gkz686.
1010 27. Lulla, V., and Firth, A. (2019). Enterovirus Competition Assay to Assess Replication Fitness.
1011 Bio Protoc *9*. 10.21769/BioProtoc.3233.
1012 28. Atkins, J.F., Loughran, G., Bhatt, P.R., Firth, A.E., and Baranov, P. V. (2016). Ribosomal
1013 frameshifting and transcriptional slippage: From genetic steganography and cryptography to
1014 adventitious use. Nucleic Acids Res, gkw530. 10.1093/nar/gkw530.
1015 29. Zhang, Y., Zhao, S., Li, Y., Feng, F., Li, M., Xue, Y., Cui, J., Xu, T., Jin, X., and Jiu, Y. (2022).
1016 Host cytoskeletal vimentin serves as a structural organizer and an RNA-binding protein
1017 regulator to facilitate Zika viral replication. Proceedings of the National Academy of Sciences
1018 *119*. 10.1073/pnas.2113909119.
1019 30. Cortese, M., Goellner, S., Acosta, E.G., Neufeldt, C.J., Oleksiuk, O., Lampe, M., Haselmann,
1020 U., Funaya, C., Schieber, N., Ronchi, P., et al. (2017). Ultrastructural Characterization of Zika
1021 Virus Replication Factories. Cell Rep *18*. 10.1016/j.celrep.2017.02.014.
1022 31. Wang, C., Ward, M.E., Chen, R., Liu, K., Tracy, T.E., Chen, X., Xie, M., Sohn, P.D., Ludwig,
1023 C., Meyer-Franke, A., et al. (2017). Scalable Production of iPSC-Derived Human Neurons to
1024 Identify Tau-Lowering Compounds by High-Content Screening. Stem Cell Reports *9*, 1221–
1025 1233. 10.1016/j.stemcr.2017.08.019.
1026 32. Szebényi, K., Wenger, L.M.D., Sun, Y., Dunn, A.W.E., Limegrover, C.A., Gibbons, G.M.,
1027 Conci, E., Paulsen, O., Mierau, S.B., Balmus, G., et al. (2021). Human ALS/FTD brain organoid
1028 slice cultures display distinct early astrocyte and targetable neuronal pathology. Nat Neurosci
1029 *24*, 1542–1554. 10.1038/s41593-021-00923-4.
1030 33. Giandomenico, S.L., Mierau, S.B., Gibbons, G.M., Wenger, L.M.D., Masullo, L., Sit, T.,
1031 Sutcliffe, M., Boulanger, J., Tripodi, M., Derivery, E., et al. (2019). Cerebral organoids at the
1032 air-liquid interface generate diverse nerve tracts with functional output. Nat Neurosci *22*, 669–
1033 679. 10.1038/s41593-019-0350-2.
1034 34. Baidaliuk, A., Miot, E.F., Lequime, S., Moltini-Conclois, I., Delaigue, F., Dabo, S., Dickson,
1035 L.B., Aubry, F., Merkling, S.H., Cao-Lormeau, V.-M., et al. (2019). Cell-Fusing Agent Virus
1036 Reduces Arbovirus Dissemination in Aedes aegypti Mosquitoes *In Vivo*. J Virol *93*.
1037 10.1128/JVI.00705-19.
1038 35. Aubry, F., Jacobs, S., Darmuzey, M., Lequime, S., Delang, L., Fontaine, A., Jupatanakul, N.,
1039 Miot, E.F., Dabo, S., Manet, C., et al. (2021). Recent African strains of Zika virus display higher
1040 transmissibility and fetal pathogenicity than Asian strains. Nat Commun *12*. 10.1038/s41467-
1041 021-21199-z.
1042 36. Andreev, D.E., Loughran, G., Fedorova, A.D., Mikhaylova, M.S., Shatsky, I.N., and Baranov,
1043 P. V (2022). Non-AUG translation initiation in mammals. Genome Biol *23*, 111.
1044 10.1186/s13059-022-02674-2.
1045 37. Noderer, W.L., Flockhart, R.J., Bhaduri, A., Diaz de Arce, A.J., Zhang, J., Khavari, P.A., and
1046 Wang, C.L. (2014). Quantitative analysis of mammalian translation initiation sites by
1047 <scp>FACS</scp> -seq. Mol Syst Biol *10*, 748. 10.15252/msb.20145136.
1048 38. Mufrrih, M., Chen, B., and Chan, S.-W. (2021). Zika Virus Induces an Atypical Tripartite
1049 Unfolded Protein Response with Sustained Sensor and Transient Effector Activation and a
1050 Blunted BiP Response. mSphere *6*. 10.1128/mSphere.00361-21.
1051 39. Ricciardi-Jorge, T., Lummertz da Rocha, E., González-Kozlova, E., Rodrigues-Luiz, G.F.,
1052 Ferguson, B.J., Irigoyen, N., Mansur, D.S. PKR-mediated stress response enhances dengue and
1053 Zika virus replication. https://doi.org/10.1101/2021.04.08.439069v2 (bioRxiv, preprint).
1054 40. Somers, J., Pöyry, T., and Willis, A.E. (2013). A perspective on mammalian upstream open
1055 reading frame function. Int J Biochem Cell Biol *45*, 1690–1700. 10.1016/j.biocel.2013.04.020.

