UCLA

UCLA Previously Published Works

Title

High-Throughput Fitness Profiling of Zika Virus E Protein Reveals Different Roles for Glycosylation during Infection of Mammalian and Mosquito Cells

Permalink

Authors

Gong, Danyang
Zhang, Tian-Hao
Zhao, Dawei
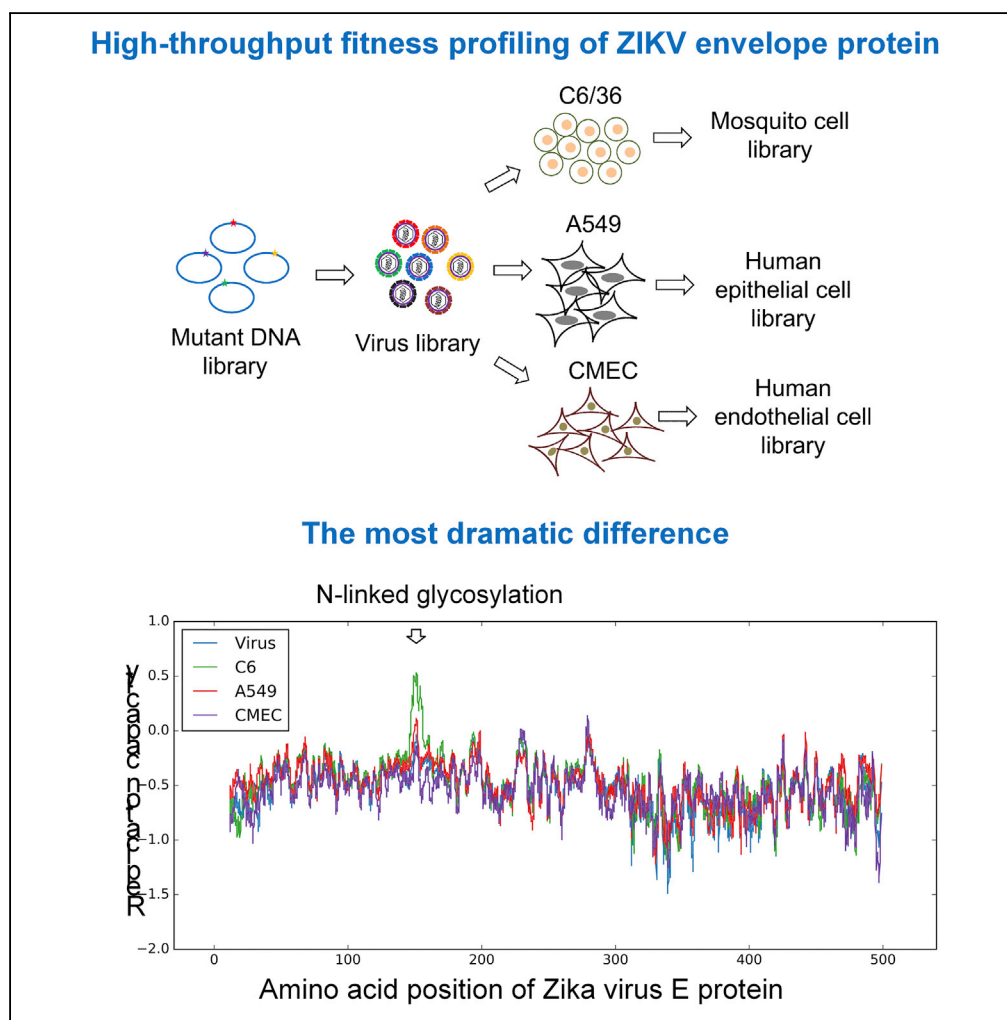et al.

Publication Date

DOI

Peer reviewed

Article

High-Throughput Fitness Profiling of Zika Virus E Protein Reveals Different Roles for Glycosylation during Infection of Mammalian and Mosquito Cells

Danyang Gong, Tian-Hao Zhang, Dawei Zhao, ..., Ting-Ting Wu, Vaithilingaraja Arumugaswami, Ren Sun

rsun@mednet.ucla.edu

HIGHLIGHTS
Gene-wide mapping of functional residues of E protein in human and mosquito cells

Mutations affecting N-linked glycosylation display the most dramatic difference

N-linked glycosylation decreases ZIKV entry into mosquito cells

N-linked glycosylation is important for DC-SIGN-mediated infection of human cells

Article

High-Throughput Fitness Profiling of Zika Virus E Protein Reveals Different Roles for Glycosylation during Infection of Mammalian and Mosquito Cells

Danyang Gong,[1] Tian-Hao Zhang,[1] Dawei Zhao,[1] Yushen Du,[1] Travis J. Chapa,[1] Yuan Shi,[1] Laurie Wang,[5] Deisy Contreras,[3] Gang Zeng,[2] Pei-Yong Shi,[4] Ting-Ting Wu,[1] Vaithilingaraja Arumugaswami,[3] and Ren Sun[1,6,*]

SUMMARY

Zika virus (ZIKV) infection causes Guillain-Barré syndrome and severe birth defects. ZIKV envelope (E) protein is the major viral protein involved in cell receptor binding and entry and is therefore considered one of the major determinants in ZIKV pathogenesis. Here we report a gene-wide mapping of functional residues of ZIKV E protein using a mutant library, with changes covering every nucleotide position. By comparing the replication fitness of every viral mutant between mosquito and human cells, we identified that mutations affecting glycosylation display the most divergence. By characterizing individual mutants, we show that ablation of glycosylation selectively benefits ZIKV infection of mosquito cells by enhancing cell entry, whereas it either has little impact on ZIKV infection on certain human cells or leads to decreased infection through the entry factor DC-SIGN. In conclusion, we define the roles of individual residues of ZIKV envelope protein, which contribute to ZIKV replication fitness in human and mosquito cells.

INTRODUCTION

Zika virus (ZIKV) is a mosquito-borne human pathogen of the *Flaviviridae* family, which also includes the dengue virus (DENV), West Nile virus (WNV), Japanese encephalitis virus (JEV), and yellow fever virus (YFV) (Lindenbach, 2007). ZIKV was first isolated from the serum of a sentinel rhesus monkey in the Zika forest of Uganda in 1947 and was subsequently recovered from the mosquito *Aedes africanus* in the same forest (Dick et al., 1952). ZIKV infection is mostly asymptomatic, but it can cause influenza-like symptoms, such as fever, headache, joint pain, and maculopapular rash (Simpson, 1964; Duffy et al., 2009). The recent outbreak of ZIKV in the Americas has demonstrated the potential for ZIKV to cause more serious disease, including microcephaly, other congenital malformations, and Guillain-Barré syndrome.

The ZIKV genome consists of a 10.8-kilobase single-stranded positive-sense RNA that codes for three structural proteins (capsid [C], membrane [prM/M], and envelope [E]) and seven non-structural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). In addition, there are short UTRs on both the 5′ and 3′ ends of the genome (Kuno and Chang, 2007). The mature ZIKV virion is roughly spherical and ∼50 nm in diameter. It contains a nucleocapsid that is surrounded by an icosahedral shell consisting of 180 copies of both E glycoprotein and M protein anchored in a lipid bilayer (Sirohi et al., 2016; Kostyuchenko et al., 2016). The flavivirus E protein, arranged as dimers on the surface of the mature virion, is the major viral protein involved in host-cell entry factor binding and fusion (Lindenbach, 2007). Each E protein monomer consists of four domains—three ectodomains (DI, DII, and DIII) and a transmembrane domain (TM). The structurally central DI acts as a bridge between DII and DIII and contains one N-linked glycosylation site (N154). The N-linked glycosylation at residue 153/154 of the E protein is conserved across most flaviviruses and has been shown to be important for optimal infection of mosquito and mammalian cells (Lee et al., 2010; Roehrig et al., 2007; Post et al., 1992; Heinz and Allison, 2003). DII includes the dimerization interface and a fusion loop that interacts with the endosomal membrane after conformation change. The IgG-like DIII is a continuous polypeptide segment and is thought to be important for binding to entry factors. Several host entry factors, including DC-SIGN, AXL, and TYRO3, have been shown to be important for mediating ZIKV infection (Hamel et al., 2015; Nowakowski et al., 2016). However, the detailed mechanism by which the E protein interacts with host-cell entry factors or the sequence determinants that contribute to human versus mosquito cell tropisms is not fully known.

[1]Department of Molecular and Medical Pharmacology, University of California, Los Angeles, CA 90095, USA

[2]Department of Urology, University of California, Los Angeles, CA 90095, USA

[3]Board of Governors Regenerative Medicine Institute, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA

[4]Department of Biochemistry & Molecular Biology, University of Texas Medical Branch, Galveston, TX 77555, USA

[5]Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA

[6]Lead Contact

*Correspondence: rsun@mednet.ucla.edu

https://doi.org/10.1016/j.isci.2018.02.005

We have developed a high-throughput fitness profiling approach that combines high-density mutagenesis with the power of next-generation sequencing to identify functional residues in the context of virus infection (Qi et al., 2014, 2017; Remenyi et al., 2014; Wu et al., 2016). In this study, we applied this approach to systematically analyze the functional residues of ZIKV E protein during infection of mosquito and human cells. We achieved high sensitivity in identifying residues essential for ZIKV E protein function. Surprisingly, we found that N-linked glycosylation at position N154 had differential effects on ZIKV infection between mosquito and human cells. Ablation of this glycosylation had little impact on viral infection of human cells (A549 and hCMEC), whereas it significantly increased infection of mosquito cells (C6/36), most probably by enhancing ZIKV entry. Last, N154 glycosylation was found to be important for ZIKV infection of mammalian cells through the entry factor DC-SIGN, further broadening our current knowledge concerning the glycosylation of E protein in the mammalian and invertebrate ZIKV life cycles.

RESULTS

Establishing Infectious cDNA Clone of ZIKV Strain PRVABC59

To facilitate mutational analysis of the ZIKV genome, we generated a plasmid (pZ-PR) carrying ZIKV cDNA, which was generated from an early passage of PRVABC59 virus. The ZIKV PRVABC59 strain was isolated from a human serum specimen from Puerto Rico in December 2015. The plasmid contains cDNA coding the entire ZIKV RNA genome, a T7 promoter on the 5′ end to drive ZIKV RNA *in vitro* transcription, and a BstBI site on the 3′ end to linearize the plasmid (Figures 1A and 1B, and Data S1). For reconstituting ZIKV virions, 5′-capped viral genomic RNA was generated by *in vitro* transcription (Figure 1C) and then electroporated into BHK21 cells. Culture supernatant containing recombinant ZIKV was harvested at 3 days post electroporation and further subjected to single-round amplification in C6/36 cells to generate the viral stock. Next, we compared recombinant and parental ZIKVs in cell culture and found that there was no obvious difference in either the percentages of cells positive for ZIKV infection or the signal intensity of E proteins (Figure 1D). In addition, they showed comparable growth kinetics in mammalian (Figure 1E) and mosquito cells (Figure 1F). These results indicate that the recombinant ZIKV can be reconstituted from the ZIKV cDNA plasmid and that it displays infectivity and growth comparable with that of the parental virus.

Constructing a Library of Mutations in ZIKV E Protein

After generating the ZIKV cDNA clone, we aimed to systematically analyze the role of each residue of the E protein using a high-throughput fitness profiling approach. In this study, error-prone PCR was employed to introduce random point mutations into the ZIKV E protein coding region (1,512 nt) by using an enzyme blend to minimize the mutation bias (Figure 2A). To better control the mutation rate of the error-prone PCR, the E protein was divided into three fragments, which together cover the whole sequence of the E protein. A mutant plasmid DNA library of ZIKV E protein was constructed by combining ∼40,000 clones from each fragment sub-library. To reconstitute the mutant virus library, the viral RNA library was generated from the plasmid DNA library by *in vitro* transcription and then electroporated into BHK-21 cells.

To evaluate the quality of the ZIKV E mutant libraries, both the plasmid DNA library and the reconstituted virus library were subjected to deep sequencing analysis. In addition, we sequenced a technical replicate for the DNA library, and triplicates for the virus library, to estimate the reproducibility of individual steps (Table S2). In the DNA library, 99.27% (4,503 of 4,536) of the expected mutations were identified by deep sequencing, of which 3,849 mutations showed an average relative frequency higher than 20 per million reads (Data S2). Here the relative frequency of each mutation was determined by dividing its mutation read count by the total population read count (see Transparent Methods). These 3,849 mutations covered all 1,512 nt positions of the ZIKV E protein and could be categorized as 2,729 missense (70.88%), 928 silent (24.11%), and 192 nonsense (4.99%) mutations. The 2,729 missense mutations introduced 2,494 unique amino acid changes and covered all residues of the ZIKV E protein. In addition, the mutations showed relatively even distributions across the E protein coding sequence (Figure 2C) and no strong bias on the nucleotide composition (Table S3). As the error-prone PCR enzyme blend introduces mutations randomly, the frequency of mutations in the plasmid DNA library should follow a Poisson distribution, where all mutation events are independent and random. Based on the deep sequencing result from the DNA plasmid library, each individual viral genome has 0.82 mutations on an average. About 44.2% of the genomes are expected to be wild type (WT), 36.1% of the genomes are expected to have a single mutation, and 14.7% of the genomes are expected to have two mutations. Only 5.0% of the genomes are expected to have more than two mutations. In the reconstituted virus library, 93.56% of the expected mutations were identified, with 2,860 mutations showing an average relative frequency greater than 20 per
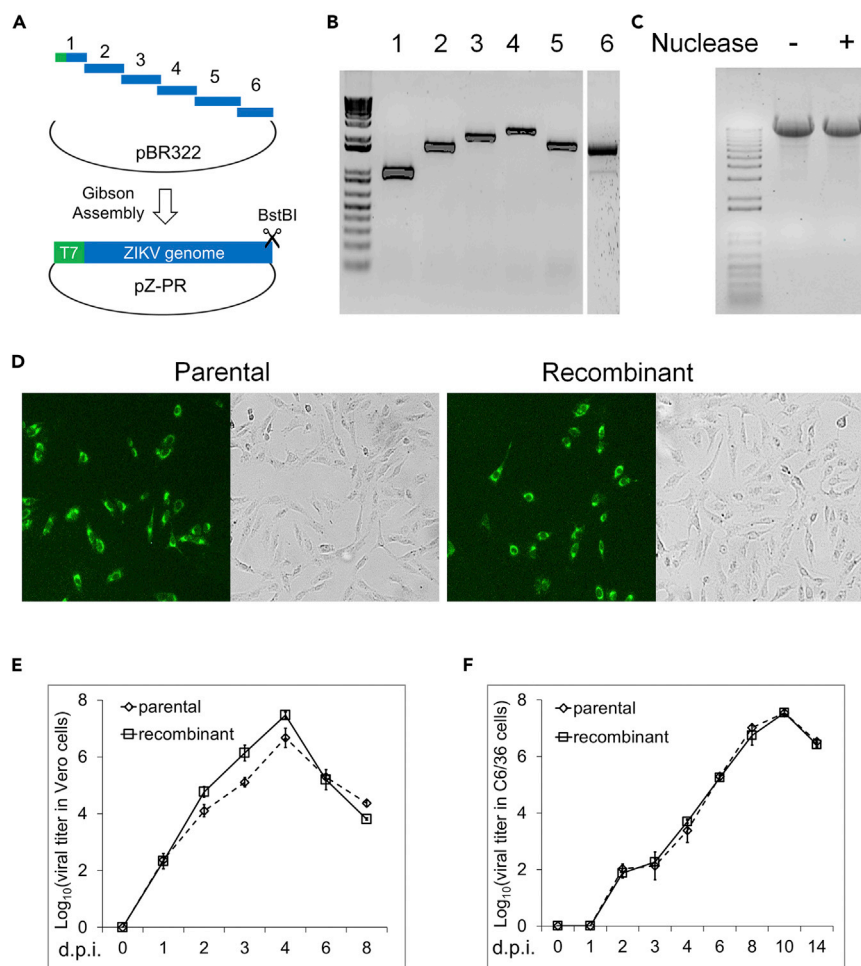
Figure 1. Construction of ZIKV Infectious cDNA Clone Using Strain PRVABC59

(A) Illustration of the cloning strategy.

