Mapping the evolutionary landscape of Zika virus infection in immunocompromised mice

Katherine E. E. Johnson,1,† Maria G. Noval,2,† Margarita V. Rangel,2,† Elfie De Jesus,2 Adam Geber,1 Samantha Schuster,2,3 Ken Cadwell,2,3 Elodie Ghedin,1,4,‡ and Kenneth A. Stapleford2,*,§

1Department of Biology, Center for Genomics & Systems Biology, New York University, New York, NY, USA, 2Department of Microbiology, New York University School of Medicine, New York, NY 10016, USA, 3Kimmel Center for Biology and Medicine at the Skirball Institute, New York University School of Medicine, New York, NY 10016, USA and 4Department of Epidemiology, School of Global Public Health, New York University, New York, NY, USA

*Corresponding author: E-mail: Kenneth.stapleford@nyulangone.org
†These authors contributed equally to this work.
‡Present address: National Institutes of Health/NIAID, elodie.ghedin@nih.gov.
§https://orcid.org/0000-0002-7796-2254

Abstract

The fundamental basis of how arboviruses evolve in nature and what regulates the adaptive process remain unclear. To address this problem, we established a Zika virus (ZIKV) vector-borne transmission system in immunocompromised mice to study the evolutionary characteristics of ZIKV infection. Using this system, we defined factors that influence the evolutionary landscape of ZIKV infection and show that transmission route and specific organ microenvironments impact viral diversity and defective viral genome production. In addition, we identified in mice the emergence of ZIKV mutants previously seen in natural infections, including variants present in currently circulating Asian and American strains, as well as mutations unique to the mouse infections. With these studies, we have established an insect-to-mouse transmission model to study ZIKV evolution in vivo. We also defined how organ microenvironments and infection route impact the ZIKV evolutionary landscape, providing a deeper understanding of the factors that regulate arbovirus evolution and emergence.

Key words: Zika virus; transmission; diversity.

1. Introduction

Arboviruses (arthropod-borne viruses) exist in a complex and dynamic life cycle between an arthropod vector (e.g. mosquitoes, ticks) and a host (e.g. mammals, birds, plants) (Weaver et al. 2018). It is within these disparate hosts that arboviruses replicate and undergo genomic evolution, which plays an essential role in transmission and pathogenesis (Weaver 2006; Parvez and Parveen 2017). To date, we understand little of how arboviruses evolve, are transmitted, or cause disease in nature and thus it is crucial to study these aspects of virus biology in a controlled laboratory setting.

Zika virus (ZIKV) is a positive-strand RNA virus and a member of the Flaviviridae family (Miner and Diamond 2017), which includes other human pathogens such as West Nile virus...
(WNV), dengue virus (DENV), and Yellow Fever virus. ZIKV has been responsible for several outbreaks including a recent and explosive epidemic in 2015–16 where it swept through the Americas causing severe disease (Campos, Bandeira, and Sardi 2015; Fauci and Morens 2016; Lucchesi and Kanduc 2016). Importantly, although we still do not completely understand the molecular basis of ZIKV emergence and disease, viral genomic evolution has been implicated as a driving force for ZIKV disease and transmission in these recent epidemics (Shi et al. 2016; Aldunate et al. 2017; Delatorre, Mir, and Bello 2017; Liu et al. 2017; Metsky et al. 2017; Yuan et al. 2017). This gap in knowledge, coupled with overall increases in the evolution of vector-borne pathogens and their spread to naïve vector and host populations (Hermance and Thangamani 2017; Gill et al. 2019), highlights the need to study arbovirus evolution and emergence during vector-borne transmission.