1056 41. Ratinier, M., Boulant, S., Combet, C., Targett-Adams, P., McLauchlan, J., and Lavergne, J.-P.
1057 (2008). Transcriptional slippage prompts recoding in alternate reading frames in the hepatitis C
1058 virus (HCV) core sequence from strain HCV-1. J Gen Virol *89*, 1569–1578.
1059 10.1099/vir.0.83614-0.
1060 42. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool,
1061 K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction
1062 with AlphaFold. Nature *596*, 583–589. 10.1038/s41586-021-03819-2.
1063 43. Rizvi, V.A., Sarkar, M., Roy, R. Translation regulation of Japanese encephalitis virus revealed
1064 by ribosome profiling. https://doi.org/10.1101/2020.07.16.206920 (bioRxiv, preprint).
1065 44. Valgeirsdottir, S., Claesson-Welsh, L., Bongcam-Rudloff, E., Hellman, U., Westermark, B., and
1066 Heldin, C.H. (1998). PDGF induces reorganization of vimentin filaments. J Cell Sci *111*, 1973–
1067 1980. 10.1242/jcs.111.14.1973.
1068 45. Lamb, N.J., Fernandez, A., Feramisco, J.R., and Welch, W.J. (1989). Modulation of vimentin
1069 containing intermediate filament distribution and phosphorylation in living fibroblasts by the
1070 cAMP-dependent protein kinase. Journal of Cell Biology *108*, 2409–2422.
1071 10.1083/jcb.108.6.2409.
1072 46. Rathje, L.-S.Z., Nordgren, N., Pettersson, T., Rönnlund, D., Widengren, J., Aspenström, P., and
1073 Gad, A.K.B. (2014). Oncogenes induce a vimentin filament collapse mediated by HDAC6 that
1074 is linked to cell stiffness. Proceedings of the National Academy of Sciences *111*, 1515–1520.
1075 10.1073/pnas.1300238111.
1076 47. Yabe, J.T., Chan, W.K.-H., Wang, F.-S., Pimenta, A., Ortiz, D.D., and Shea, T.B. (2003).
1077 Regulation of the transition from vimentin to neurofilaments during neuronal differentiation.
1078 Cell Motil Cytoskeleton *56*, 193–205. 10.1002/cm.10137.
1079 48. Compagnucci, C., Piemonte, F., Sferra, A., Piermarini, E., and Bertini, E. (2016). The
1080 cytoskeletal arrangements necessary to neurogenesis. Oncotarget *7*. 10.18632/oncotarget.6838.
1081 49. Gabriel, E., Ramani, A., Karow, U., Gottardo, M., Natarajan, K., Gooi, L.M., Goranci-Buzhala,
1082 G., Krut, O., Peters, F., Nikolic, M., et al. (2017). Recent Zika Virus Isolates Induce Premature
1083 Differentiation of Neural Progenitors in Human Brain Organoids. Cell Stem Cell *20*.
1084 10.1016/j.stem.2016.12.005.
1085 50. Saade, M., Ferrero, D.S., Blanco-Ameijeiras, J., Gonzalez-Gobartt, E., Flores-Mendez, M.,
1086 Ruiz-Arroyo, V.M., Martínez-Sáez, E., Ramón y Cajal, S., Akizu, N., Verdaguer, N., et al.
1087 (2020). Multimerization of Zika Virus-NS5 Causes Ciliopathy and Forces Premature
1088 Neurogenesis. Cell Stem Cell *27*. 10.1016/j.stem.2020.10.002.
1089 51. Yuan, L., Huang, X.-Y., Liu, Z.-Y., Zhang, F., Zhu, X.-L., Yu, J.-Y., Ji, X., Xu, Y.-P., Li, G.,
1090 Li, C., et al. (2017). A single mutation in the prM protein of Zika virus contributes to fetal
1091 microcephaly. Science (1979) *358*. 10.1126/science.aam7120.
1092 52. Loughran, G., Howard, M.T., Firth, A.E., and Atkins, J.F. (2017). Avoidance of reporter assay
1093 distortions from fused dual reporters. RNA *23*, 1285–1289. 10.1261/rna.061051.117.
1094 53. Wilkinson, K.A., Merino, E.J., and Weeks, K.M. (2006). Selective 2′-hydroxyl acylation
1095 analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide
1096 resolution. Nat Protoc *1*, 1610–1616. 10.1038/nprot.2006.249.
1097 54. Karabiber, F., McGinnis, J.L., Favorov, O. V., and Weeks, K.M. (2013). QuShape: Rapid,
1098 accurate, and best-practices quantification of nucleic acid probing information, resolved by
1099 capillary electrophoresis. RNA *19*, 63–73. 10.1261/rna.036327.112.
1100 55. Reuter, J.S., and Mathews, D.H. (2010). RNAstructure: software for RNA secondary structure
1101 prediction and analysis. BMC Bioinformatics *11*, 129. 10.1186/1471-2105-11-129.
1102 56. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of
1103 image analysis. Nat Methods *9*, 671–675. 10.1038/nmeth.2089.
1104 57. Chung, B.Y., Hardcastle, T.J., Jones, J.D., Irigoyen, N., Firth, A.E., Baulcombe, D.C., and
1105 Brierley, I. (2015). The use of duplex-specific nuclease in ribosome profiling and a user-friendly
1106 software package for Ribo-seq data analysis. RNA *21*. 10.1261/rna.052548.115.
1107