(B) Analysis of cDNA fragments on an agarose gel. The first five fragments were amplified from cDNA synthesized with random hexamer, and the sixth fragment was generated from cDNA synthesized with specific oligo against the last 20 nt of ZIKV 3′ end (Table S1).

(C) Analysis of *in vitro* transcribed ZIKV RNA on a denaturing agarose gel. pZ-PR plasmids were linearized by BstBI, treated (or mock treated) with mung bean nuclease to remove the two extra nucleotides, *in vitro* transcribed, and then separated on a 0.8% formaldehyde denaturing gel.

(D) Immunostaining of ZIKV E proteins in parental or recombinant ZIKV-infected cells. Vero cells were infected at an MOI of 0.5, and after 16 hr, cells were fixed and subjected to IFA with αE (4G2) antibody.

(E and F) Multiple-step growth curve analysis of recombinant and parental ZIKVs. Vero (E) and C6/36 (F) cells were infected at an MOI of 0.01, and supernatants were harvested at the indicated time points post infection to titrate virion production. Viral titers with error bars are plotted as mean ± SD (n = 3).

See also Table S1 and Data S1.

million reads. Compared with the plasmid DNA library, these mutations included similar numbers of silent (908% and 31.75%), slightly fewer missense (1,845% and 64.51%), and significantly fewer nonsense (107% and 3.74%) mutations (Data S2). The 1,845 missense mutations introduced 1,753 unique amino acid changes and covered 503 of the 504 residues of the ZIKV E protein (except Gly5). These results clearly show that we have generated a high-quality plasmid DNA library and that we have reconstituted a virus library with high complexity. To evaluate the reproducibility of the experimental procedures, we examined the frequency distribution of all mutant types across replicate libraries; we observed similar distribution between both the DNA libraries and among the three virus libraries (Figure S1A). In addition, we compared the relative frequency of individual mutations between replicate libraries. Pearson's correlation coefficients
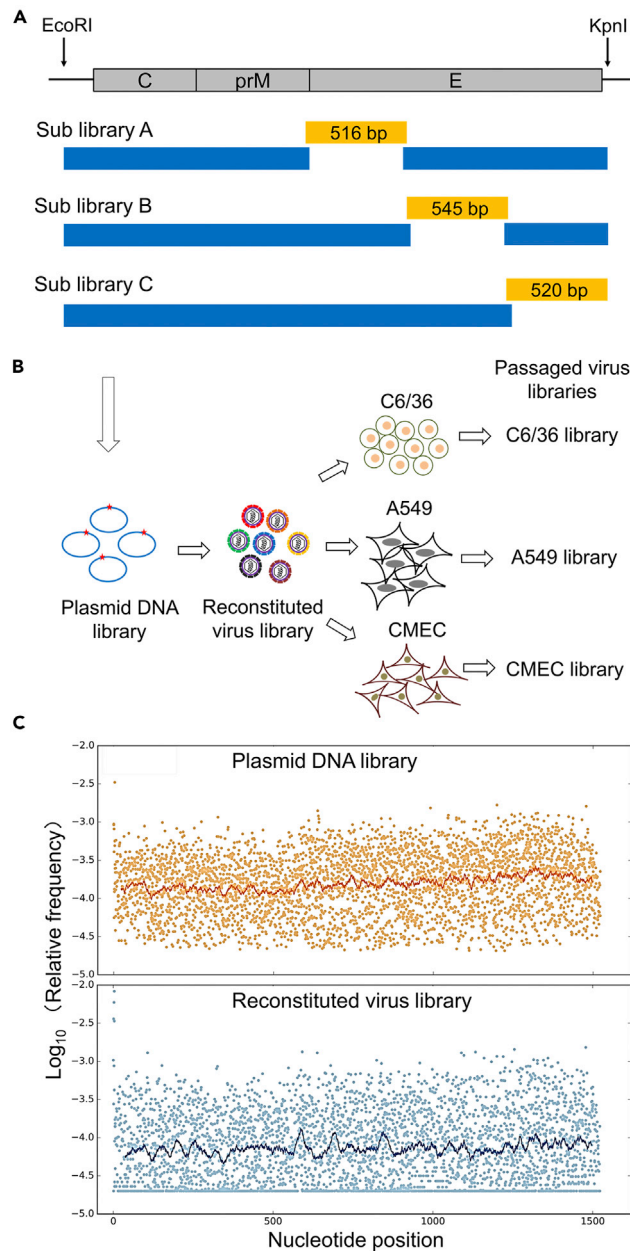
Figure 2. Construction and Characterization of ZIKV E Protein Mutational Library

(A) Construction of the plasmid DNA library. Three segments (yellow) covering the entire E coding region were generated by error-prone PCR, whereas the corresponding left and right fragments (blue) were generated by high-fidelity PCR. Fragments were further assembled and cloned into pZ-PR plasmids.

(B) A schematic representation of virus library construction and fitness profiling. Virus library was reconstituted by *in vitro* transcription from DNA library, followed by electroporation of viral RNA into BHK21 cells. Virus library was then passaged in C6/36, A549, or hCMEC cells. Deep sequencing of plasmid DNA library, reconstituted virus library, and passaged libraries was performed.

(C) Relative frequencies of individual point mutations are plotted across the E protein. Mutations were examined by deep sequencing of the ZIKV E gene, and the relative frequency of each mutation was calculated by dividing the mutation occurrence by the total population count (WT plus mutations). A smooth curve was fitted by cross-correlation and plotted. See also Figures S1 and S2, Tables S2 and S3, and Data S2.

of 0.97 for the technical replicates of the DNA library and 0.89 or 0.88 for the virus libraries (Figure S1B) were obtained. These strong correlations validated the reproducibility of our high-throughput fitness profiling platform.

To collect sufficient quantities of ZIKV mutant viruses for subsequent experiments, our virus library was harvested at 4 days after RNA electroporation. Trans-complementation after RNA transfection should enable the reconstitution of most mutant viruses. Owing to some viral amplification, more likely in the late period, there could be some selection effect on mutants. To evaluate the potential selection effect on the virus library, we calculated the relative fitness (RF) score (expressed as relative frequency$_{virus}$/relative frequency$_{DNA}$) for each mutant in the virus library. As shown in Figure S1C, the RF score of silent mutations (mean = 0.88) was higher than those of nonsense (mean = 0.06) and missense (mean = 0.21) mutations, suggesting that some fitness selection did occur during the virus reconstitution step. However, the high complexity of missense mutations in the virus library, and the depth of our sequencing, can support a high-resolution fitness profiling of ZIKV E protein in the three culture cell lines.

Fitness Profiling of E Protein Mutants in Mosquito and Human Cells

The ZIKV E protein mediates viral entry into host cells, defines tropism, and is an attractive target to counteract infection. As an arbovirus, the life cycle of ZIKV involves transmission between *Aedes* mosquitoes and human hosts (Ayres, 2016; Marchette et al., 1969). Therefore, it is important to identify host-specific functional residues of the E protein, which in turn can facilitate studies of the viral life cycle, especially regarding its entry mechanisms. In this study, one mosquito (C6/36) and two human (A549 and hCMEC/D3) cell lines were employed for the following reasons: (1) C6/36 cells were generated from the larvae of *Aedes albopictus* and have been widely used to study ZIKV and other flaviviruses; (2) epithelial cells are natural targets for ZIKV infection (Ma et al., 2016; Singh et al., 2017; Tang et al., 2016b), and as a human epithelial cell line, A549 expresses a majority of reported ZIKV entry factors (except DC-SIGN) (unpublished RNAseq result), supporting efficient infection of ZIKV; and (3) ZIKV targets neuronal cells by passing the human blood-brain barrier (Tang et al., 2016a; Nowakowski et al., 2016; Miner and Diamond, 2017), and barrier endothelial cells are permissive for ZIKV infection (Richard et al., 2017; Liu et al., 2016; Singh et al., 2017); thus, hCMEC/D3, which is a human blood-brain barrier endothelial cell line, was also chosen for our study. We first examined the fitness of E protein mutants by passaging the reconstituted virus library in these three cell lines, respectively (Figure 2B and Data S2), and noticed that the RF score distributions of silent mutations were clearly different from those of nonsense and missense mutations (Figure S2), which indicated a fitness selection in all three cell lines. It is notable that there are some nonsense mutations identified in the passaged virus libraries. To have rapid adaptation and evolution, RNA viruses have an average mutation rate (closely equivalent to mutational frequency) of 10^{-4}–10^{-5} mutations per nucleotide per round of RNA replication (Domingo et al., 1996; Drake et al., 1998; Meyerhans and Vartanian, 1999; Ramirez et al., 1995). In our passaged virus libraries from C6/36, A549, and CMEC cells, the mean frequency of nonsense mutations is 3.08×10^{-5}, 3.42×10^{-5}, and 3.23×10^{-5}, respectively, which is consistent with the spontaneous mutation rates that occur during genome replication. Therefore, we conclude that nonsense mutations identified in our virus libraries were most likely generated by spontaneous mutations instead of contamination by transfected RNA.

As a control for our fitness profile, we evaluated residues that have high variability in nature. We downloaded all ZIKV E protein sequences from Virus Pathogen Resource and extracted 52 unique complete sequences. After calculating the polymorphism score of each residue, we found that nine residues were highly diverse with a SNP score ≥ 100 (Data S3 and Figure 3A). The high diversity of these residues may suggest the ability of tolerating mutations. Indeed, the average RF scores of mutations on these nine residues were close to 1 in all three cell lines (Figure 3B). Moreover, we found that the frequently occurring natural variants on these nine residues did not cause or only slightly caused fitness loss (Figure 3C).

To guide our analysis of the fitness data, we focused on the pH-induced conformation change of E protein. Previous studies have shown that H146 and H323 in tick-borne encephalitis virus (equivalent to H144 and H323 in ZIKV) functioned as molecular switches (Figure 3D; Prakash et al., 2010; Fritz et al., 2008). Additionally, R9 in domain I (DI) and E373 (E377 in ZIKV) in domain III (DIII) form a salt bridge through their side chains in neutral pH, and that this interaction is essential for the dimeric conformation of E in mature virions (Figure 3D). After extensive conformation change triggered by low pH, a different salt bridge forms between E373 and H323, which may contribute to the stability of the post-fusion trimer (Bressanelli et al.,
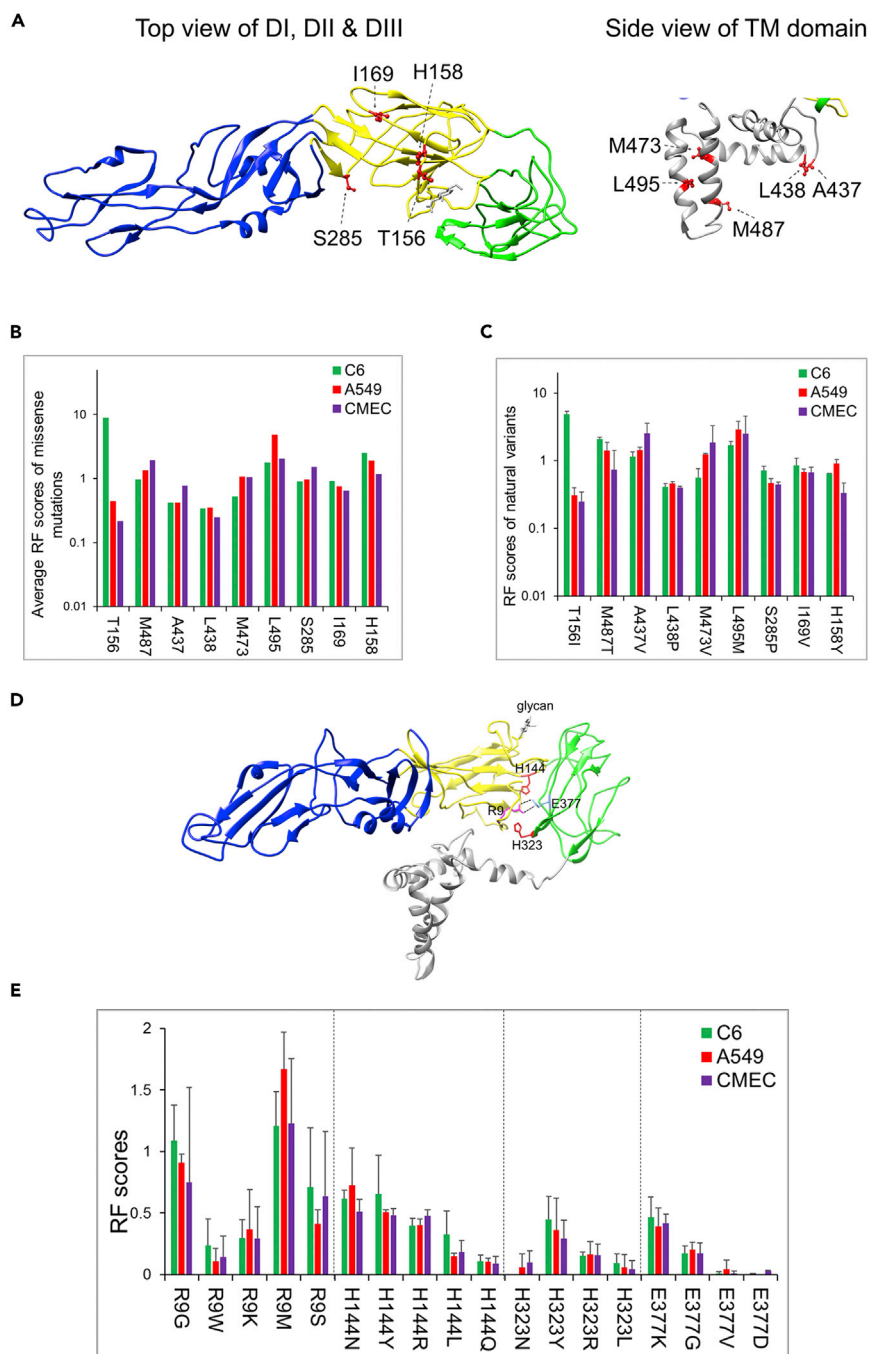
Figure 3. Analysis of Residues Highly Polymorphic or Involved in Conserved Function

(A) Ribbon diagrams of E protein monomer and localization of nine residues highly polymorphic in natural ZIKV sequence. Yellow, domain I; blue, domain II; green, domain III; gray, transmembrane domain.