ZIKV can be transmitted by a variety of routes, including through sexual contact (Hastings and Fikrig 2017) and mother to fetus (Coyne and Lazear 2016), yet the primary route of infection is by the widespread and invasive mosquito species Aedes (Ae.) aegypti and Ae. albopictus (Magalhaes et al. 2018). To date, several important studies using mouse models have addressed the physiological effects of vector-borne arbovirus transmission (Pingen et al. 2016; Schmid, Harris, & McKimmie, 2017). These seminal studies, using human pathogens such as DENV (Cox et al. 2012; Schmid et al. 2016), WNV (Schneider et al. 2006, 2010), Semliki Forest virus (Pingen et al. 2016), chikungunya virus (CHIKV) (Agarwal et al. 2016), and Rift Valley fever virus (Le Coupance et al. 2013) revealed that mosquito bites can enhance viral infections, increase pathology, and alter host immune responses when compared to conventional needle inoculations. However, with the exception of several studies using rhesus macaques (Dudley et al. 2017; Aliota et al. 2018), the impact of vector-borne transmission on ZIKV viral evolution and adaptation has not been extensively explored, underscoring the need to study these processes in the context of their natural route of transmission, from mosquito to mammal. In particular, the role that specific organ microenvironments play in shaping ZIKV evolution by influencing viral diversity and defective viral genomes (DVGs) in mammals has not been investigated and is essential to our understanding of how arboviruses evolve and emerge.

In this study, we infected interferon receptor deficient (Ifnar−/−) mice either by the natural transmission route via mosquito bite or by subcutaneously infection via the footpad. We then monitored disease, isolated and quantified viral RNA and infectious particles in multiple organs, and performed a viral genome deep sequencing analysis to compare ZIKV diversity between the two inoculation routes. We found that while signs of disease and viral RNA accumulation do not differ between inoculation route, viral diversity and infectious viral particle production were impacted. Interestingly, viral diversity as well as the production of DVGs was organ specific, suggesting that organ-specific selective pressures and bottlenecks may drive the emergence of new viral variants, with potential impacts on viral disease and spread. Finally, we observed the emergence of unique ZIKV variants as well as variants present in currently circulating strains. Taken together, these studies highlight the importance of transmission route and organ-specific microenvironments in regulating viral diversity and the emergence of new viral variants. These results underscore the need to understand the driving forces of viral diversity throughout the complex and complete viral life cycle of ZIKV.

2. Materials and methods

2.1 Cells and viruses

293T cells (ATCC CRL3216) were grown in Dulbecco’s modified Eagle medium (DMEM) supplemented with 1% penicillin/streptomycin (P/S), 1% non-essential amino acids, and 10% fetal bovine serum (Atlanta Biologicals) at 37 °C with 5% CO2. Vero cells (ATCC CCL-81) were maintained in DMEM supplemented with 1% P/S and 10% newborn calf serum (Gibco) at 37 °C with 5% CO2. All cells were verified to be mycoplasma free.

The Ugandan (MR766) strain of ZIKV was generated from a plasmid-based infectious clone obtained from Dr. Matthew Evans at the Icahn School of Medicine at Mt. Sinai (Schwarz et al. 2016). To generate infectious virus, 293T cells were transfected with 0.5 μg of plasmid via Lipofectamine 2000 transfection reagent (Invitrogen) and virus-containing supernatants were harvested 48 h post transfection, centrifuged at 1,200 × g for 5 min, aliquoted and stored at −80 °C. To generate a working viral stock, transfection virus stocks were used to infect Vero cells and virus containing supernatants were harvested 48 h post infection centrifuged at 1,200 × g for 5 min, aliquoted, and stored at −80 °C. Viral titers were determined by plaque assay on Vero cells, as described below.

2.2 Viral titrations

Viral titers were determined by plaque assay on Vero cells (Cifuentes Kottkamp et al. 2019). In brief, virus was subjected to ten-fold serial dilutions in DMEM and added to a monolayer of Vero cells for 1 h at 37 °C. Following incubation, a 0.8% agarose overlay was added, and cells were incubated for 5 days at 37 °C. Five days post infection, cells were fixed with 4% formalin, the agarose overlay removed, and plaques were visualized by staining with crystal violet (10% crystal violet and 20% ethanol). Viral titers were determined on the highest dilution virus could be counted.