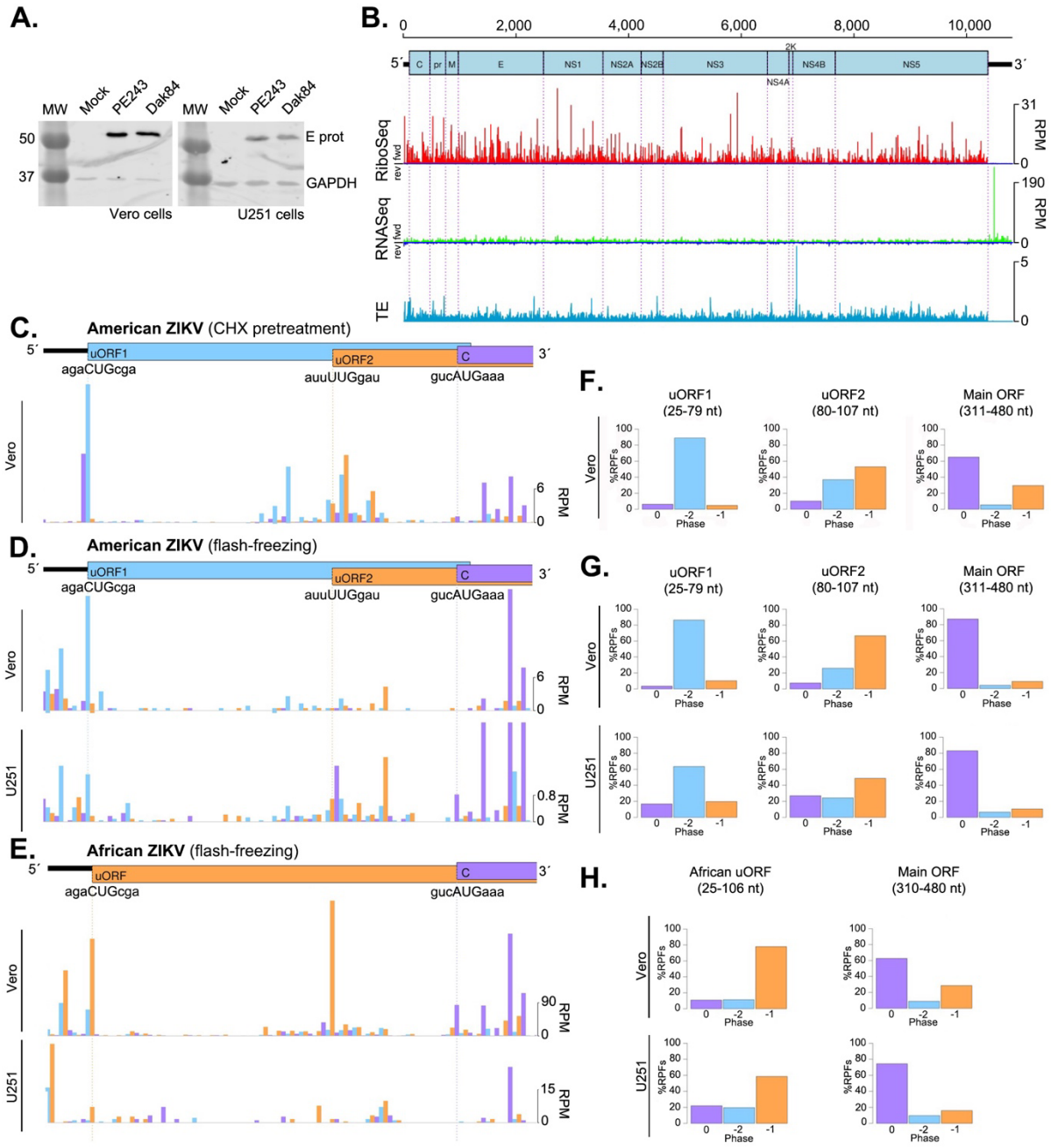
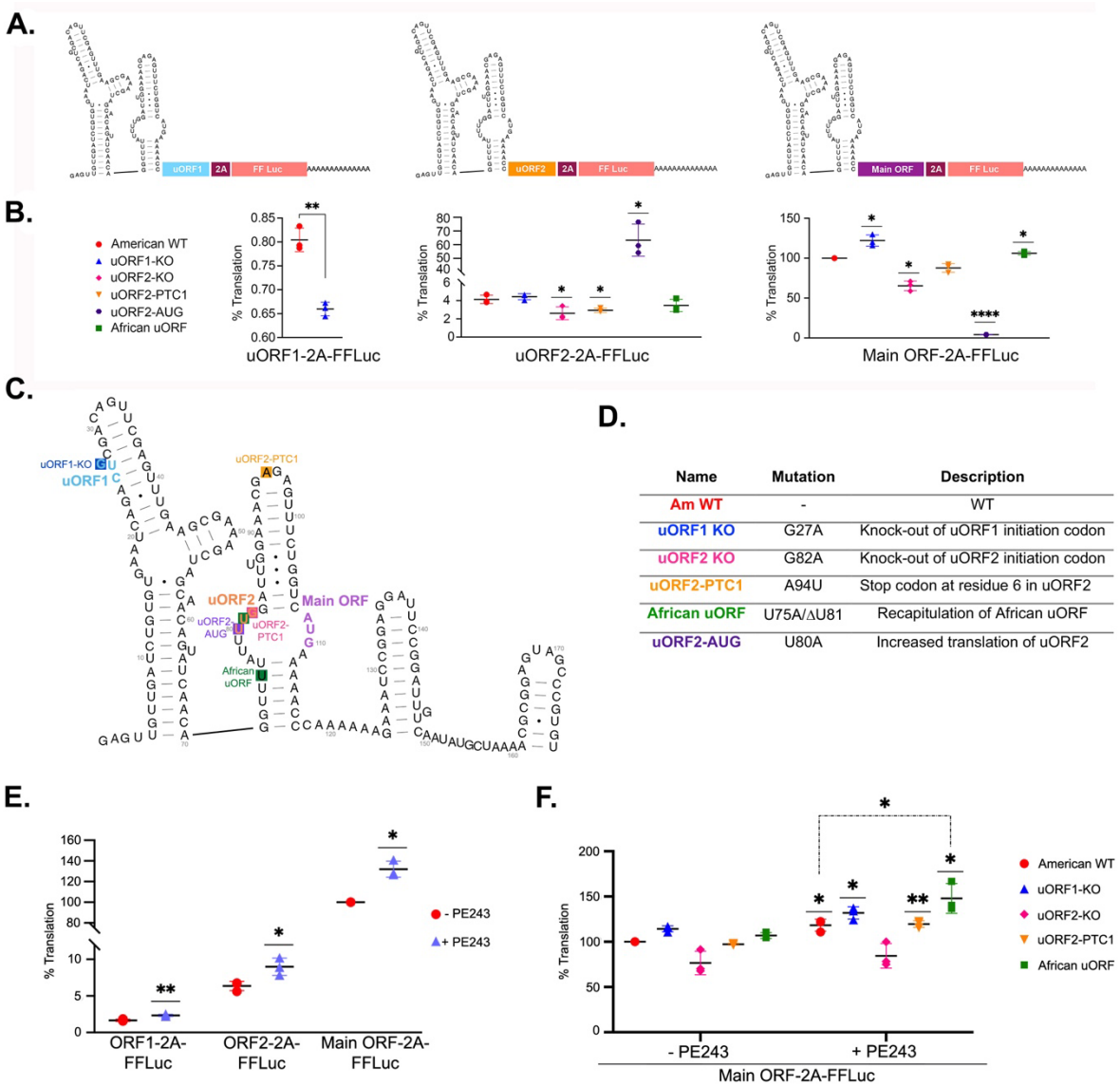