(B) Average relative fitness (RF) score of mutations on residues. The RF score of mutation is expressed as the ratio of relative frequency in viral library to the relative frequency in DNA library. The average RF score of mutations on residue is defined as an average of RF scores of all missense mutations on this residue.
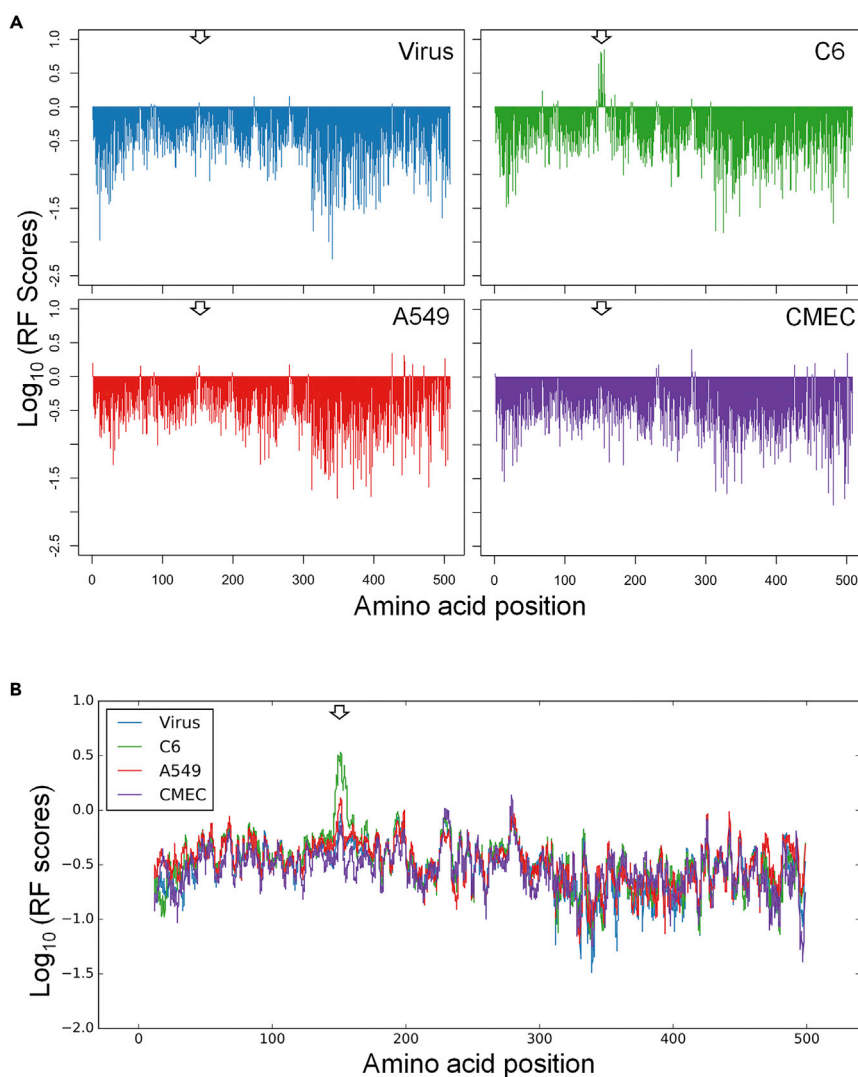
(C) RF scores of naturally occurring mutations. Data are plotted as mean + SD (n = 3).

(D) Localization of four conserved residues potentially functioning in pH-induced conformational change. A salt bridge between R9 and E377 is highlighted.

(E) RF scores of mutations on residues R9, H144, H323, and E377.

See also Data S3.

Figure 4. Fitness Profiling of ZIKV E Protein Mutant Virus Libraries

(A) Plot of averaged relative fitness (RF) scores of mutations at each residue position.

(B) Smoothed local average plot of mutants' RF scores. Smoothed local average was expressed as the rolling mean of the mutations' RF scores with a window size of 20. Arrow, N-linked glycosylation site (N154).

2004). After analyzing the fitness of mutations on these four residues, we found that (1) every mutation showed comparable fitness among the three cell lines; (2) all mutations on H323 and E377 showed RF scores lower than 0.5, and four specific mutations (H323N, H323L, E377V, and E377D) reduced the RF scores to < 0.1; and (3) in contrast, five mutations on R9 and H144 (R9G, R9M, R9S, H144N, and H144Y) showed RF scores greater than 0.5 and no mutation on these two residues reduced the RF scores below 0.1 (Figure 3E). These results suggest that H323 and E377 are important for the survival of ZIKV independent of the cell type, whereas R9 and H144 play relatively minor roles. In addition, the above findings validate the sensitivity of our high-throughput fitness profiling approach.

Next, we plotted the RF scores of mutations across the entire ZIKV E protein (Figure 4A) and found that the majority of mutations resulted in comparable fitness among the original reconstituted virus and the three passaged libraries. Interestingly, there was a cluster of beneficial mutations in C6/36 cells around the N-linked glycosylation site (N154). Smoothed local average analysis further confirmed that mutations in the glycosylation region presented the most dramatic difference in comparable fitness, wherein they showed significantly higher fitness in mosquito C6/36 cells but not in the two human cells tested (Figure 4B).
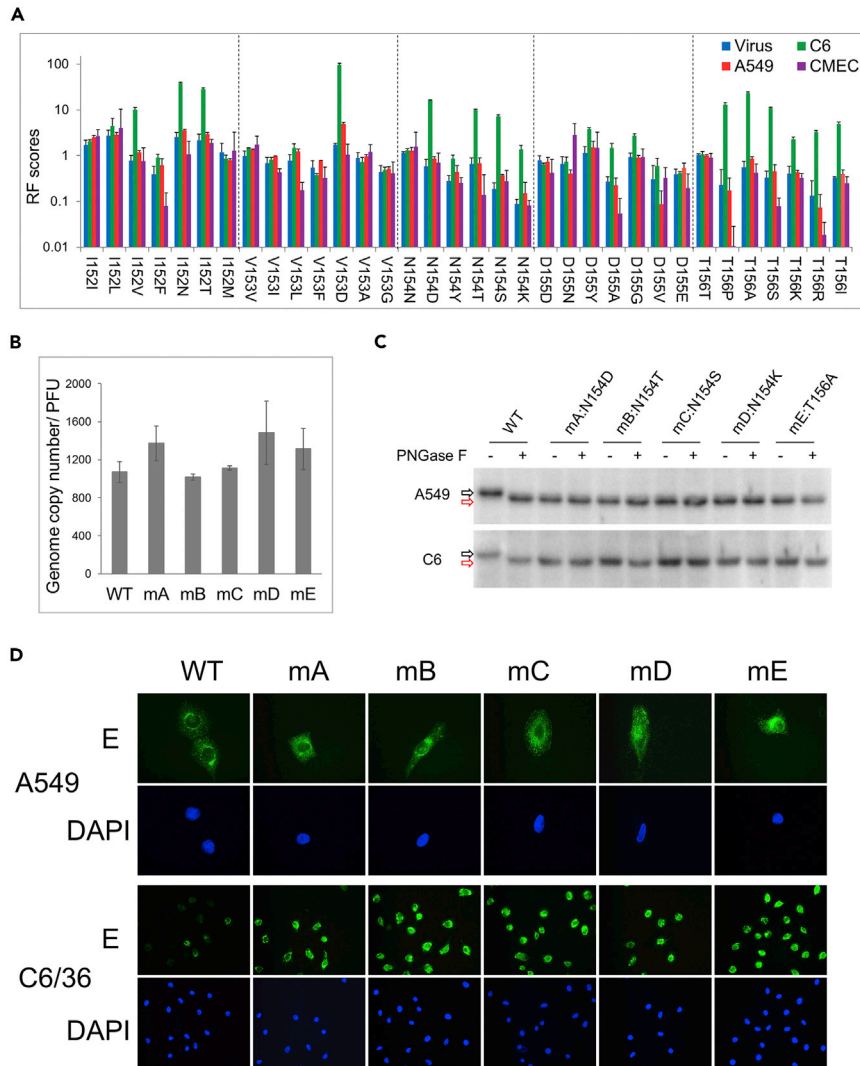
Figure 5. Analysis of Mutations in the N-Linked Glycosylation Region

(A) RF scores of mutations covering the five residues on and surrounding the N154 position.

(B) Infectivity of ZIKV WT and mutants. Titers of ZIKV WT and mutants were determined by plaque-forming assay on A549 cells, whereas genome copies of the same viruses were quantified by qRT-PCR. The ratio of viral genome copies to PFU was plotted as mean $\pm$ SD (n = 3).

(C) Analysis of E protein N-linked glycosylation using PNGase F treatment and Western blotting. Black arrow, N-linked glycosylation; red arrow, absence of N-linked glycosylation.

(D) Cells from same infection as (C) were also subjected to immunostaining using αE (4G2) antibody.

See also Figure S3.

Constructing and Characterizing Individual N-Linked-Glycosylation-Defective Mutants

Previous studies on the cryoelectron microscopy (cryo-EM) structure of ZIKV mature virion showed that the E protein was glycosylated at the N154 position (Sirohi et al., 2016; Kostyuchenko et al., 2016). This is consistent with the knowledge that N-linked glycosylation requires an N-X-S/T consensus sequence (Taylor and Drickamer, 2011), which is present as N154-D155-T156 in the ZIKV strain PRVABC59. In our deep sequencing analysis, many mutations at or around the N154 position were beneficial for ZIKV replication in mosquito C6/36 cells, likely due to the ablation of N-linked glycosylation. As shown in Figure 5A, all silent mutations (I152I, V153V, N154N, D155D, and T156T) did not obviously affect viral fitness, further validating our high-throughput screening approach. Eleven missense mutations were identified on two residues, N154 and T156, and in general, they selectively increased ZIKV fitness in C6/36 cells by 10-fold. In addition,

four missense mutations on I152 (I152V, I152N, and I152T) and V153 (V153D) caused a similar phenotype. Taken together, the profiling results suggest that glycosylation removal at N154 of ZIKV E protein might contribute to increased viral fitness in mosquito C6/36 cells.

To confirm the profiling results and to explore the mechanism of how N-linked glycosylation affects ZIKV fitness, five point mutations, four on N154 and one on T156, were individually introduced into the pZ-PR plasmid. ZIKV mutants were reconstituted in BHK21 cells and were further amplified in C6/36 cells to generate viral stocks. In plaque-forming assays on A549 cells, all mutants generated plaques of size comparable with that of WT ZIKV (data not shown). To quantify the infectivity of ZIKV WT and mutants, we calculated the particle-to-plaque-forming unit (PFU) ratios by utilizing genome copy number and the results of the plaque assay (Figure 5B). A ratio of 1.0×10^3 was obtained for the WT virus, and ratios between 1.0×10^3 and 1.4×10^3 were obtained for the mutants, suggesting comparable infectivity among WT and mutants.

To test whether these five mutations affected the N-linked glycosylation, we employed peptide-N-glycosidase F (PNGase F) treatment and Western blotting. As shown in Figure 5C, PNGase F treatment increased the mobility of WT E protein, suggesting the removal of N-linked glycosylation. All mutant E proteins displayed the same mobility as PNGase F-treated WT E protein and were not sensitive to PNGase F treatment, indicating that all five mutations abolished N-linked glycosylation. We further examined the E protein localization and expression by immunofluorescence assay (IFA) and found that E proteins from ZIKV WT and mutants showed similar localization and expression levels during infection of A549 cells (Figures 5D and S3). However, when infecting C6/36 cells, ZIKV mutants expressed more E proteins than the WT virus, which is consistent with the results from Western blotting (Figures 5C and S3).

Ablation of N-Linked Glycosylation Benefits ZIKV Replication in Mosquito Cells by Enhancing Virus Entry

To evaluate the roles of the N-linked glycosylation during ZIKV infection of mosquito and human cells, we characterized the growth kinetics of WT and mutant viruses by quantifying intracellular viral genome replication (Figures 6A and S4A) and extracellular infectious virion production (Figures 6B and S4B). The data showed that all viruses grew comparably in A549 cells. In contrast, the growth of mutant viruses was selectively enhanced in C6/36 cells. The genome copies of mutant viruses were ~10 times higher than that of the WT virus at days 1 and 2 post infection, leading to ~10 times more virion release at days 2 and 3 post infection. These results clearly demonstrate that ablation of N-linked glycosylation specifically benefited ZIKV growth in C6/36 cells, but not in A549 cells, validating our findings from the high-throughput fitness profiling analysis. Furthermore, we found that mutant viruses also showed significantly enhanced growth (Figure S5) in *Aedes aegypti* CCL-125 cells, indicating that this phenotype may be common to mosquito cell lines but not something specific to C6/36 cells. However, we also noticed that at late times post infection of C6/36 cells, WT and mutant viruses showed similar genome replication and virus production (day 4 in Figures 6A and 6B, days 4 and 6 in Figure S4). One possible reason for this observation may be that infected C6/36 cells do not undergo cell lysis; thus, once the monolayer is completely infected, fitness advantages become no longer apparent. At 4 and 6 days post infection, all C6/36 cells were positive for ZIKV infection, maintained similar levels of intracellular viral genomes, and released comparable amounts of virions. Therefore, a possible mechanism that explains why ZIKV mutants show higher fitness in C6/36 cells is that ablation of N-linked glycosylation might enhance the early events of virus infection, including attachment and entry.

To test this hypothesis, we first quantified the efficiency of ZIKV attachment and found that mutant viruses bound to the cell surface as efficiently as the WT virus did (Figures 6C and S6). Next, we examined the efficiency of viral genome delivery into the cytoplasm by quantifying cell-associated viral RNA. C6/36 cells were absorbed with WT or mutant viruses at 4°C and then returned to 28°C to initiate virus entry followed by viral protein expression and genome replication. Cells were collected at various times post infection to quantify the viral genome. Right after 4°C incubation and washing, cell-associated genome copies of WT and mutant viruses were similar (Figure 6D, 0 hr). At 0.5 hr after returning to 28°C, despite an overall drop of ~80%, the genome copies of all viruses were still comparable (Figure 6D, 0.5 hr). It is likely that ZIKV loses a large portion of virions in acidic endosomes, as was reported for DENV (van der Schaar et al., 2007), which would explain the drop of viral genome copies between 0 and 0.5 hr. Genome copies of the WT virus kept decreasing until 2 hr and then maintained a steady level until 8 hr, a level that may reflect the amount of viral

A

Genome replication

B

Virion production

C

D
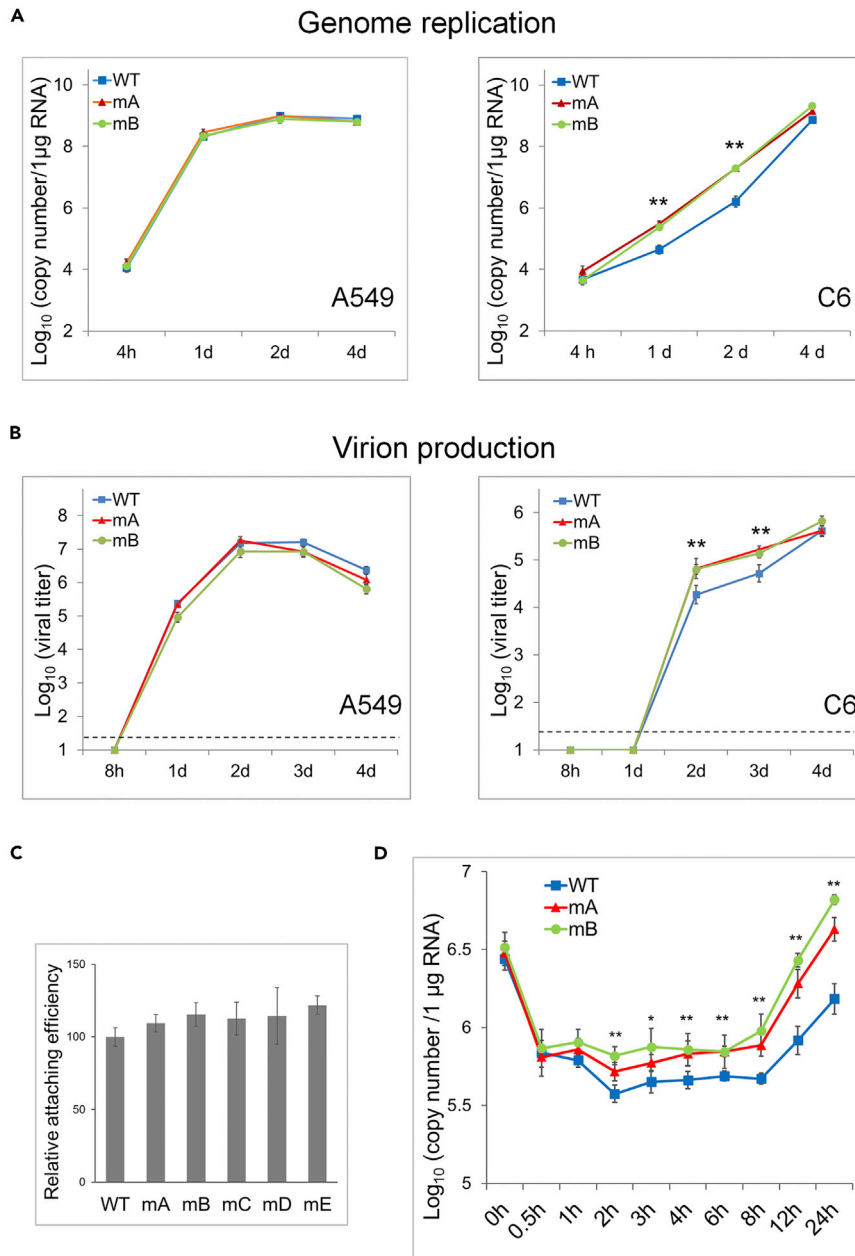
Figure 6. Characterization of ZIKV N-Linked-Glycosylation-Defective Mutants

(A) Intracellular genome replication of ZIKV WT and mutants. A549 or C6/36 cells were infected with ZIKV WT, mA, or mB at an MOI of 0.05 and harvested at indicated times for RNA purification and qRT-PCR analysis.