2.3 Mosquito infections and manipulations

Ae.aegypti mosquitoes (Poza Rica, Mexico, F18-20) were a kind gift from Gregory Ebel at Colorado State University (Ruckert et al. 2017; Noval et al. 2019). Mosquitoes were reared and maintained in the NYU School of Medicine ABSL3 facility at 28 °C and 70% humidity with a 12:12 h diurnal light cycle. The day before infection, female mosquitoes were sorted and starved overnight. The day of infection, mosquitoes were exposed to an infectious bloodmeal containing freshly washed rabbit blood, 5 mM ATP, and 106 plaque forming units (PFU)/ml virus for approximately 30 min, subsequently cold anesthetized, and engorged female mosquitoes were sorted into new cups. Engorged mosquitoes were incubated at 28 °C with 70% humidity for 14 days and fed ad libitum with 10% sucrose. Following incubation, mosquitoes were allowed to feed on naïve Ifnar−/− mice as described below. After mouse feeding, mosquitoes were placed into a 2-ml tubes containing 200 μl phosphate buffered saline (PBS) and a steel ball. Mosquitoes were ground using a Tissuelyser (Qiagen) with 30 shakes/second for 2 min. After grinding, each mosquito homogenate was mixed with an equal volume of Trizol (Invitrogen) for RNA extraction.

2.4 Mouse infections and transmission studies

Animal experiments were performed in accordance with all NYU School of Medicine Institutional Animal Care and Use
Committee guidelines (IACUC). About 6-week old male and female Ifnar<sup>−/−</sup> mice crossed to the C57BL/6 background were bred and housed in the animal facility at the NYU School of Medicine and subsequently transferred to the NYU School of Medicine ABSL3 animal facility for all experiments. For needle inoculations, mice were anesthetized briefly with isoflurane and inoculated in the right rear footpad with 50 PFU of ZIKV. Mice were weighed and monitored daily for signs of disease and euthanized at a defined humane endpoint where the mice had lost at least 20 per cent of their initial body weight. Mice were dissected and organs placed in 2 ml round bottom tubes containing 500 µl of PBS and a steel ball. Organs were homogenized as described above, clarified by centrifugation at 8,000 rpm for 10 min, and virus containing supernatants were used directly to quantify viral titers by plaque assay or mixed with equal volume of Trizol for RNA extraction and RT-qPCR.

For transmission studies (Carrau et al. 2019), ZIKV-infected mosquitoes (~5/cup) were allowed to feed on the tail of male and female Ifnar<sup>−/−</sup> mice for 30 min. Following feeding, mice were placed back in their cages and all mosquitoes, regardless of being engorged, were processed as described above. Mice were weighed and monitored daily for signs of disease and euthanized at the humane endpoint. Mice were dissected and organs were processed as described above.

2.5 RNA extractions and RT-qPCR

Total RNA from homogenized mosquitoes and mouse organs were isolated by Trizol following the manufacturer’s instructions and resuspended in 50 µl of nuclease-free water. Total RNA was quantified, diluted to equal amounts of RNA for each organ, and viral genomes were quantified using the Taqman RNA-to-Ct kit (Applied Biosystems) using primers in Supplementary Table S1. A standard curve was generated from in vitro transcribed ZIKV RNA for each experiment, as previously described.

2.6 Genome amplification and library preparation

The ZIKV genome was amplified by PCR in three overlapping fragments for deep sequencing analysis. In brief, 200 ng of RNA from each sample was used to generate virus cDNA using a Maxima H minus strand kit (Invitrogen) and ZIKV-specific primers (Supplementary Table S1). cDNA was then used immediately to amplify the genome using Phusion high-fidelity DNA polymerase (Thermo) with the primers in Supplementary Table S1. PCR products were purified using a PCR cleanup kit (Macherey-Nagel) and with 0.9× AMPure beads (Beckman, Culture, Inc.). All amplicons were normalized to 2.5 ng before being used as input into the Nextera DNA Library Preparation Kit (Illumina). Libraries were purified with 0.55× AMPure beads, quantified, pooled in equimolar ratios, and sequenced on the NextSeq500 at MidOutput 2 × 150—300 Cycle v2.