Figure 1. ZIKV RNA synthesis and translation. (**A**) Western blot analysis of ZIKV E protein and GAPDH in Vero and U251 cells infected with PE243 and Dak84 (MOI:3) for 24 h. GAPDH was used as a loading control. Molecular masses (kDa) are indicated on the left. (**B**) Map of the 10807-nt ZIKV/Brazil/PE243/2015 genome. The 5′ and 3′ UTRs are in black, and the polyprotein ORF is in pale blue with subdivisions showing mature cleavage products. Histograms show the read densities, in reads per million mapped reads (RPM), of Ribo-Seq (red) and RNA-Seq (green) reads at 24 h p.i. (repeat 1) in Vero cells pretreated with CHX. The positions of the 5′ ends of reads are plotted with a +12 nt offset to map (for RPFs) approximate P-site positions. Negative-sense reads are shown in dark blue below the horizontal axis. In light blue, the translational efficiency (TE) is calculated as the positive-sense RiboSeq/RNASeq ratio. (**C-D**) The 5′ region of the ZIKV PE243 genome shows two non-AUG uORFs in Vero cells pretreated with CHX (**C**), flash-frozen Vero cells (**D**, upper panel) and flash-frozen U251 cells (**D**, lower panel). (**E**) The 5′ region of the ZIKV Dak84 genome shows the African uORF

in flash-frozen Vero (upper panel) or flash-frozen U251 cells (lower panel). Ribo-Seq counts at 24 h p.i. Histograms show the positions of the 5′ ends of reads with a +12 nt offset to map the approximate P-site. Reads whose 5′ ends map to the first, second or third phases relative to codons in the polyprotein reading frame are indicated in blue, orange or purple, respectively. (**F-H**) Bar charts of the percentage of ribosome-protected fragments (RPFs) in each phase relative the polyprotein ORF. Regions with the least amount of overlapping, (coordinates given in plot titles), were selected. Reads whose 5′ ends map to the -2, -1 or 0 phases are indicated in blue, orange or purple, respectively. CHX-treated Vero cells infected with PE243 (**F**); flash-frozen Vero (**G**, upper panel) and U251 (**G**, lower panel) cells infected with PE243; and flash-frozen Vero (**H**, upper panel) and U251 (**H**, lower panel) cells infected with Dak84.



Name	Mutation	Description
Am WT	-	WT
uORF1 KO	G27A	Knock-out of uORF1 initiation codon
uORF2 KO	G82A	Knock-out of uORF2 initiation codon
uORF2-PTC1	A94U	Stop codon at residue 6 in uORF2
African uORF	U75A/ΔU81	Recapitulation of African uORF
uORF2-AUG	U80A	Increased translation of uORF2

Figure 2. Analysis of ZIKV uORF translation using luciferase reporters. (**A**) Scheme of firefly luciferase (FF-Luc) reporter constructs where the 2A-FFLuc cassette is positioned downstream of and in frame with uORF1 (left), uORF2 (middle) or main ORF (right). (**B**) Relative FF-Luc activity of different mutants for uORF1 (left), uORF2 (middle) and main ORF (right) normalised to Renilla luciferase (Ren-Luc) used as a transfection control. Vero-transfected cells were harvested at 30 hours post-transfection (h p.t.). 100% translation accounted for the main ORF wild-type (WT) translation (right panel in red). (**C**) Scheme of the 5′ region of PE243 with initiation codons for uORF1, uORF2 and main ORF indicated in blue, orange and purple, respectively. Modified nucleotides in the different mutants are indicated by a coloured-filled square associated with the mutant name, briefly described in **D**. (**E**) Relative FF-Luc/Ren-Luc ratio of uORF1-, uORF2-, and main-ORF-2A-FFLuc reporter mRNAs in Vero cells infected with PE243 (MOI:3, purple triangle) or mock-infected (red circle) at 6 h p.t.. Cells were harvested at 24 h p.i.. (**F**) Relative FF-Luc/Ren-Luc ratio of main-ORF-2A-FFLuc mutant reporters in Vero cells infected with PE243 (MOI:3) or mock-infected at 6 h p.t.. Cells were harvested at 24 h p.i.. Experiments were performed in triplicate with three biological replicates. All *t*-tests were two-tailed and did not assume equal variance for the two populations being compared (*$p < 0.05$, ** $p < 0.01$, **** $p < 0.0001$). All *p*-values are from comparisons of the mutant with the respective wild-type in the same ORF.

A.