(B) Infectious virus production. Supernatants were harvested from same cells as (A) to determine viral titers.

(C) Attachment efficiency of ZIKV WT and mutants on surface of C6/36 cells. C6/36 were incubated with ZIKVs at an MOI of 10 for 1.5 hr at 4°C, followed by washing with ice-cold PBS three times, and then immediately subjected to RNA purification and qRT-PCR to determine viral genome copies. Relative attaching efficiency to the WT virus was plotted as mean $\pm$ SD (n = 3).

(D) Entry efficiency of ZIKV WT and mutants. C6/36 cells were incubated with ZIKVs as (C). After washing with ice-cold PBS, cells were replenished with warm culture medium before returning to 28°C, harvested at indicated times, and subjected to qRT-PCR analysis. *p<0.05; **p<0.01.

See also Figures S4, S5, and S6.

genome successfully delivered to the cytoplasm. In contrast, genome copies of mutant viruses did not obviously change between 0.5 and 8 hr. Starting from 1 hr, mutant viruses consistently showed higher genome copies than the WT virus, implying a higher efficiency of viral genome delivery. After 8 hr, genome copies of WT and mutant viruses robustly increased, indicating successful genome replication in each. Taken together, during ZIKV infection of mosquito cells, N-linked glycosylation seems to impair virus entry, resulting in less viral genome delivery into the cytoplasm than in non-glycosylated mutant viruses.

N-Linked Glycosylation is Important for ZIKV Infection of Mammalian Cells through DC-SIGN

The conserved N-D-T sequence motif was identified in the majority of ZIKV strains (Faye et al., 2014; Lanciotti et al., 2008), suggesting that N-linked glycosylation of E protein should play some role(s) in certain step(s) of the ZIKV life cycle, although we did not identify this using the above *in vitro* cell systems. DC-SIGN, also known as CD209, mediates DENV infection of dendritic cells by binding E protein glycan(s) (Navarro-Sanchez et al., 2003; Tassaneetrithep et al., 2003; Pokidysheva et al., 2006). Furthermore, expression of DC-SIGN makes cells susceptible to ZIKV infection (Hamel et al., 2015). Therefore, by using the mutant viruses, we sought to test whether the N-linked glycosylation of E protein was required for ZIKV infection of DC-SIGN-expressing cells. Because 293T cells are not susceptible to ZIKV infection owing to the lack of existing entry factor(s) (Hamel et al., 2015), we first established DC-SIGN-expressing cells by transfecting 293T cells with DC-SIGN expression plasmid. As shown in Figure S7, control 293T cells were largely non-permissive to ZIKV infection. The expression of DC-SIGN strongly enhanced WT ZIKV infection, resulting in ~10% cells positive for ZIKV E protein. Furthermore, we tested the infection efficiency of ZIKV WT and mutants by using the DC-SIGN-expressing 293T cells. The results showed that when compared with the WT virus, infection with mutant viruses resulted in significantly fewer E-protein-positive cells (Figure 7A), as well as an ~80% decrease of viral genome replication (Figure 7B). Moreover, similar results were also observed during ZIKV infection of human immature dendritic cells (Figures 7C and 7D). These observations suggest that N-linked glycosylation of E protein plays an important role in ZIKV infection through the entry factor DC-SIGN.

DISCUSSION

In this study, we generated an infectious cDNA clone of ZIKV strain PRVABC59 and used a high-throughput fitness profiling approach to assess the impact of mutations at every amino acid position of E protein on viral fitness. We found that nine residues showing high polymorphism in natural sequences could tolerate mutations, and we also evaluated the role of four conserved residues involved in E protein pH-induced conformation change. By comparing ZIKV replication fitness in mosquito and human cells, we discovered that mutations affecting N-linked glycosylation of E protein presented the most dramatic difference. These mutations either only slightly affected or did not affect ZIKV infection of human cells, whereas they significantly increased ZIKV growth in mosquito cells. In an attempt to explore the underlying mechanism, we discovered that ablation of N-linked glycosylation enhanced ZIKV entry into mosquito cells. Furthermore, N154 glycosylation was found to be important for ZIKV infection of mammalian cells through the entry factor DC-SIGN, indicating functional conservation among flaviviruses.

High-Throughput Fitness Profiling of ZIKV E Protein

Traditionally, mutagenesis at a specified position is the common approach for examining functional residues of a virus. The underlying principle of our high-throughput fitness profiling approach is to mutagenize the entire viral gene (or even genome) toward a single-nucleotide resolution, aiming to interrogate every position across the viral gene, passage the mutant library in a desired selection condition, and quantify the frequency change of each mutation to evaluate its phenotype. By using a mutant library of ZIKV E protein in which every individual amino acid is substituted with other residues, this study represents the first application of a high-throughput approach to analyze the functional residues of an entire ZIKV protein. In the future, this high-resolution profiling approach could be applied to other proteins to systematically understand ZIKV infection and its related pathogenesis. In addition, our library of high-density ZIKV E mutants could be applied to examine the roles of E protein in other aspects. In recent studies, mutant libraries of viral envelope proteins were used to characterize viral mutations escaping from monoclonal antibody neutralization and predict antibody-binding regions (Doud et al., 2017; Lin et al., 2012; Fulton et al., 2017). Through a similar strategy, our high-resolution profiling approach, facilitated by the library of ZIKV E mutants, could be applied to define protein epitope(s) for neutralizing antibody(s), including monoclonal antibodies and polyclonal antibodies in patient sera.
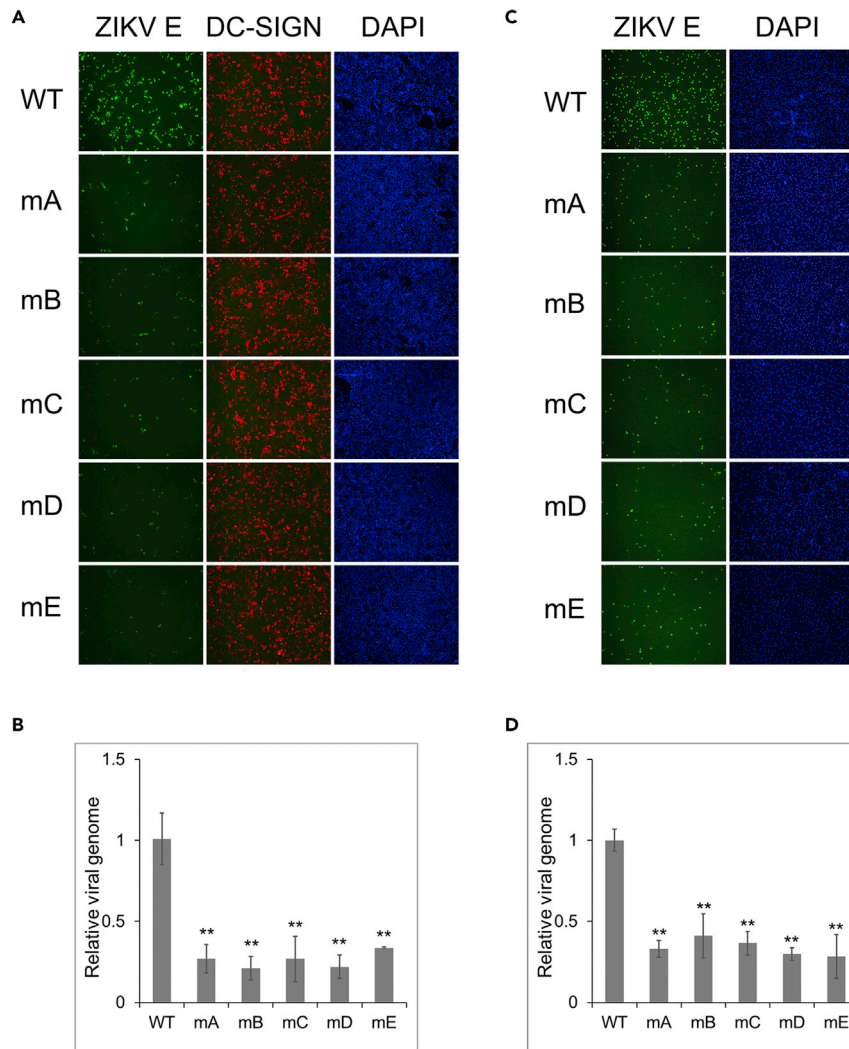
Figure 7. N-Linked Glycosylation Enhances ZIKV Infection through Entry Factor DC-SIGN

(A) Immunostaining of ZIKV E protein. DC-SIGN-cFLAG-transfected 293T cells were infected with ZIKV WT or mutants at an MOI of 0.2 for 16 hr and then fixed and subjected to IFA analysis with αE (4G2) and αFLAG antibodies.

(B) Total RNA was purified from the same cells as (A) and subjected to qRT-PCR analysis to quantify viral genome copies.

(C) Immunostaining of ZIKV E protein in immature dendritic cells infected with ZIKV WT or mutants.

(D) Quantification of viral genome copies in the infected immature dendritic cells. The relative viral genome copies to the WT virus was plotted as means $\pm$ SD (n = 3). *p<0.05; **p<0.01..

See also Figure S7.

N-Linked Glycosylation at Position N154 of ZIKV E Protein

Employing the high-throughput fitness profiling analysis, we are able to identify mutations that impair viral fitness in a specific cell background, which enables us to select informative mutants to dissect the multiple functions of a protein. In this study, by comparing ZIKV fitness in mosquito and human cells, we discovered that mutations affecting the N-linked glycosylation of E protein presented the most dramatic difference. In studies of other flaviviruses, N-linked glycosylation at position N153/154 of the E protein was shown to be important for viral growth in both mammalian and mosquito cells (Hanna et al., 2005; Mondotte et al., 2007; Lee et al., 2010; Beasley et al., 2005; Scherret et al., 2001). For example, ablation of N153 glycosylation of the E protein reduced DENV production over 10-fold in cultured cells, including mosquito cells (C6/36), mammalian epithelial cells (Vero), and fibroblast cells (BHK21) (Mondotte et al., 2007; Lee et al., 2010). Similarly, N154 glycosylation was required for the efficient growth of WNV in multiple cell lines, including C6/36, Vero, and DF-1 (Beasley et al., 2005; Scherret et al., 2001; Moudy et al., 2009). Surprisingly, N-linked

glycosylation at position N154 of E protein was not necessary for ZIKV replication in either human epithelial cells (A549) or endothelial cells (hCMEC). In addition, ablation of N-linked glycosylation significantly enhanced ZIKV growth in C6/36 cells (Figures 4, 6A and 6B). Further exploring the mechanism, we found that N-linked glycosylation of ZIKV E protein did not affect the initial attachment of virions to the cell surface (Figure 6C) but did notably impair viral entry into C6/36 cells (Figure 6D). In a study of WNV using subviral particles (SVPs), SVPs lacking glycosylation of E protein were found to be modestly more efficient in entry on BHK-21 and QT6 cells, whereas this absence enhanced the infection of C6/36 cells. However, the latter experiments showed that this increase of entry did not lead to more viral production of non-glycosylated mutants, probably owing to deficiency in virion assembly and egress (Hanna et al., 2005).

The initial step in the life cycle of flaviviruses is attachment of the virion to host-cell entry factor(s), followed by entry into cells via endocytosis, through which the virus particles are transported to endosomes. The acidic environment in the endosome lumen promotes rearrangement and conformational change of E protein, resulting in the fusion of the virion envelope with the endosomal membrane, followed by release of the viral genome into the cytoplasm (Heinz et al., 2004; van der Schaar et al., 2007; Zhang et al., 2004). Based on cryo-EM structures of ZIKV, the glycans at the N154 position of the E protein lie over the dimer interface in close proximity to the fusion loop of domain II (Sirohi et al., 2016; Kostyuchenko et al., 2016). Considering their substantial size, glycans might generate steric effects impairing ZIKV E protein conformational change in endosome and/or the membrane fusion process between virion and endosome of mosquito cells, which can explain why ablation of N-linked glycosylation enhanced ZIKV entry in C6/36 cells (Figure 7D). However, we also found that ablation of N-linked glycosylation did not enhance ZIKV infection of human A549 and hCMEC cells, implying that it plays different role(s) in entry of mosquito and human cells. In future research, the mechanisms underlying the impairment of N-linked glycosylation of E protein on ZIKV entry in mosquito cells and the different roles of glycosylation among flaviviruses need to be investigated.

Two sequences in the ZIKV PRVABC59 strain, N154-D155-T156 and N481-G482-S483, match the N-linked glycosylation consensus element (N-X-S/T). However, the residue N481 is embedded in the lipid bilayer of ZIKV envelope and is therefore unlikely to be glycosylated. The N154-D155-T156 sequence exists in many but not all ZIKV strains (Faye et al., 2014; May and Relich, 2016). E protein variants defective in N-linked glycosylation, especially a Thr to Ile (T156I) change, were frequently observed among ZIKV strains. It is suggested that acquisition of the N154 glycosylation site is a recurrent event in the history of ZIKV (Faye et al., 2014). However, this hypothesis could be limited, as some loss-of-glycosylation variants might be a consequence of passaging ZIKV strains before sequencing. In our study of the ZIKV Puerto Rico strain, mutations of residue N154 did not obviously change viral growth in two human cells, significantly enhanced viral fitness in mosquito cells (Figures 5 and 6), and impaired virus infection through the DC-SIGN entry factor (Figure 7). However, in ZIKV African isolates, substitution or deletion of N154 is rarely observed. In addition, we noticed that, when compared with the corresponding silent mutations, two of five mutations (N154Y and N154K) of N154 did not obviously increase viral fitness in mosquito cells, whereas all six mutations of T156 did so significantly (Figure 5A). All the above information suggests that residue N154 may also serve other function(s), independent of N-linked glycosylation, in ZIKV replication. In addition, it also validates the advantage of the high-resolution fitness profiling platform in unbiasedly analyzing functional residues.