2.7 Deep sequencing analysis

Sequencing reads were trimmed using Trimmomatic v0.36 and the Nextera adaptors file; a sliding window of 4 bp scanned the reads to make sure the average quality of the base pair call was above 15 and that the minimum length was at least 36 bp long. The trimmed reads from the three individual amplicons were initially aligned separately to the reference genome (MR766GenomeReference; GenBank: KX309660.1) using bowtie2 v2.2.9 (Kim et al. 2013) –no-mixed –very-sensitive –local parameters. The alignments were then sorted with Samtools v1.6 (Langmead and Salzberg 2012) and deduplicated using Picard Tools MarkDuplicates v2.8.2. Minor variants were identified using an in-house variant caller (https://github.com/GhedinLab/ZIKV_Analysis). Coverage and minority variant calls were checked to ensure overlapping regions were identical in their nucleotide composition before merging the fastq files and then realigning the three amplicons to the reference file at once. Minority variants were called again on the merged alignment files and the amino acid position was added using the positions indicated on the MR766 NCBI site. Minority variants present at a nucleotide position with 500× or above coverage and a frequency of 1 per cent or greater were used for richness calculations. All pairwise genetic distance and Shannon entropy calculations used minority variants present at a position with at least 200× coverage and a frequency of 3 per cent or greater. Further, the minor variant had to be present in both forward and reverse reads and have a 25 or above quality score.

2.8 Pairwise genetic distance

Variant files generated from the complete coding sequence were used as input to calculate the L2-norm, which uses the Euclidean distance to perform an all-versus-all pairwise comparison of each sample at each nucleotide position.

\[
d_k(p, q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}
\]

Here, \(d_k\) is the distance between two samples at the given position \(k\), where \(n\) is the total number of possible nucleotides (G, C, U, A) and \(p\) and \(q\) are the relative frequencies of the different alleles. Only frequencies of major and minor nucleotide variants were considered, all remaining nucleotides were considered to have a frequency of 0. The total distance measured between two samples, or \(D\), was calculated by summing all nucleotide site distances (\(d_k\)) across the length, \(N\), of the coding sequence.

\[
D = \sum_{k=1}^{N} d_k
\]

2.9 Within-host diversity

Shannon entropy was used as the measurement for within-host diversity within each of the samples. In short, entropy scores \((h)\) are calculated using the frequency, \(p_i\), for each variant at position \(i\) and summed across the number of alleles, \(S\).

\[
H = - \sum_{i=1}^{S} p_i \log_2 p_i
\]

Nucleotide diversity was measured as outlined in Zhao and Illingworth (2019).

2.10 Defective viral genome (DVG) identification

Deletion coordinates of DVGs were identified by aligning the individual PCR libraries to the reference genome using the split-read aligner, STAR v2.5.3a (Dobin et al. 2013). Coordinates of the deletions were pulled using the CIGAR string from the alignment files (https://github.com/GhedinLab/ZIKV_Analysis). A simulated MR766 DVG dataset was used to determine the
necessity input parameters to account for noise generated during alignment. For a deletion to be called, at least 25 nucleotides had to align to the reference genome before and after the deletion. We allowed for one insertion or deletion to be present and grouped the start and end coordinates when located 10 nucleotides apart. We further filtered out low-confidence DVG calls using the frequency of the most abundant DVG identified in each PCR amplicon of the plasmid as thresholds for our remaining PCR amplicon libraries.

2.11 Data subsampling
To determine how sampling depth impacts diversity measurements, each merged and deduplicated alignment file was randomly subsampled at five different frequencies (75%, 50%, 25%, 10%, and 5%); https://github.com/GhedinLab/ZIKV_Analysis. Therefore, the final number of reads remaining after sampling depends on the number of reads present before sampling. To ensure that sampling did not introduce random artifacts, we also randomly sampled 100 per cent of the reads. Variants were called using the subsampled alignment files, and diversity measurements were calculated as outlined. To determine the impact of sampling depth on DVG diversity and dynamics, we subsampled individual PCR alignments at five different frequencies (75%, 50%, 25%, 10%, and 5%) before calling DVGs. We filtered out low-confidence DVG calls using the frequency of the most abundant DVG identified in each subsampled plasmid PCR amplicon library as a threshold for the remaining amplicon libraries.

2.12 Zika virus sequence alignments
The Ugandan (MR766, KX830960.1), Puerto Rican (PRVABC59, MK713748.1), Brazilian (Paraiba, KX280026.1), and Cambodian (FSS13025, MH158236.1) ZIKV genome sequences were aligned using MegAlign (www.dnastar.com).