Name	Mutation	SHAPE analysis	RT-PCR analysis
Am WT	-	WT	WT
uORF1-KO	G27A	WT	uORF1-KO
uORF2-KO	G82A	WT	Revertant
uORF2-PTC1	A94U	WT	uORF2-PTC1
African uORF	U75A/ΔU81	Shorter cHP	African uORF
uORF2-AUG	U80A	No internal loop in SLB	No virus recovered

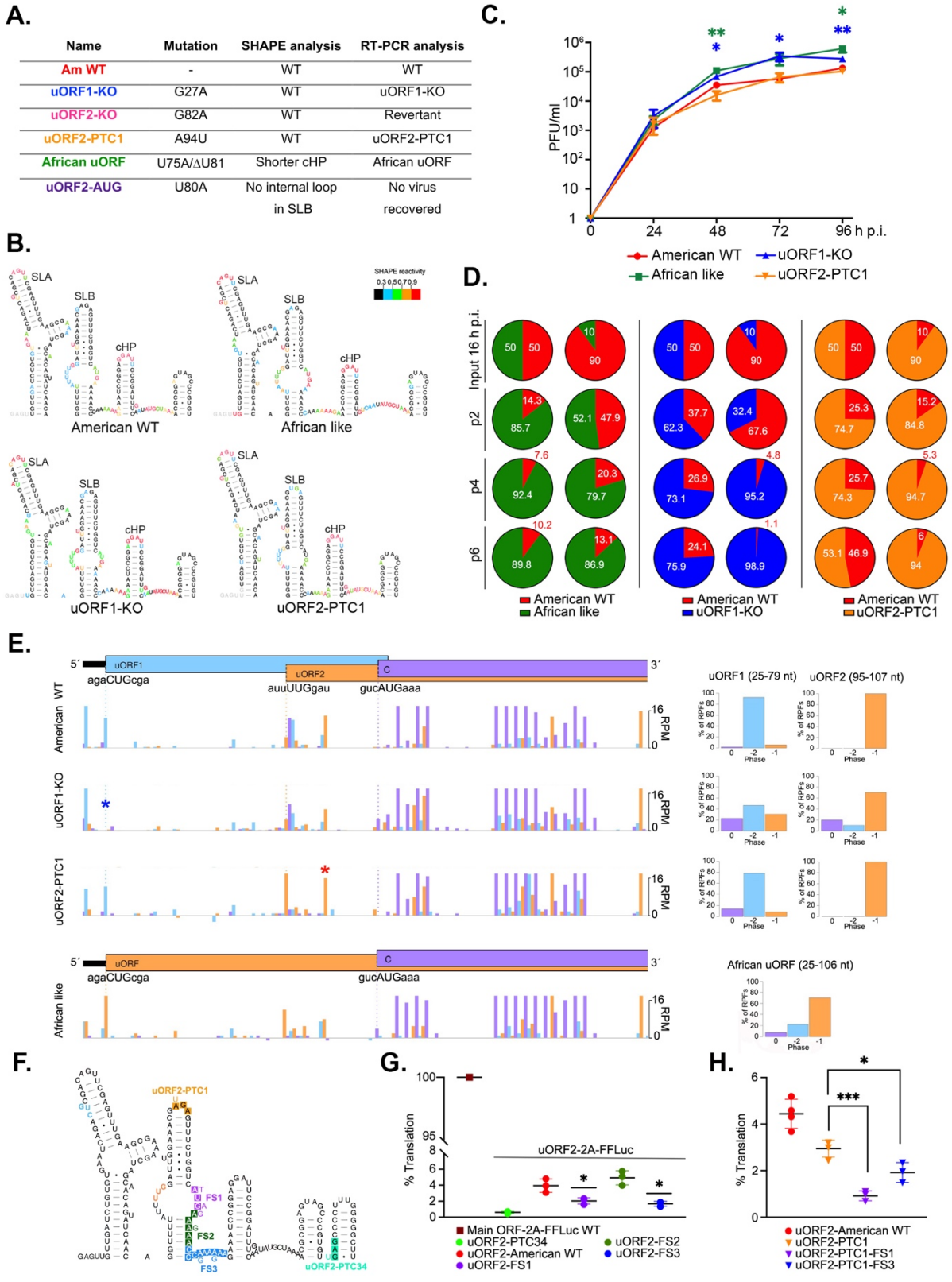


Figure 3. The significance of uORF translation in virus infection. (**A**) Summarising table of SHAPE and RT-PCR data for ZIKV uORFs mutant viruses derived from the pCCI-SP6-American ZIKV infectious clone. (**B**) SHAPE RNA secondary structure of the 5′ region (first 180 nucleotides) of the American WT, African-like, uORF1-KO and uORF2-PTC mutant viruses. Nucleotides are colour-coded based on SHAPE reactivity. SLA (stem-loop A), SLB