It has been shown that DC-SIGN mediates flavivirus infection of human dendritic cells by binding E protein glycan(s) (Hamel et al., 2015; Tassaneetrithep et al., 2003), and dendritic cells were important targets for flavivirus infection in patients (Wu et al., 2000; Marovich et al., 2001). In addition, N-linked glycosylation of the E protein of WNV could possibly enhance its neuroinvasion (Beasley et al., 2005; Shirato et al., 2004). In our study, N-linked glycosylation of the E protein was shown to be important for ZIKV infection of immature dendritic cells and also 293T cells expressing DC-SIGN (Figure 7). The above information suggests that glycosylation of ZIKV E protein may contribute to its virulence in patients and mouse models. However, all the data in our study were collected from *in vitro* infection of human and mosquito cell lines, and future detailed investigation with mouse models and mosquito vectors will be required to examine the potential connection between E protein N-linked glycosylation and ZIKV virulence, particularly neuroinvasion.

In summary, this is the first study to systematically map functional residues of the ZIKV E protein and characterize the roles of its N-linked glycosylation at the N154 position. The study validates the advantage of a

high-resolution fitness profiling platform in unbiasedly analyzing functional residues. Future research will be required to elucidate the molecular mechanism of the viral entry affected by N-linked glycosylation of ZIKV E protein, in cell culture and in animal models. Our library of ZIKV E mutants could be employed to examine the roles of E protein in other conditions, for example, mapping neutralizing antibody-binding epitopes and interactions with entry factor(s). We also expect that the high-throughput fitness profiling approach will be applied to study other ZIKV viral proteins.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND SOFTWARE AVAILABILITY

Sequence of the ZIKV infectious cDNA clone (pZ-PR) is available from NCBI GenBank: KY583506. Raw deep sequencing data is available from NIH Short Read Archive: PRJNA373927.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Transparent Methods, seven figures, three tables, and three data files and can be found with this article online at https://doi.org/10.1016/j.isci.2018.02.005.

AUTHOR CONTRIBUTIONS

All experiments were performed by D.G. with the help of D.Z. and L.W. D.G. and R.S. designed the project. D.G., T.-H.Z., and Y.D. analyzed the data. Y.S., T.J.C, D.C., G.Z., P.-Y.S., T.-T.W., and V.A. provided reagents and advices. D.G. wrote the manuscript. T.J.C, L.W., and R.S. revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Ayres, C.F. (2016). Identification of Zika virus vectors and implications for control. Lancet Infect. Dis. *16*, 278–279.

Beasley, D.W., Whiteman, M.C., Zhang, S., Huang, C.Y., Schneider, B.S., Smith, D.R., Gromowski, G.D., Higgs, S., Kinney, R.M., and Barrett, A.D. (2005). Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. J. Virol. *79*, 8339–8347.

Bressanelli, S., Stiasny, K., Allison, S.L., Stura, E.A., Duquerroy, S., Lescar, J., Heinz, F.X., and Rey, F.A. (2004). Structure of a flavivirus envelope glycoprotein in its low-pH-induced membrane fusion conformation. EMBO J. *23*, 728–738.

Dick, G.W., Kitchen, S.F., and Haddow, A.J. (1952). Zika virus. I. Isolations and serological specificity. Trans. R. Soc. Trop. Med. Hyg. *46*, 509–520.

Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S.F., Quer, J., Novella, I.S., and Holland,

J.J. (1996). Basic concepts in RNA virus evolution. FASEB J. *10*, 859–864.

Doud, M.B., Hensley, S.E., and Bloom, J.D. (2017). Complete mapping of viral escape from neutralizing antibodies. PLoS Pathogen *13*, e1006271.

Drake, J.W., Charlesworth, B., Charlesworth, D., and Crow, J.F. (1998). Rates of spontaneous mutation. Genetics *148*, 1667–1686.

Duffy, M.R., Chen, T.H., Hancock, W.T., Powers, A.M., Kool, J.L., Lanciotti, R.S., Pretrick, M., Marfel, M., Holzbauer, S., Dubray, C., et al. (2009). Zika virus outbreak on Yap Island, Federated States of Micronesia. N. Engl. J. Med. *360*, 2536–2543.

Faye, O., Freire, C.C., Iamarino, A., Faye, O., de Oliveira, J.V., Diallo, M., Zanotto, P.M., and Sall, A.A. (2014). Molecular evolution of Zika virus during its emergence in the 20(th) century. PLoS Negl. Trop. Dis. *8*, e2636.

Fritz, R., Stiasny, K., and Heinz, F.X. (2008). Identification of specific histidines as pH sensors in flavivirus membrane fusion. J. Cell Biol. *183*, 353–361.

Fulton, B.O., Sachs, D., Schwarz, M.C., Palese, P., and Evans, M.J. (2017). Transposon mutagenesis of the Zika virus genome highlights regions essential for RNA replication and restricted for immune evasion. J. Virol. *91*, e00698–e00717.

Hamel, R., Dejarnac, O., Wichit, S., Ekchariyawat, P., Neyret, A., Luplertlop, N., Perera-Lecoin, M., Surasombatpattana, P., Talignani, L., Thomas, F., et al. (2015). Biology of Zika virus infection in human skin cells. J. Virol. *89*, 8880–8896.

Hanna, S.L., Pierson, T.C., Sanchez, M.D., Ahmed, A.A., Murtadha, M.M., and Doms, R.W. (2005). N-linked glycosylation of West Nile virus envelope proteins influences particle assembly and infectivity. J. Virol. *79*, 13262–13274.

Heinz, F.X., and Allison, S.L. (2003). Flavivirus structure and membrane fusion. Adv. Virus Res. *59*, 63–97.

Heinz, F.X., Stiasny, K., and Allison, S.L. (2004). The entry machinery of flaviviruses. Arch. Virol. Suppl. 133–137.

Kostyuchenko, V.A., Lim, E.X., Zhang, S., Fibriansah, G., Ng, T.S., Ooi, J.S., Shi, J., and Lok, S.M. (2016). Structure of the thermally stable Zika virus. Nature 533, 425–428.

Kuno, G., and Chang, G.J. (2007). Full-length sequencing and genomic characterization of Bagaza, Kedougou, and Zika viruses. Arch. Virol. 152, 687–696.

Lanciotti, R.S., Kosoy, O.L., Laven, J.J., Velez, J.O., Lambert, A.J., Johnson, A.J., Stanfield, S.M., and Duffy, M.R. (2008). Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. Emerg. Infect. Dis. 14, 1232–1239.

Lee, E., Leang, S.K., Davidson, A., and Lobigs, M. (2010). Both E protein glycans adversely affect dengue virus infectivity but are beneficial for virion release. J. Virol. 84, 5171–5180.

Lin, T.Y., Dowd, K.A., Manhart, C.J., Nelson, S., Whitehead, S.S., and Pierson, T.C. (2012). A novel approach for the rapid mutagenesis and directed evolution of the structural genes of West Nile virus. J. Virol. 86, 3501–3512.

Lindenbach, B.D., Thiel, H.J., and Rice, C.M. (2007). Flaviviridae: the viruses and their replication. In Fields Virology, Fifth edition, D.M. Knipe and P.M. Howely, eds. (Lippincott-Raven), pp. 1101–1152.

Liu, S.F., Delalio, L.J., Isakson, B.E., and Wang, T.T. (2016). AXL-mediated productive infection of human endothelial cells by Zika virus. Circ. Res. 119, 1183–1189.

Ma, W.Q., Li, S.H., Ma, S.Q., Jia, L.N., Zhang, F.C., Zhang, Y., Zhang, J.Y., Wong, G., Zhang, S.S., Lu, X.C., et al. (2016). Zika virus causes testis damage and leads to male infertility in mice. Cell 167, 1511–1524.e10.

Marchette, N.J., Garcia, R., and Rudnick, A. (1969). Isolation of Zika virus from Aedes aegypti mosquitoes in Malaysia. Am. J. Trop. Med. Hyg. 18, 411–415.

Marovich, M., Grouard-Vogel, G., Louder, M., Eller, M., Sun, W., Wu, S.J., Putvatana, R., Murphy, G., Tassaneetrithep, B., Burgess, T., et al. (2001). Human dendritic cells as targets of dengue virus infection. J. Investig. Dermatol. Symp. Proc. 6, 219–224.

May, M., and Relich, R.F. (2016). A comprehensive systems biology approach to studying Zika virus. PLos One 11, e0161355.

Meyerhans, A., and Vartanian, J.P. (1999). The fidelity of cellular and viral polymerases and its manipulation for hypermutagenesis. In Origin and Evolution of Viruses, R.W. Domingo and J. Holland, eds. (Academic Press, New York), pp. 87–114.

Miner, J.J., and Diamond, M.S. (2017). Zika virus pathogenesis and tissue tropism. Cell Host Microbe 21, 134–142.

Mondotte, J.A., Lozach, P.Y., Amara, A., and Gamarnik, A.V. (2007). Essential role of dengue virus envelope protein N-glycosylation at asparagine-67 during viral propagation. J. Virol. 81, 7136–7148.

Moudy, R.M., Zhang, B., Shi, P.Y., and Kramer, L.D. (2009). West Nile virus envelope protein glycosylation is required for efficient viral transmission by Culex vectors. Virology 387, 222–228.

Navarro-Sanchez, E., Altmeyer, R., Amara, A., Schwartz, O., Fieschi, F., Virelizier, J.L., Arenzana-Seisdedos, F., and Despres, P. (2003). Dendritic-cell-specific ICAM3-grabbing non-integrin is essential for the productive infection of human dendritic cells by mosquito-cell-derived dengue viruses. EMBO Rep. 4, 723–728.

Nowakowski, T.J., Pollen, A.A., Di Lullo, E., Sandoval-Espinosa, C., Bershteyn, M., and Kriegstein, A.R. (2016). Expression analysis highlights AXL as a candidate Zika virus entry receptor in neural stem cells. Cell Stem Cell 18, 591–596.

Pokidysheva, E., Zhang, Y., Battisti, A.J., Bator-Kelly, C.M., Chipman, P.R., Xiao, C., Gregorio, G.G., Hendrickson, W.A., Kuhn, R.J., and Rossmann, M.G. (2006). Cryo-EM reconstruction of dengue virus in complex with the carbohydrate recognition domain of DC-SIGN. Cell 124, 485–493.

Post, P.R., Santos, C.N., Carvalho, R., Cruz, A.C., Rice, C.M., and Galler, R. (1992). Heterogeneity in envelope protein-sequence and N-linked glycosylation among yellow-fever virus-vaccine strains. Virology 188, 160–167.

Prakash, M.K., Barducci, A., and Parrinello, M. (2010). Probing the mechanism of pH-induced large-scale conformational changes in dengue virus envelope protein using atomistic simulations. Biophys. J. 99, 588–594.

Qi, H., Olson, C.A., Wu, N.C., Ke, R., Loverdo, C., Chu, V., Truong, S., Remenyi, R., Chen, Z., Du, Y., et al. (2014). A quantitative high-resolution genetic profile rapidly identifies sequence determinants of hepatitis C viral fitness and drug sensitivity. PLoS Pathog. 10, e1004064.

Qi, H., Chu, V., Wu, N.C., Chen, Z., Truong, S., Brar, G., Su, S.Y., Du, Y., Arumugaswami, V., Olson, C.A., et al. (2017). Systematic identification of anti-interferon function on hepatitis C virus genome reveals p7 as an immune evasion protein. Proc. Natl. Acad. Sci. USA 114, 2018–2023.

Ramirez, B.C., Barbier, P., Seron, K., Haenni, A.L., and Bernardi, F. (1995). Molecular mechanisms of point mutations in RNA viruses. In Molecular Basis of Virus Evolution, A.J. Gibbs, C.H. Calisher, and F. García-Arenal, eds. (Cambridge University Press), pp. 105–118.

Remenyi, R., Qi, H., Su, S.Y., Chen, Z., Wu, N.C., Arumugaswami, V., Truong, S., Chu, V., Stokelman, T., Lo, H.H., et al. (2014). A comprehensive functional map of the hepatitis C virus genome provides a resource for probing viral proteins. MBio 5, e01469–14.

Richard, A.S., Shim, B.S., Kwon, Y.C., Zhang, R., Otsuka, Y., Schmitt, K., Berri, F., Diamond, M.S., and Choe, H. (2017). AXL-dependent infection of human fetal endothelial cells distinguishes Zika virus from other pathogenic flaviviruses. Proc. Natl. Acad. Sci. USA 114, 2024–2029.

Roehrig, J.T., Bryant, J.E., Calvert, A.E., Mesesan, K., Crabtree, M.B., Volpe, K.E., Silengo, S., Kinney, R.M., Huang, C.Y., and

Miller, B.R. (2007). Glycosylation of the dengue 2 virus E protein at N67 is critical for virus growth in vitro but not for growth in intrathoracically-inoculated Aedes aegypti mosquitoes. Am. J. Trop. Med. Hyg. 77, 1.

van der Schaar, H.M., Rust, M.J., Waarts, B.L., van der Ende-Metselaar, H., Kuhn, R.J., Wilschut, J., Zhuang, X., and Smit, J.M. (2007). Characterization of the early events in dengue virus cell entry by biochemical assays and single-virus tracking. J. Virol. 81, 12019–12028.

Scherret, J.H., Mackenzie, J.S., Khromykh, A.A., and Hall, R.A. (2001). Biological significance of glycosylation of the envelope protein of Kunjin virus. Ann. N. Y. Acad. Sci. 951, 361–363.

Shirato, K., Miyoshi, H., Goto, A., Ako, Y., Ueki, T., Kariwa, H., and Takashima, I. (2004). Viral envelope protein glycosylation is a molecular determinant of the neuroinvasiveness of the New York strain of West Nile virus. J. Gen. Virol. 85, 3637–3645.

Simpson, D.I. (1964). Zika virus infection in man. Trans. R. Soc. Trop. Med. Hyg. 58, 335–338.

Singh, P.K., Guest, J.M., Kanwar, M., Boss, J., Gao, N., Juzych, M.S., Abrams, G.W., Yu, F.S., and Kumar, A. (2017). Zika virus infects cells lining the blood-retinal barrier and causes chorioretinal atrophy in mouse eyes. JCI Insight 2, e92340.

Sirohi, D., Chen, Z., Sun, L., Klose, T., Pierson, T.C., Rossmann, M.G., and Kuhn, R.J. (2016). The 3.8 A resolution cryo-EM structure of Zika virus. Science 352, 467–470.

Tang, H.L., Hammack, C., Ogden, S.C., Wen, Z.X., Qian, X.Y., Li, Y.J., Yao, B., Shin, J., Zhang, F.R., Lee, E.M., et al. (2016a). Zika virus infects human cortical neural progenitors and attenuates their growth. Cell Stem Cell 18, 587–590.

Tang, W.W., Young, M.P., Mamidi, A., Regla-Nava, J.A., Kim, K., and Shresta, S. (2016b). A mouse model of Zika virus sexual transmission and vaginal viral replication. Cell Rep. 17, 3091–3098.