2.13 Statistics
All statistical analyses were performed using GraphPad Prism and R Studio. Data represent three independent mosquito feeds and three independent mouse infections (N = 3 total mice) and two independent needle infections (N = 7 total mice). Data are represented as the average ± the standard error of the mean (SEM). Mann–Whitney U and Kruskal–Wallis with Dunn’s post-test, considering a type-I error were performed and indicated in the figure legends. P values >0.05 were considered non-significant (ns).

3. Results
3.1 ZIKV inoculation route does not impact pathogenesis or viral RNA levels in immunocompromised mice
The study of arbovirus evolutionary trajectories within and between hosts is essential to understand the fundamental mechanisms these viruses use for transmission and pathogenesis. To gain insight into the ZIKV evolutionary trajectories, we infected type I interferon receptor deficient (lnfαr−/−) mice using either a natural transmission route via mosquito bite or subcutaneously via footpad injection. We took advantage of the well-established infαr−/− mouse model (Lazear et al. 2016), which favors unrestricted viral replication, and the prototypical African lineage Ugandan strain of ZIKV (MR766) that has enhanced replication and pathogenesis compared to other contemporary strains of ZIKV. We hypothesized that using an evolutionary distant virus from the currently circulating strains would provide better insight into the evolutionary trajectory of ZIKV in nature, and that generating a relatively homogeneous virus stock from an infectious clone would allow us to follow the de novo evolution of ZIKV. We infected Ae aegypti mosquitoes via a bloodmeal containing 10⁶ PFU/ml of the infectious clone-derived MR766 virus. After an incubation period of 14 days, we allowed infected mosquitoes to feed on lnfαr−/− mice. In parallel, we infected lnfαr−/− mice subcutaneously in the footpad with 50 PFU of the same stock of virus and compared the disease progression and viral RNA levels generated by these two different routes of inoculation (Fig. 1A). We found no significant differences in weight loss (Fig. 1B) and overall signs of disease (Fig. 1C) between mice infected by either mosquito bite or needle inoculation. We euthanized these mice at 7 days post-infection, which coincided with the humane endpoint, and quantified viral RNA genomes from different target organs. Interestingly, we did not observe any significant differences in viral RNA levels between inoculation routes (Fig. 1D). These data are in contrast to what has been previously reported for ZIKV vector-borne infections in rhesus macaques (Dudley et al. 2017) and may suggest that in the absence of a type I interferon response, vector-borne transmission could follow a similar course of infection as seen in subcutaneous inoculations.

3.2 ZIKV genomic changes in lnfαr−/− mice have similar evolutionary characteristics as found in nature
Next, we asked whether the infection of these mice recapitulated the evolutionary dynamics of ZIKV observed in nature. To do this, we amplified by PCR the ZIKV genome (nt 21 to 10,528), including the protein-coding region, in three over-lapping amplicons from: (1) our viral stock (used to infect mosquitoes for vector-borne transmission and to inoculate mice via the footpad), (2) organs isolated from individual needle-inoculated mice, (3) whole infected mosquitoes, and (4) organs isolated from individual mosquito-infected mice. We successfully amplified the complete coding region from the majority of our samples and performed a deep sequencing analysis of the viral subpopulations, where we obtained similar numbers of sequencing reads for each sample (Supplementary Fig. S1). Upon sequencing, we observed that our original ZIKV stock (Fig. 2A, Stock 1—Red) had one consensus change (variant frequency >50%) in NS5 (A9133T), the viral RNA-dependent RNA polymerase, which was also found in the plasmid and thus this variant was removed from further analysis (Supplementary Table S2). In addition, in stock 1 we identified three synonymous and six non-synonymous minority variants (variant frequency between 1% and 50% of the population) located either in the capsid or in NS5 (Fig. 2A and Supplementary Table S3). In mice that were needle inoculated with stock 1, we identified two novel non-synonymous consensus changes in Mouse 7, one in NS4B (T7582C, F2496L, 63%) and one in NS5 (G8221T, 95.9%) that were not found in the viral stock (Fig. 2C, Mouse 7). In mosquitoes (Fig. 2C), we found only a few minor variants above background (>1%) with mosquito 1 having the majority of these mutations, including two novel consensus changes in NS5 (G8025A, G2640E, 92.9%); and C8221T, 95.9%) that were not found in the viral stock (Fig. 2C and Supplementary Table S2). However, in mice infected via these mosquitoes (Fig. 2D, Mouse A and Mouse B), the number of minority variants increased above background, similar to what is seen in the needle-
inoculated mice. To confirm these findings with three independent infections, we repeated the vector-borne transmission with a new virus stock (Fig. 2A, Stock 2—Black). Deep sequencing of this stock shows that it contains four synonymous consensus changes (Supplementary Table S3) and additional minority variants compared to our ZIKV plasmid and Stock 1 (Supplementary Table S3). While we were unable to obtain amplicons from mosquitoes with this infection, we observed that these variants were maintained in Mouse C, which was infected via mosquito bite from ZIKV Stock 2. Mouse C had five additional consensus changes, four of which became fixed (Supplementary Table S2). Of these new consensus changes, a non-synonymous variant in the NS3 protein (G5962T, E1940D, 99.9%) was found at 12.1 per cent in the viral stock. The enrichment of this variant in Mouse C could suggest that it was under positive selection in the mosquito or in the mammalian host. It could also be the result of the stochastic nature of transmission, which has tight bottlenecks. Finally, to rule out any bias in our analysis due to the variant frequency threshold we set (>1%), we addressed the richness of each sample (number of minority variants per sample) at different variant frequency thresholds (Supplementary Fig. S2). We did not find any significant differences in richness dynamics between samples at any threshold indicating that our frequency threshold is not introducing bias into our analysis. Taken together, these data suggest that either ZIKV viral diversity is restricted in the mosquito host and expanded in Ifnar−/−/C0 mice, independent of inoculation route, or diversity is generated in the mosquito and is transmitted to the mouse.