(stem-loop B) and cHP (capsid hairpin). (**C**) Time-course of infection of U251 cells with ZIKV mutant viruses (MOI 0.01) for 96 h. Plaque assays were performed on serial dilutions of the supernatant containing released virions. Values show the mean averages of the titration of three biological replicates. Error bars represent standard errors. PFU: plaque forming units. All *t*-tests were two-tailed and did not assume equal variance for the two populations being compared (*$p < 0.05$, ** $p < 0.01$). All *p*-values are from comparisons of the mutant virus with the American WT. (**D**) Pie charts of competition assays of American WT with different mutant viruses after 2 (p2), 4 (p4) and 6 (p6) sequential passages in U251 cells. Different proportions of each virus (50:50 and 90:10) were added as indicated in the first row (input 16 h p.i., passage 0). The chart area indicates the RNA proportion for each virus, as measured from sequencing chromatograms of RT-PCR products. Experiments were repeated independently eight times (raw data in **Supp Table 5**). (**E, left panels**) Ribo-Seq read density in the 5′ region of the ZIKV genome at 24 h p.i. of flash-frozen Vero cells infected with American WT, uORF1-KO, uORF2-PTC1 and African-like viruses (MOI:3). Histograms show the positions of the 5′ ends of reads with a +12 nt offset to map the approximate P-site as described in **Fig 1C-E**. Note that in order to properly visualise RPFs across the uORFs, the y-axis has been truncated at 20 reads per million virus mapping reads for the Ribo-Seq samples, leaving some RPF counts for the main ORF off-scale. The blue asterisk indicates the initiation codon of uORF1 and the red asterisk is the premature termination codon in uORF2. (**E, right panels**) Bar charts of the percentage of RPFs in each phase, relative to the main ORF (plot as described in **Fig 1F-H**). (**F**) Scheme of the 5′ region of the American WT. Premature termination codons (PTC) are colour-squared in orange and bright green for uORF2-PTC1 and uORF2-PTC34, respectively. The nucleotides corresponding to the main ORF initiation codon (FS1) and two potential slippery sequences (FS2 and FS3) are coloured in purple, green and blue, respectively. Mutated nucleotides are colour-coded and located adjacent to the original one. Relative FF-Luc activity of different mutants (PTC and FS) of uORF2-2A-FFLuc (**G**) and uORF2-PTC1-2A-FFLuc (**H**) in Vero-transfected cells as in **Fig 2B**. 100% translation accounted for the main ORF WT (brown square). All *t*-tests were two-tailed and did not assume equal variance for the two populations being compared (*$p < 0.05$, *** $p < 0.001$).

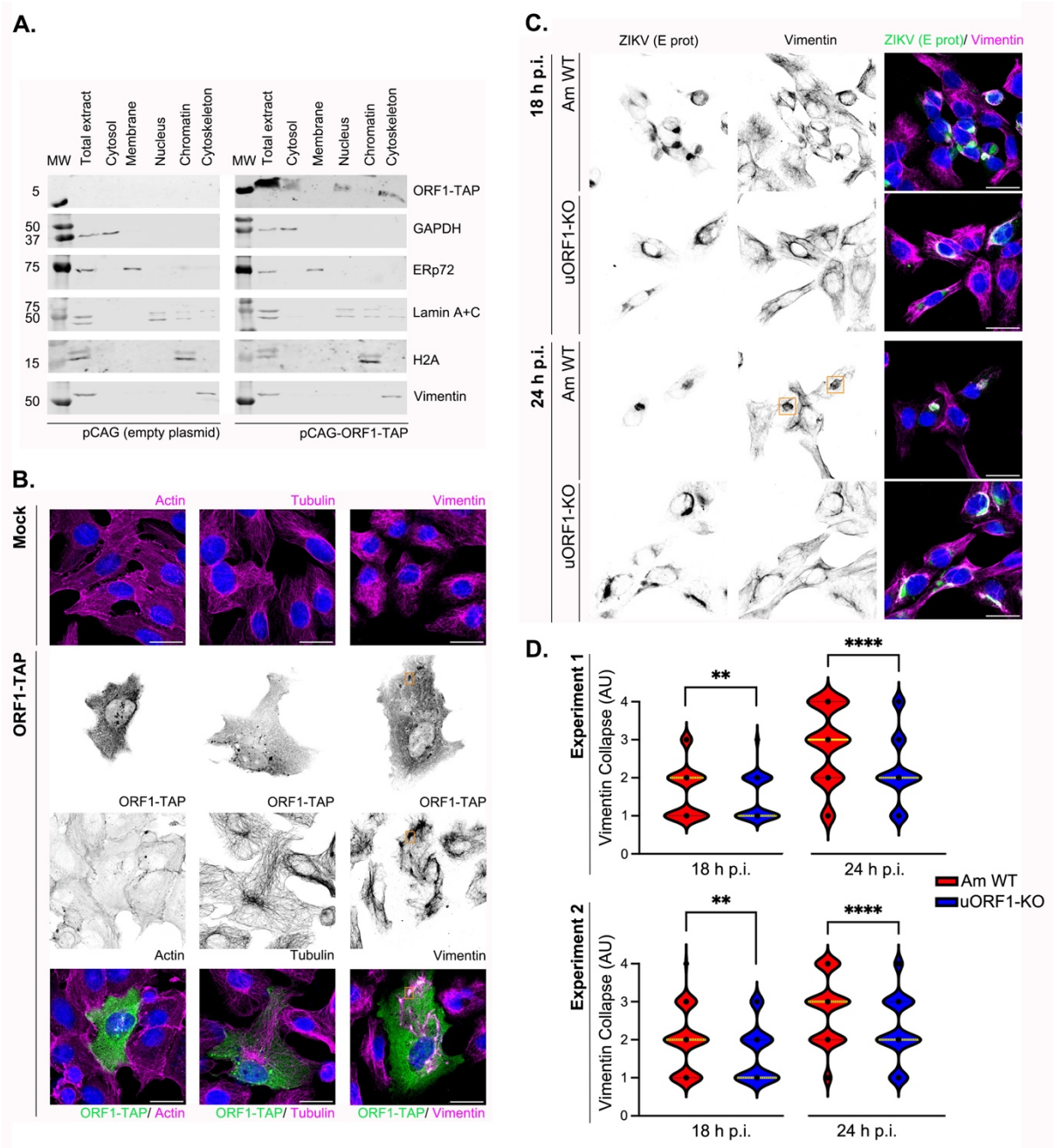


Figure 4. ZIKV uORF1 interacts with the cytoskeleton. (**A**) Subcellular fractionation analysis by Western blot of Vero cells transfected for 48 h with plasmid pCAG (left panel) or pCAG expressing the ZIKV uORF1 protein fused with TAP-tag at the C-terminus (pCAG-ORF1-TAP, right panel). The total extract, cytosolic, membrane, nuclear, chromatin and cytoskeletal fractions were probed with antibodies against FLAG (to detect TAP-tag); GAPDH (cytosolic marker); ERp72 (membrane marker); Lamin A+C (nuclear marker); H2A (chromatin marker); and vimentin (cytoskeletal marker). Molecular masses (kDa) are indicated on the left. (**B**) Representative confocal images of Vero cells transfected for 36 h with pCAG (mock) or the pCAG-ORF1-TAP plasmids. Cells were stained with antibodies against FLAG (green) and different cytoskeletal proteins (i.e. actin, tubulin and vimentin in magenta). Nuclei were counter-stained with DAPI (blue). Images are a maximum projection of a Z-stack. Scale bars, 25 µm. Experiments in **A** and **B** were repeated three times with similar results. (**C**) Representative confocal images of Vero cells infected with the American WT (Am WT) and