Tassaneetrithep, B., Burgess, T.H., Granelli-Piperno, A., Trumpfherer, C., Finke, J., Sun, W., Eller, M.A., Pattanapanyasat, K., Sarasombath, S., Birx, D.L., et al. (2003). DC-SIGN (CD209) mediates dengue virus infection of human dendritic cells. J. Exp. Med. 197, 823–829.

Taylor, M.E., and Drickamer, K. (2011). Introduction to Glycobiology (Oxford University Press).

Wu, S.J., Grouard-Vogel, G., Sun, W., Mascola, J.R., Brachtel, E., Putvatana, R., Louder, M.K., Filgueira, L., Marovich, M.A., Wong, H.K., et al. (2000). Human skin Langerhans cells are targets of dengue virus infection. Nat. Med. 6, 816–820.

Wu, N.C., Du, Y.S., Le, S., Young, A.P., Zhang, T.H., Wang, Y.Y., Zhou, J., Yoshizawa, J.M., Dong, L., Li, X.M., et al. (2016). Coupling high-throughput genetics with phylogenetic information reveals an epistatic interaction on the influenza A virus M segment. BMC Genomics 17, 46.

Zhang, Y., Zhang, W., Ogata, S., Clements, D., Strauss, J.H., Baker, T.S., Kuhn, R.J., and Rossmann, M.G. (2004). Conformational changes of the flavivirus E glycoprotein. Structure 12, 1607–1618.

Supplemental Information

High-Throughput Fitness Profiling of Zika Virus E

Protein Reveals Different Roles for Glycosylation

during Infection of Mammalian and Mosquito Cells

Danyang Gong, Tian-Hao Zhang, Dawei Zhao, Yushen Du, Travis J. Chapa, Yuan Shi, Laurie Wang, Deisy Contreras, Gang Zeng, Pei-Yong Shi, Ting-Ting Wu, Vaithilingaraja Arumugaswami, and Ren Sun
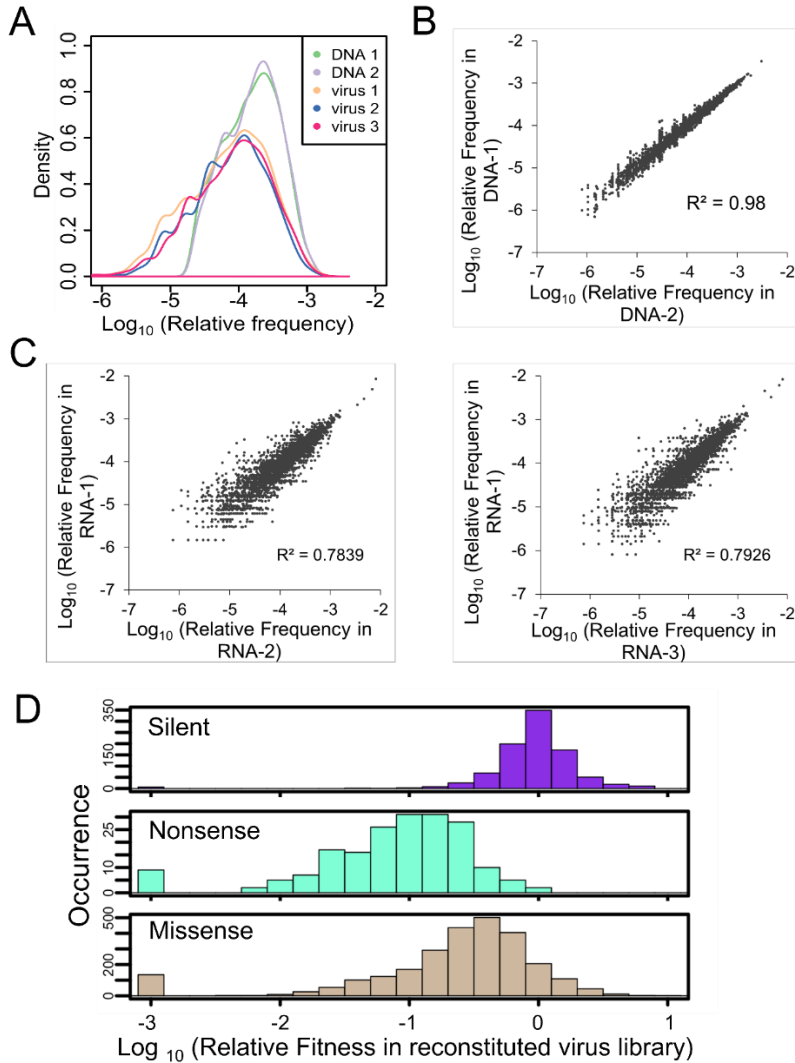
Fig. S1.

Fig. S1. Related to Fig. 2. Analysis of deep sequencing results of plasmid DNA and reconstituted virus libraries. (A) Mutation frequency distribution of DNA and virus libraries. Frequency distributions were smoothed by Gaussian kernel density estimation. Two replicates of DNA libraries are shown as orange polygons. Three replicates of input virus libraries are shown as blue polygons. (B) Correlations of relative frequency of individual point mutations between samples. (C) Log_{10} relative fitness indices for silent mutations, nonsense mutations, and missense mutations in virus library are shown as histograms.
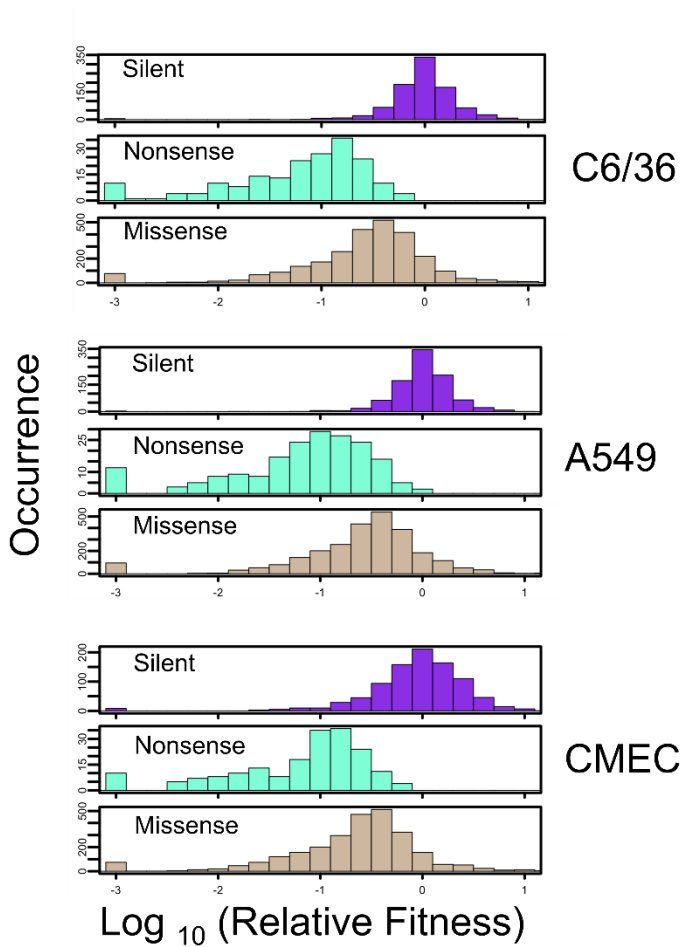
Fig. S2

Fig. S2. Related to Fig. 2. Log$_{10}$ relative fitness indices for silent mutations, nonsense mutations, and missense mutations in passaged libraries are shown as histograms.
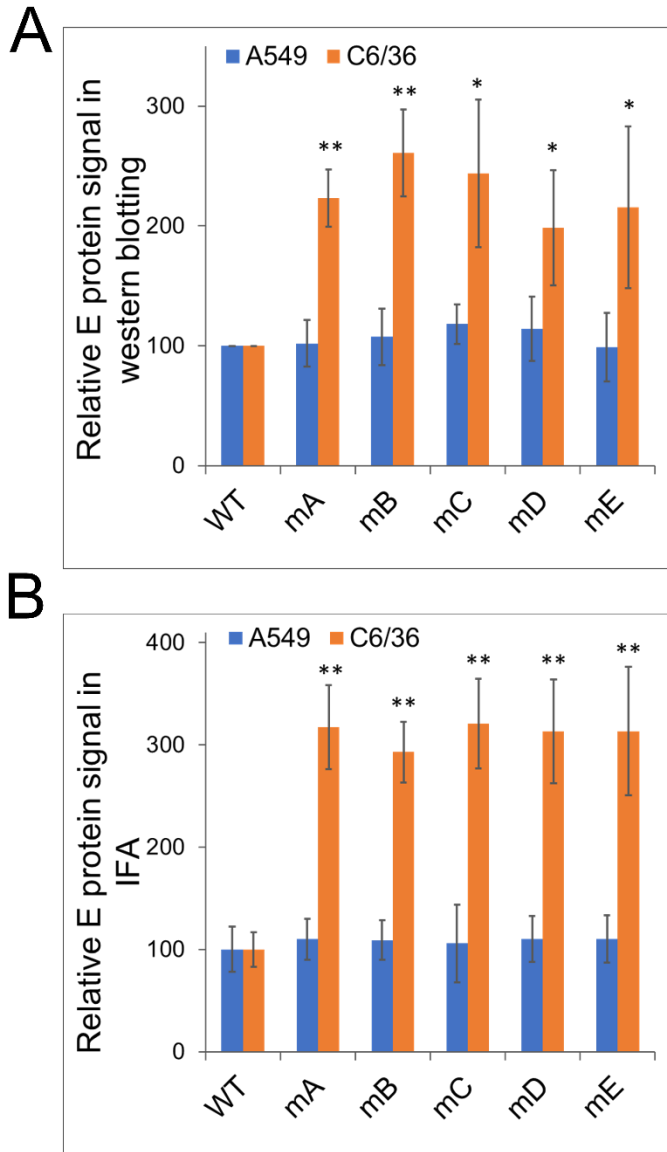
Fig. S3

Fig. S3. Related to Fig. 5. Western blotting signal intensity (A) and fluorescence (B) and was quantified with ImageJ, normalized to corresponding WT virus and plotted. Results were analyzed by unpaired Student's t test. Differences between WT and mutants were considered statistically significant when $p<0.05$ (*) or $p<0.01$ (**).

Fig. S4
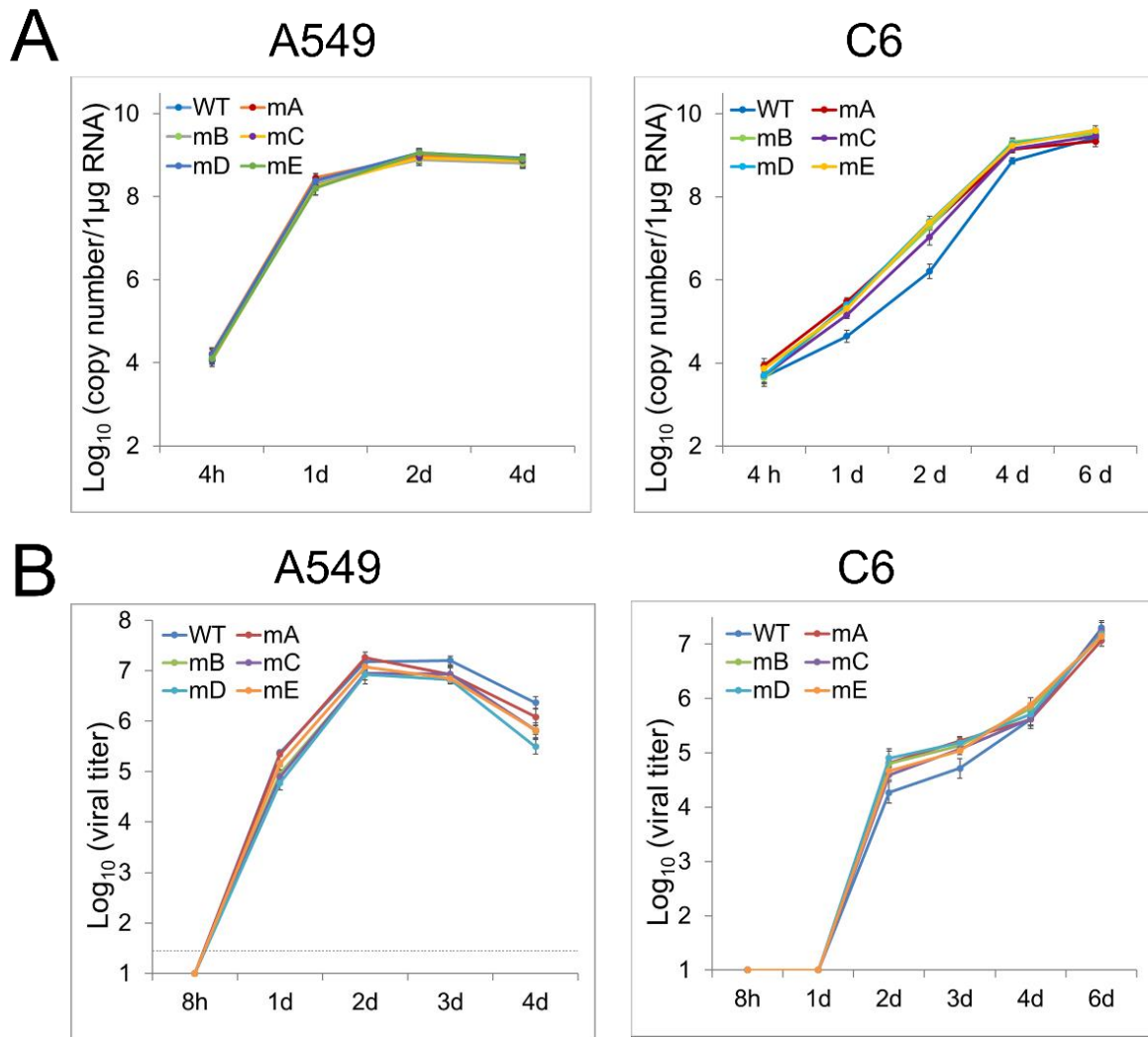
Fig. S4. Related to Fig. 6. Multiple-step growth curve analysis of ZIKV WT and five N-linked glycosylation defective mutants. (A) Intracellular genome replication of ZIKV WT and mutants. A549 or C6/36 cells were infected with ZIKV WT or mutants at a MOI of 0.05, total RNA was purified from cells at the indicated time points post-infection, and subjected to qRT-PCR analysis. (B) Infectious virus production. Supernatants were harvested from same cells as (A), and viral titers were determined by plaque assay.