the uORF1-KO mutant virus (MOI:3) for 18h (upper panel) and 24 h (lower panel). Cells were stained with antibodies against the viral E protein (green) and vimentin (magenta). Nuclei were counter-stained with DAPI (blue). Images are a maximum projection of a Z-stack. Scale bars, 25 μm. (**D**) Violin plots represent vimentin collapse in Vero cells infected with the American WT and the uORF1-KO mutant virus (MOI:3) for 18 and 24 h. Vimentin collapse is quantified in immunofluorescence images as described in **C.** Vimentin collapse units are arbitrary units (AU) which are scored as indicated in the text. 110-170 cells per condition were scored (**Supp Table 6**). Experiments in **C** and **D** were independently repeated two times. All *t*-tests were two-tailed and did not assume equal variance for the two populations being compared (** $p < 0.01$, **** $p < 0.0001$).

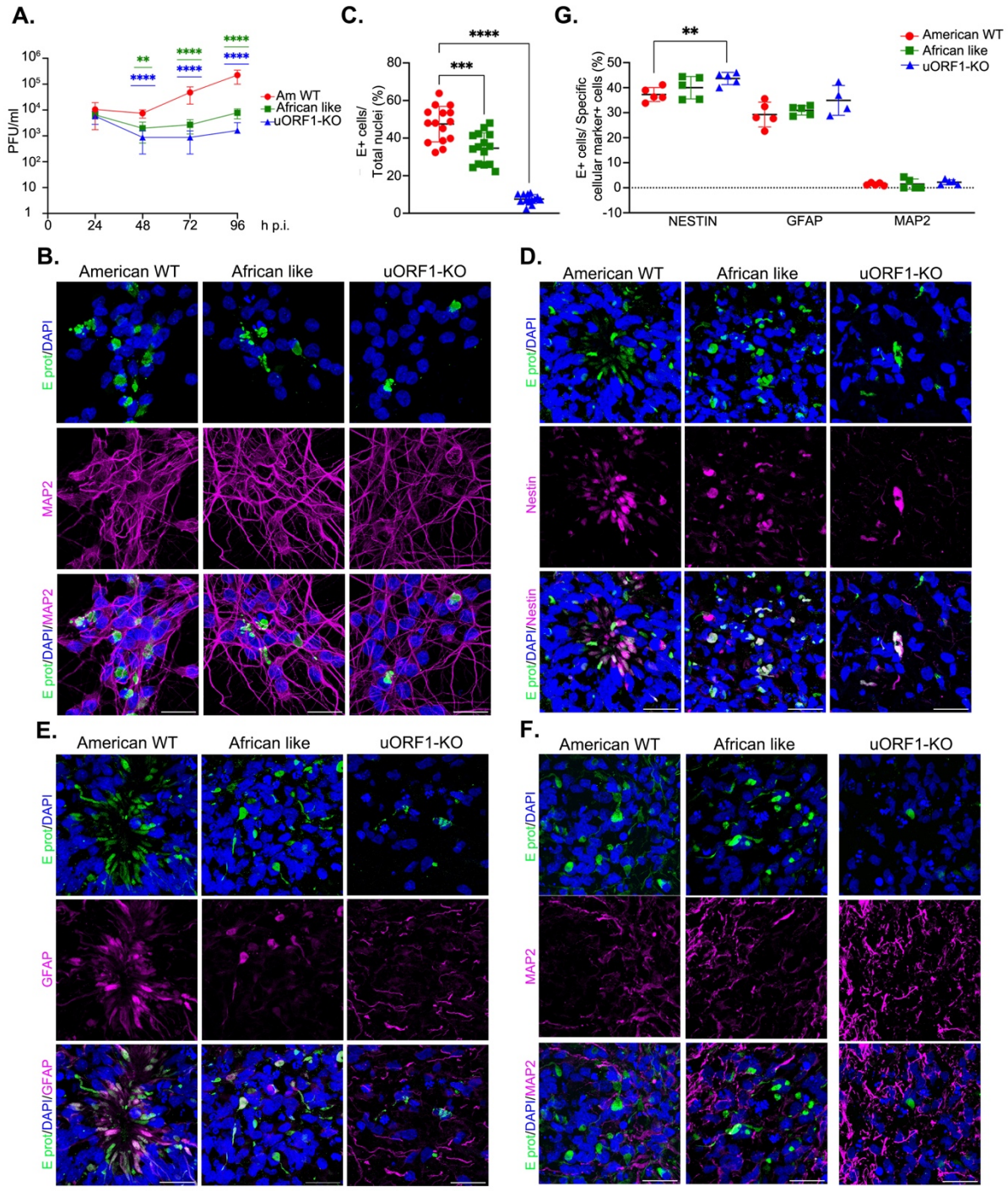


Figure 5. ZIKV uORFs are involved in neurotropism. (**A**) Time-course of i^{3}Neurons infected with ZIKV (MOI 10) for 96 h. $TCID_{50}$ were performed with serial dilutions of the supernatant to measure virion release. Values show the mean averages of the titration from five biological replicates. Error bars represent standard errors. PFU: plaque forming units. Statistical analysis was repeated measures two-way ANOVA on the log-transformed data (**$p < 0.01$, ****$p < 0.0001$). All p-values are from comparisons of the mutant virus with the American WT at that specific time-point. (**B**) Representative confocal images of i^{3}Neurons infected with the American WT, the African-like and the uORF1-KO viruses (MOI:10) for 96 h. Cells were stained with antibodies against the viral E protein (green) and the mature neuron marker MAP2 (magenta). Nuclei were counter-stained with DAPI (blue). Images represent the

maximum projection of a Z-stack. Scale bars, 25 μm. (**C**) Percentage of E^+ cells in relation to total number of nuclei. 15 images per virus type at 20X resolution (approx. 400-500 nuclei/image) were quantified for E-positive staining. Error bars represent standard errors. Statistical analysis was one-way ANOVA with Gaussian distribution and did not assume equal variance for the two populations being compared (** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$). All *p*-values are from comparisons of the mutant virus with the American WT. (**D-F**) Representative confocal images of ALI-COs infected with the American WT, the African-like and the uORF1-KO viruses (MOI:5) for 7 days, showing immunoreactivity for the viral E protein (green) and for different cellular markers (in magenta): nestin (**D**), GFAP (**E**) and MAP2 (**F**). Nuclei were counter-stained with DAPI (blue). Images represent the maximum projection of a Z-stack. Scale bars, 25 μm. (**G**) Percentage of E^+ cells that are positive for nestin, GFAP and MAP2. 5 images per virus type at 20X resolution (approx. 400-500 nuclei/image) were quantified. Statistical analysis was performed as in **C**.

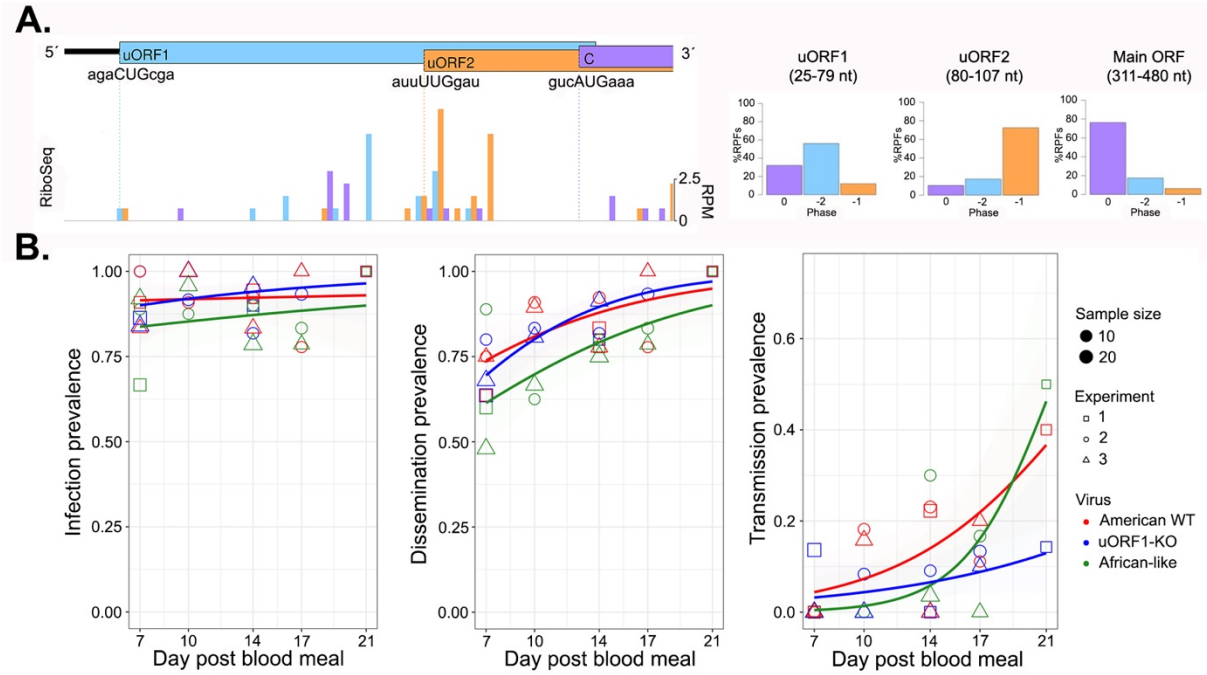


Figure 6. ZIKV mutant viruses display similar transmission dynamics in mosquitoes *in vivo*. (**A, left panel**) Ribo-Seq reads in the 5′ region of the ZIKV genome at 24 h p.i. of CHX-treated C6/36 cells infected with PE243 (MOI:3). Histograms show the positions of the 5′ ends of reads with a +12 nt offset to map the approximate P-site as described in **Fig 1C-E**. (**A, right panel**) Bar charts of the percentage of RPFs translated in each frame in relation to the main ORF as described in **Fig 1F-H**. (**B**) Prevalence of ZIKV infection (**left**), dissemination (**middle**) and transmission (**right**) over time in mosquitoes exposed to an infectious blood meal containing 2 x 10^6 PFU/mL of virus. Infection prevalence is the proportion of blood-fed mosquitoes with a body infection (determined by RT-PCR), dissemination prevalence is the proportion of infected mosquitoes with a virus-positive head (determined by RT-PCR), and transmission prevalence is the proportion of mosquitoes with a disseminated infection and infectious saliva (determined by focus-forming assay). The data represent three separate experiments depicted by different symbols, colour coded by different ZIKV mutants. The size of the symbols is proportionate to the number of mosquitoes tested, and the lines are the logistic fits of the time effect for each type of virus (ignoring the experiment effect on transmission prevalence in the visual representation).