Fig. S5
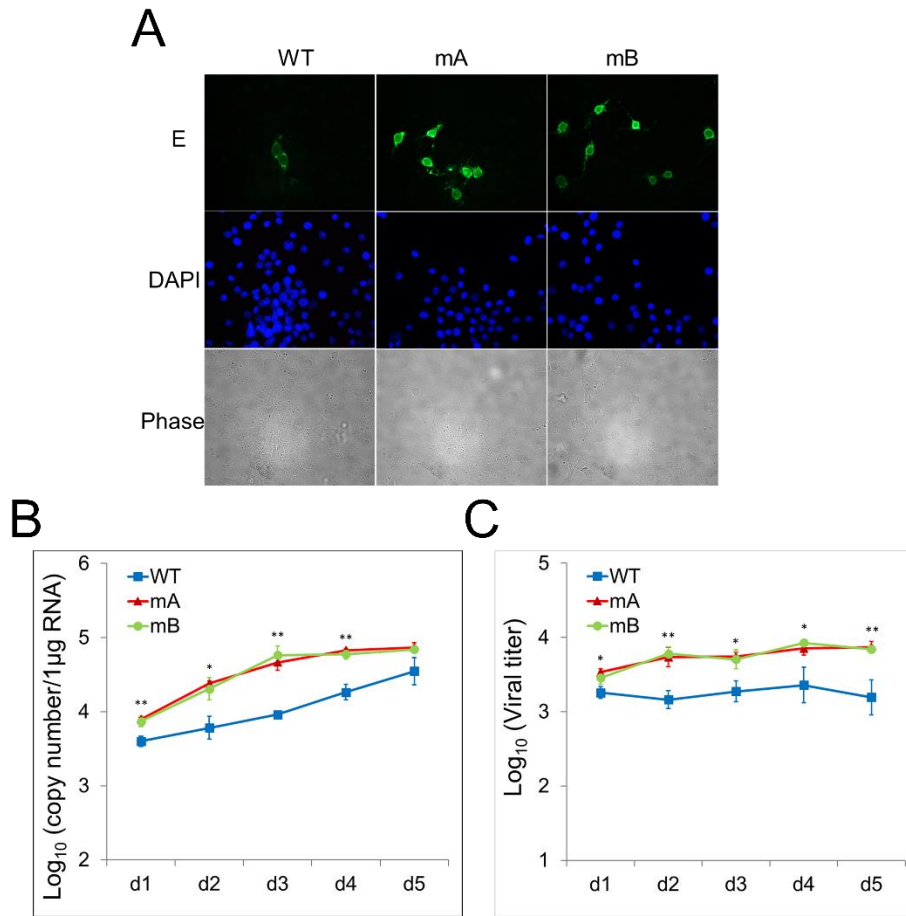
Fig. S5. Related to Fig. 6. Ablation of E protein N-linked glycosylation enhances ZIKV growth in *Aedes aegypti* CCL-125 cells. (A) Cells were infected with ZIKV WT, mA or mB at a MOI of 1 for two days, and then subjected to immunostaining using αE (4G2) antibody. (B) Cells were infected with ZIKV WT, mA or mB at a MOI of 0.1, harvested at indicated times for RNA purification and qRT-PCR analysis. Viral genome copy numbers were plotted as mean ± SD (n=3). (C) Infectious virus production. Supernatants were harvested from same cells as (B) to determine viral titers. Results were analyzed by unpaired Student's *t* test. Differences were considered statistically significant when p<0.05 (*) or p<0.01 (**).

Fig. S6.
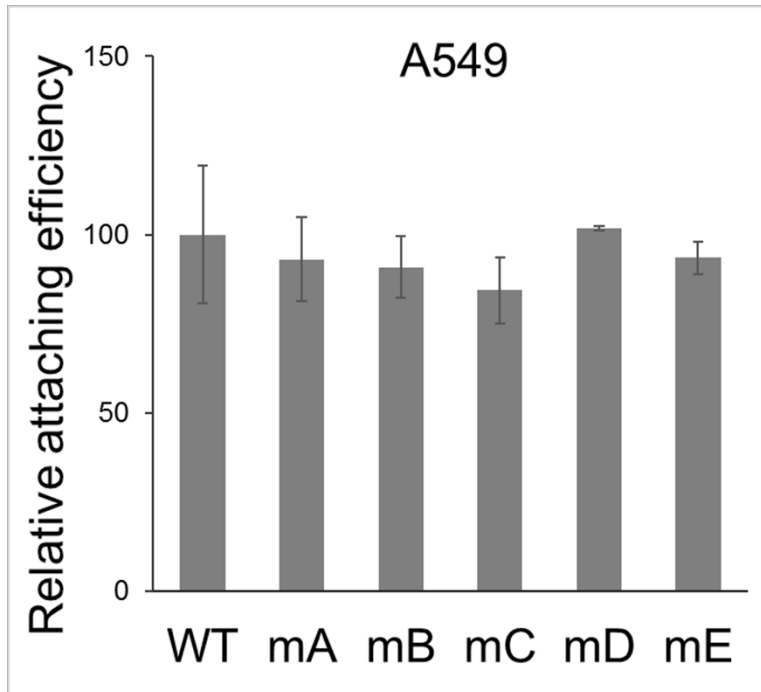
Fig. S6. Related to Fig. 6. Attachment efficiency of ZIKV WT and mutants on surface of A549 cells. A549 were culture to ~100% confluency, then incubated with ZIKV WT or mutants at a MOI of 10 at 4 ºC with constantly agitation. One and half hours later, cells were washed with ice cold PBS for three time, then immediately subjected to RNA purification and qRT-PCR to determine viral genome copy numbers. Relative attaching efficiency to WT virus was plotted as mean ± SD (n=3).

Fig. S7
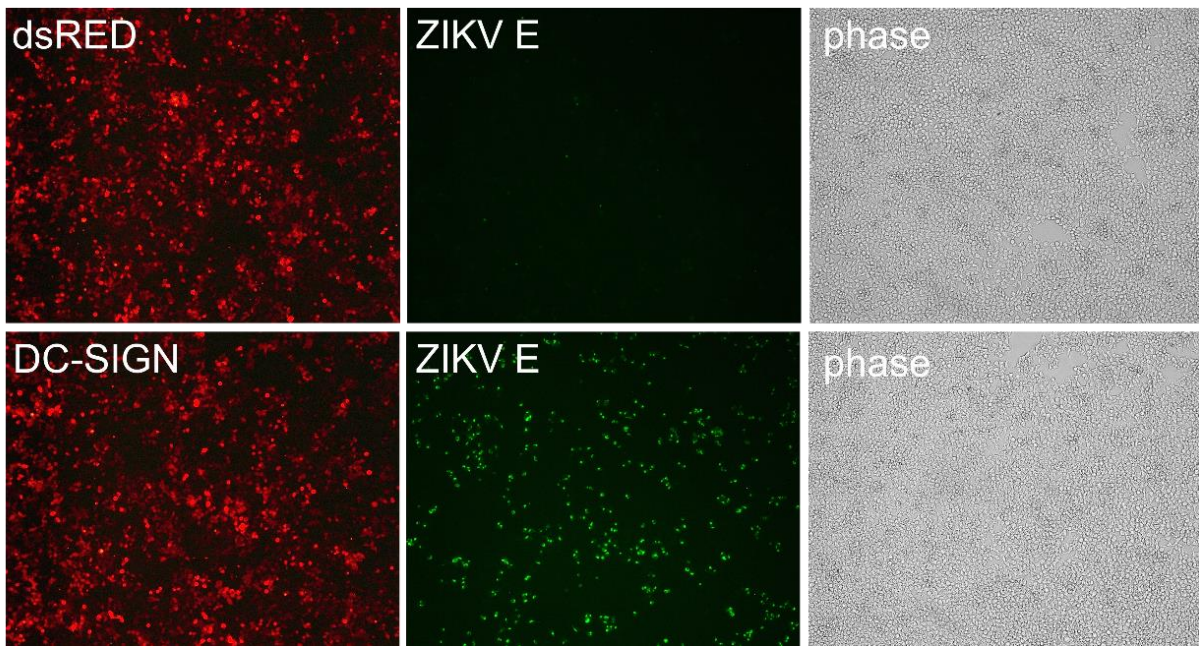
Fig. S7. Related to Fig. 7. Expression of DC-SIGN mediated ZIKV infection. 293T cells were transfected with pCMV-DCSIGN-cFLAG or pCMV-DsRed-express (control) plasmids and one day later, cells were infected with ZIKV WT at a MOI of 0.2. Cells were fixed at 24 hours post infection, and subjected to IFA analysis with αE (4G2) and αFLAG antibodies.

Supplemental Tables

Table S1. Sequence of primers. Related to Figure 1.

| primers	sequence	comments
BR-F	AGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATT CTAATACGACTCACTATAGAGTTGTTGATCTGTGTG AATCAG	Amply the first ZIKV genomic fragment
BR-R1	CTTCCCAAAAGCCAAGCGATGGCGGCCGCTGCTAA CGCGAAGCCAGG	
BR-F1	CCTGGCTTCGCGTTAGCAGCGGCCGCCATCGCTTG GCTTTTGGGAAG	Amply the second ZIKV genomic fragment
BR-R2	CGTTATAGACGAACACCCCGGTACCGCATCTCGTC TCCT	
BR-F2	AGGAGACGAGATGCGGTACCGGGGTGTTCGTCTAT AACG	Amply the third ZIKV genomic fragment
BR-R3	GTTGAACCTAGCAGTCTACGCGTCATTACTCTGTAC ACTCC	
BR-F3	GGAGTGTACAGAGTAATGACGCGTAGACTGCTAGG TTCAAC	Amply the fourth ZIKV genomic fragment
BR-R4	GCGCAGCTGCTGCCTGCAGGCCTGGGATCAAGTA CATGTA	
BR-F4	TACATGTACTTGATCCCAGGCCTGCAGGCAGCAGC TGCGC	Amply the fifth ZIKV genomic fragment
BR-R5	GTTCAAGAATCCAAGGGCTTCAAACTCTAGAAATCT AGCC	
BR-F5	GGCTAGATTTCTAGAGTTTGAAGCCCTTGGATTCTT GAAC	Amply the sixth ZIKV genomic fragment
BR-R	GCTTATCGATGATAAGCTGTCAAACATGAGAATTCG AAGACCCATGGATTTCCCCACACCGG	
ep-A-F	ATTGCCCCGGCATACAGC	Generate sub library A by error prone PCR
ep-A-R	GGCTTCGGCTCTCGGTGA	
ep-B-F	GAGCGAAAGTTGAGATAACGC	Generate sub library B by error prone PCR
ep-B-R	TCAACCTCCCAACTGGGGT	
ep-C-F	GCTCAGATGGCGGTGGAC	Generate sub library C by error prone PCR
ep-C-R	gattGGTACCGCATCTCGTCTCCTTCTTTGAGAAGTC CACCGAGCACCCCACATCAG	
Sub-Full-F	ACGAGGCCCTTTCGTCTTC	Amply whole fragment for cloning
Sub-Full-R	gattGGTACCGCATCTCGTCTCCT	

SubA-Left-R	GCTGTATGCCGGGGCAAT	With Sub-Full-F, generate left fragment of subA
SubA-Right-F	TCACCGAGAGCCGAAGCC	With Sub-Full-R, generate right fragment of subA
SubB-Left-R	GCGTTATCTCAACTTTCGCTC	With Sub-Full-F, generate left fragment of subB
SubB-Right-F	ACCCCAGTTGGGAGGTTGA	With Sub-Full-R, generate left fragment of subB
SubC-Left-R	GTCCACCGCCATCTGAGC	With Sub-Full-F, generate left fragment of subC
ZD1-F	TGATACTGCTGATTGCCC	Deep sequencing primer set 1
ZD1-R	TCGACAGTCGGTTTGTCC	
ZD2-F	TGTCACCGTAATGGCACA	Deep sequencing primer set 2
ZD2-R	GTAGGCTTCACCTTGTGT	
ZD3-F	TCTGACAGCCGCTGCCCA	Deep sequencing primer set 3
ZD3-R	ATGCAAACTTAGCGCATG	
ZD4-F	GCAAAGGGAGCCTGGTGA	Deep sequencing primer set 4
ZD4-R	GTTTCATGTCCTGTGTCA	
ZD5-F	CAGTGGGATGATCGTTAA	Deep sequencing primer set 5
ZD5-R	TGAAAAGTCAAGGCCTGT	
ZD6-F	CTTGATTGTGAACCGAGG	Deep sequencing primer set 6
ZD6-R	TGTTGTTCCAGTGTGGAG	
ZD7-F	CTGGGGCAGACACCGGAA	Deep sequencing primer set 7
ZD7-R	TCCATCTCAGCCTCCAGA	
ZD8-F	CACGGCCCTTGCTGGAGC	Deep sequencing primer set 8
ZD8-R	GATCTTGGTGAATGTGAA	
ZD9-F	TCCTTGTGTACTGCAGCG	

ZD9-R	TCCCAACTGGGGTCAGAG	Deep sequencing primer set 9
ZD10-F	TGGCGGTGGACATGCAAA	Deep sequencing primer set 10
ZD10-R	ATCTTCTTCTCCCCGACT	
ZD11-F	CTCTTACATTGTCATAGG	Deep sequencing primer set 11
ZD11-R	TCCAACTGATCCAAAGTC	
ZD12-F	TTGGGAGACACAGCCTGG	Deep sequencing primer set 12
ZD12-R	ACCACATCAGCAACGTTC	
ZD13-F	TCTCACAAATTCTCATTG	Deep sequencing primer set 13
ZD13-R	TTTGAGAAGTCCACCGAG	
N154D-F	AGTGGGATGATCGTT gAT GACACAGGACATGAA	A489G, mA
N154D-R	TTCATGTCCTGTGTCATCAACGATCATCCCACT	
N154T-F	AGTGGGATGATCGTT AcT GACACAGGACATGAA	A490C, mB
N154T-R	TTCATGTCCTGTGTCAGTAACGATCATCCCACT	
N154S-F	AGTGGGATGATCGTT AgT GACACAGGACATGAA	A490G, mC
N154S-R	TTCATGTCCTGTGTCACTAACGATCATCCCACT	
N154K-F	AGTGGGATGATCGTT AAa GACACAGGACATGAA	T491A, mD
N154K-R	TTCATGTCCTGTGTCTTTAACGATCATCCCACT	
T156A-F	ATGATCGTTAATGAC gCA GGACATGAAACTGAT	A495G, mE
T156A-R	ATCAGTTTCATGTCCTGCGTCATTAACGATCAT	
Y shape adapter-F	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNT	Anneal to generate Y shape adaptor
Y shape adapter-R	/5Phos/NNNAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAG	
UniFcell-F	AATGATACGGCGACCACCGAGATCTACACTCTTTCCTACACGAC	Amply adaptor ligated amplicon
UniFcell-R	CAAGCAGAAGACGGCATACGAGATCGGTCTCGGCATTCCTGCTGAACC	
DCSIGN-Full-F	agaGAATTccacc ATGAGTGACTCCAAGGAAC	Clone full length DC-SIGN with C-terminal FLAG tag into pCMV-noHA vector
DCSIGN-Full-R	GAGgtcga cta CTTGTCGTCATCGTCTTTGTAGTC CGCAGGAGGGGGGTTTG	

Table S2. Summary of deep sequencing reads count. Related Figure 2.

Samples	Total reads (Million)	WT Reads (Million)	Mutation Reads (Million)	Mutation Rate (%)
DNA-1	6.59	6.19	0.40	6.13
DNA-2	20.37	19.09	1.27	6.24
RNA-1	5.14	4.92	0.22	4.36
RNA-2	1.95	1.88	0.06	3.31
RNA-3	3.56	3.42	0.14	4.04
C6-1	3.97	3.76	0.20	5.15
C6-2	5.48	5.19	0.29	5.32
C6-3	6.48	6.17	0.31	4.81
A549-1	5.29	5.07	0.22	4.09
A549-2	5.37	5.12	0.24	4.50
A549-3	4.65	4.44	0.21	4.46
CMEC-1	3.24	3.09	0.16	4.85
CMEC-2	4.83	4.63	0.21	4.27
CMEC-3	12.52	11.94	0.58	4.61

Table S3. Frequency of mutation types. Related to Figure 2.

Original nt	Mutation frequency
C	0.15
G	0.18
T	0.22
A	0.27

Transparent Methods

Cells, viruses and plaque assays

Aedes albopictus C6/36 and *Aedes aegypti* CCL-125 cells were obtained from ATCC. C6/36 cells were maintained in DMEM medium supplemented with 10% FBS at 28 ^{o}C with 5% CO2; CCL-125 cells were maintained in MEM medium supplemented with 20% FBS. Vero and BHK21 cells were cultured in DMEM with 10% FBS. Human epithelial A549 cells were cultured in RPMI1640 with 10% FBS. Human cerebral microvascular endothelial cells (hCMEC/D3) were maintained in EndoGRO-MV complete culture media (Millipore). The parental ZIKV Puerto Rico strain PRVABC59 (GenBank number KU501215) was obtained from United States Centers for Disease Control and Prevention. Titers of infectious viral particles were measured by a standard plaque assay as descried previously (Qi et al., 2015). Briefly, 6X10^{4} Vero or A549 cells were seeded in 12-well plates one day before infection, then infected with virus for 1 h and overlaid with 1% methylcellulose (Sigma). Four days later, cells were fixed and stained with 0.2% crystal violet in 20% ethanol, and plaques were counted to determine the titer.

Human dendric cells generation

Human PBMCs were isolated from healthy donors by density centrifugation over Ficoll-Paque Plus in UCLA virology core. Monocytes were isolate from PBMCs with CD14 magnetic beads (Miltenyi Biotec) and cultured at 10^{6} cells/ml in RPMI1640 medium supplemented with 10% human serum, 1000U/ml each of human GM-CSF and IL-4 for 7 days.

Construction of ZIKV infectious cDNA clone

After receiving the virus from CDC, ZIKV was amplified in C6/36 cells, and total RNA of infected cells was extracted with Purelink RNA Mini Kit (Ambion). ZIKV cDNA covering the complete genome was synthesized using SuperScript III reverse transcriptase (Thermo Fisher) with random hexamer (Fragments 1 to 5) or ZIKV specific primer against the last 20 nt of ZIKV 3'-end (Fragment 6). Then, six ZIKV sub-genomic fragments covering the whole genome were amplified using KOD polymerase (Millipore). Six

fragments were assembled into EcoR I linearized pBR322 plasmid to generate pZ-PR plasmid using HiFi DNA Assembly Cloning Kit (NEB), and further transformed into DH10B E.coli. The whole sequence of pZ-PR plasmid was verified by Sanger DNA sequencing. Sequence of pZ-PR is available from NCBI GenBank: KY583506 and also in supplemental Materials and methods.

Reconstitution of ZIKV from pZ-PR plasmid

Recombinant ZIKV was generated from plasmid pZ-PR, containing the full-length cDNA of ZIKV Puerto Rico strain, using the previously describe method (Shan et al., 2016) with some modifications. In Brief, pZ-PR plasmid was amplified in DH10B E.coli and then purified using PureLink HiPure Plasmid Midiprep Kit (Thermo Fisher). To generate DNA template for RNA *in vitro* transcription, 100 μg of pZ-PR was linearized with BstBI, followed by end blunting with Mung Bean Nuclease (NEB). Then, the end-blunted template DNA was purified with phenol-chloroform, precipitated with ethanol, and resuspended in 20 μl RNase-free water. 5'-capped RNA was *in vitro* transcribed using mMESSAGE mMACHINE T7 kit (Ambion) with 1 μg template DNA and an additional 1 μl of 30 mM GTP solution. *In vitro* transcribed RNA was further purified with phenolchloroform, precipitated with ethanol, resuspended in 100 μl RNase-free water, aliquoted, and stored in -80 °C freezer.

For cell transfection, 12 μg of RNA was mixed with $5X10^6$ BHK21 cells in 200 μl Electroporation Solution (Bioland), and electroporated in 4-mm cuvette with the

GenePulser apparatus (Bio-Rad) at the setting of 240 V and 950 μF, pulsing once. After 10 min recovery at room temperature, transfected cells were resuspended in 10 ml warm culturing medium, incubated in cell incubator overnight, and then washed once and replenished with fresh medium. Three to four days post electroporation when ~40% of cells showing CPE, supernatants were collected, span at 8000 g for 10 min at 4 °C to remove cellular debris, aliquoted and stored at -80 °C.

Generation of the plasmid DNA library of ZIKV E protein mutations

To generate the mutant plasmid DNA library of ZIKV E protein, we designed an errorprone PCR strategy utilizing the pZ-PR plasmid as the template and the GeneMorph II Random Mutagenesis Kit (Agilent) to generate the point mutations during PCR

amplification (Al-Mawsawi et al., 2014, Wu et al., 2015). The ZIKV E coding region was divided into three segments ranging from 516 to 545 bp (Fig. 2A) for the purpose of controlling mutational rate in error-prone PCR reaction. The mutated inserts were generated with 12-cycles amplification using the GeneMorph II Random Mutagenesis Kit, 1 µg pZ-PR plasmid and corresponding primers. To generate full length fragments for cloning, the corresponding left and right fragments were amplified using high fidelity KOD DNA polymerase (Millipore) for 18 cycles. Each mutated insert and its corresponding left and right fragments were combined, and the resulting full length DNA inserts were generated by amplification with KOD polymerase for 20 cycles. Then, the full length inserts were digested with KpnI and EcoRI, while the vector pZ-PR plasmid was digested with KpnI and EcoRI followed by treatment with Shrimp Alkaline Phosphatase (NEB). Ligation was performed for each of the three sub libraries with T4 DNA ligase (NEB) with a vector: insert ratio of 1:5. Ligated products were purified with phenol-chloroform, precipitated with ethanol, resuspended in 10 µl sterilized water, and electroporated into DH10B E.coli competent cells. For each of the three sub libraries, ~40,000 colonies were collected from LB plates, and directly subjected to plasmid DNA purification. Equal amounts of the three sub libraries were then mixed to generate the DNA library, and stored at -80 $^{\circ}$C in aliquot.

Preparation of virus library and passages
To reconstitute the virus library, 5' capped RNA was *in vitro* transcribed from the plasmid DNA library, electroporated into BHK21 cells, and cell supernatant was collected at 4 days post transfection. Cell supernatant containing the library of virus mutants was span at 8000 g for 10 min at 4 $^{\circ}$C to remove cellular debris, titrated and stored at -80 $^{\circ}$C in aliquot. Virus library was further passaged in C6/36, A549 or hCMEC cells by infecting 40 million cells at a MOI of 0.05. Virus supernatants were collected at 4 days post infection, debris clarified, RNase A treated, and then subjected to RNA purification followed by deep sequencing analysis.

Next-generation sequencing of virus mutations and data analysis

Viral supernatants were treated with RNase A, then virion RNA was extracted using QIAamp Viral RNA Mini Kit (Qiagen), and reverse transcribed to cDNA using Superscript III Reverse Transcriptase (Thermo Fisher). The plasmid DNA library or cDNA from the viral libraries (virus or passaged libraries) were used as templates to generate thirteen158-bp deep sequencing amplicons. The resulting PCR amplicons were 3'-end dA-tailed, and ligated to Y-shape sequencing adapters using T4 DNA ligase (NEB). Yshape adapters were generated by annealing two oligos: 5'-ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT NNN NNN T-3' and 5'-/5Phos/NNN NNN AGA TCG GAA GAG CGG TTC AGC AGG AAT GCC GAG-3'. The six "N" stands for the multiplex ID for distinguishing different samples. The adapter-ligated products were further PCR amplified for 18 cycles using KOD polymerase and UniFcell primers. The final PCR products were gel purified, and pooled for deep sequencing on an Illumina Hiseq 3000 platform with 150 bp paired-end reads. Raw sequencing reads were demultiplexed by using the six-nucleotide barcodes. Sequencing error was corrected by filtering unmatched forward and reverse reads. Each mutation was called by comparing individual reads to the WT pZ-PR reference sequence. The relative frequency and relative fitness score (RF score) for a given mutation was calculated as follows: For a mutation i in deep sequencing amplicon n of sample t (where t could be DNA plasmid, or virus libraries):

Relative frequency$_{i,n,t}$ = Read count$_{i,n,t}$/ Coverage$_{n,t}$, where Read count$_{i,n,t}$ represented the number of read in amplicon n of sample t that carried mutation i and Coverage$_{n,t}$ represented the total read count (including WT and all mutations) of the amplicon n of sample t.

RF score = Relative frequency$_{i,n,Virus}$ / Relative frequency$_{i,n,DNA}$, where Relative frequency$_{i,n,Virus}$ represented relative frequency of mutation i in viral library (virus or passaged library) and Relative frequency$_{i,n,DNA}$ represented relative frequency of mutation i in DNA plasmid library. Data and Software Availability

All analysis was performed with customized python scripts, which are available upon request. Raw sequencing data is available from NIH Short Read Archive: PRJNA373927. Polymorphism calculation

Polymorphism score of each residue was calculated using a formula previously reported (Crooks et al., 2004). S = -100 * Sum (Pi * logPi) where Pi is the frequency of mutations at position i. The score is the normalized entropy of the observed mutation distribution.

Construction of individual ZIKV mutants

Single nucleotide mutations were introduced into pZ-PR plasmid individually. To generate mutant virus, BHK21 cells were electroporated with *in vitro* transcribed RNA, and viral supernatant were collected at 4 days post transfection. Mutant viruses were further amplified in C6/36 cells, then supernatants were collected, clarified of debris and stored at -80 ℃ in aliquot. Virion RNA was also extracted, and reverse transcribed to cDNA. E protein coding sequence of all mutants were PCR amplified individually, followed by Sanger DNA sequencing to verify the sequence.

Reverse transcription and real-time PCR

Quantitative RT-PCR was performed as previously described (Gong et al., 2016). Briefly, total RNA was extracted from infected cells with Purelink RNA Mini Kit (Ambion) or from RNAase A treated cell supernatant with QIAamp Viral RNA Mini Kit (Qiagen), reverse transcribed using Superscript III Reverse Transcriptase (Thermo Fisher) and random hexamer, and then subjected to real-time PCR analysis using primers: 5'-TTG TGG AAG GTA TGT CAG GTG-3' and 5'- ATC TTA CCT CCG CCA TGT TG-3'.

Western blotting

ZIKV infected C6/36 or A549 cells were lysed in denaturing buffer, mock treated or treated with PNGase F(NEB) to remove N-linked glycan from proteins. Then proteins were diluted 5 times in SDS-PAGE sample buffer, heated at 95 ℃, resolved by SDS-PAGE gel electrophoresis, and transferred onto PVDF membrane. ZIKV E protein was detected with the monoclonal antibody D1-4G2-4-15 (Absolute Antibody). HRP-conjugated secondary antibodies were used and detection was performed with SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher).

Indirect immunofluorescence assay

Expression and localization of ZIKV E proteins was determined by indirect immunofluorescence assay. Cells were infected with ZIKV WT or mutants, fixed in 2% paraformaldehyde, permeabilized with 0.1% Triton-X100, and then blocked with 3% BSA plus 10% FBS. ZIKV E proteins were detected with mouse monoclonal antibody clone D1-4G2-4-15 (Absolute Antibody), followed by Alexa Fluor 488-conjugated anti-mouse IgG antibody (Life Sciences). DAPI was used for the staining of DNA. Samples were examined using a Nikon microscope using 63X/1.40-0.60 oil lens.

Statistical analysis

All numerical data were calculated and plotted with mean+/- SD. Results were analyzed by unpaired Student's *t* test. Differences were considered statistically significant when $p<0.05$ (*) or $p<0.01$ (**).

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal anti-ZIKV E	Absolute Antibody	D1-4G2-4-15
CD14 magnetic beads	Miltenyi Biotec	130-050-201
Monoclonal anti-FLAG	Sigma	F3165
Bacterial and Virus Strains		
Zika virus Puerto Rico strain	Centers for Disease Control and Prevention	PRVABC59
Biological Samples		
Human PBMCs from healthy donor	UCLA virology core	N/A
Chemicals, Peptides, and Recombinant Proteins		
Critical Commercial Assays		
GeneMorph II Random Mutagenesis Kit	Agilent	200550
mMESSAGE mMACHINE T7 kit	Ambio	mMESSAGE mMACHINE T7 kit

Deposited Data		
Sequence of pZ-PR plasmid	NCBI GenBank	KY583506
Deep sequencing raw reads on an Illumina Hiseq 3000 platform	NIH Short Read Archive	PRJNA373927
Experimental Models: Cell Lines		
Aedes albopictus C6/36	ATCC	CRL-1660
Aedes aegypti CCL-125 cells	ATCC	CCL-125
Human: A549	ATCC	CRM-CCL-185
Hamster: BHK-21	ATCC	CCL-10
Monkey: Vero	ATCC	CRL-1586
Human cerebral microvascular endothelial cells (hCMEC/D3)	Millipore	SCC066
Experimental Models: Organisms/Strains		
Oligonucleotides		
Primers for cloning, qPCR, library construction, and deep sequencing	This paper	N/A
Recombinant DNA		
ZIKV cDNA plasmid pZ-R	This paper	N/A
DC-SIGN expression plasmid pCMV-DCSIGN-cFLAG	This paper	N/A
pCMV-DsRed-expression	Clontech	632416
pCMV-noHA	Gong et al 2016	https://www.ncbi.nlm.nih.gov/pubmed/27832591
Software and Algorithms		
customized python scripts	This paper	N/A

Supplemental References

AL-MAWSAWI, L. Q., WU, N. C., OLSON, C. A., SHI, V. C., QI, H., ZHENG, X., WU, T. T. & SUN, R. 2014. High-throughput profiling of point mutations across the HIV-1 genome. *Retrovirology,* 11**,** 124.

CROOKS, G. E., HON, G., CHANDONIA, J. M. & BRENNER, S. E. 2004. WebLogo: a sequence logo generator. *Genome Res,* 14**,** 1188-90.

GONG, D., KIM, Y. H., XIAO, Y., DU, Y., XIE, Y., LEE, K. K., FENG, J., FARHAT, N., ZHAO, D., SHU, S., DAI, X., CHANDA, S. K., RANA, T. M., KROGAN, N. J., SUN, R. & WU, T. T. 2016. A Herpesvirus Protein Selectively Inhibits Cellular mRNA Nuclear Export. *Cell Host Microbe,* 20**,** 642-653.

QI, J., HAN, C., GONG, D., LIU, P., ZHOU, S. & DENG, H. 2015. Murine Gammaherpesvirus 68 ORF48 Is an RTA-Responsive Gene Product and Functions in both Viral Lytic Replication and Latency during In Vivo Infection. *J Virol,* 89**,** 5788-800.

SHAN, C., XIE, X., MURUATO, A. E., ROSSI, S. L., ROUNDY, C. M., AZAR, S. R., YANG, Y., TESH, R. B., BOURNE, N., BARRETT, A. D., VASILAKIS, N., WEAVER, S. C. & SHI, P. Y. 2016. An Infectious cDNA Clone of Zika Virus to Study Viral Virulence, Mosquito Transmission, and Antiviral Inhibitors. *Cell Host Microbe,* 19**,** 891-900.

WU, N. C., OLSON, C. A., DU, Y., LE, S., TRAN, K., REMENYI, R., GONG, D., AL-MAWSAWI, L. Q., QI, H., WU, T. T. & SUN, R. 2015. Functional Constraint Profiling of a Viral Protein Reveals Discordance of Evolutionary Conservation and Functionality. *PLoS Genet,* 11**,** e1005310.