Interestingly, in these analyses we observed that most minority variants were generally present at low frequencies (<30%) in mice, similar to what has been seen in human isolates (Metsky et al. 2017). Given these similarities to natural infections, we asked whether infections in mice can begin to recapitulate the evolutionary trajectories found in nature during the most recent outbreaks. When we compared minority variants observed in our studies with consensus differences between the Ugandan (MR766), and Cambodian, Brazilian, and Puerto Rican
ZIKV strains, we found eighty-six unique minority variants that corresponded to exact nucleotide differences between MR766 and each of these strains (Fig. 2E and F, Supplementary Table S4). Of these minority variants, eighty-two were common differences between MR766 and the Cambodian, Brazilian, and Puerto Rican strains, two were unique differences between MR766 and the Brazilian and Puerto Rican strains, one was a difference between MR766 and only the Brazilian strain, and one was a difference between MR766 and the Cambodian strain. Moreover, we observed that the majority of these variants arose in the spinal cord of mice potentially highlighting the importance of certain tissues for virus emergence (Fig. 2G). Finally, in addition to point mutations in currently circulating strains as compared to MR766, we also observed the emergence of 17 variants that have been described as minority variants or common variants in nature (Metsky et al. 2017; Collins et al. 2019) (Supplementary Table S5).

Taken together, these data suggest that ZIKV infection in Ifnar−/− can mimic various aspects of natural ZIKV infections and represents a relevant model to study how arboviruses evolve.

3.3 ZIKV inoculation route impacts organ-specific diversity and infectious virus production

To understand how different organ microenvironments shape the genetic composition of ZIKV, we analyzed ZIKV genetic diversity in individual organs across all mice (Fig. 3A). We observed that the brain, spinal cord, and liver generated more variants (single nucleotide variant (SNVs) > 1%) in comparison to kidney, reproductive tract, and serum. In addition, in the brain, spinal cord, and kidney there were more high-frequency variants generated in the vector-borne transmission than with needle inoculation (Fig. 3A, white and black dots, respectively). Given this organ-specific richness, we expected to find host and organ-specific variants during infection. Of the variants identified in Stock 1, five (at positions 9029, 9033, 9037, 9064, and 9837) were present in both the mosquito and needle-inoculated mice but were absent (or below our limits of detection) in all liver samples (Supplementary Fig. S3). Further, two variants that were not present in the stock were found in both mosquito- and needle-inoculated mice but were enriched (two–three-fold)

Figure 2. Total ZIKV SNV and evolutionary analysis. The ZIKV genome was amplified in three amplicons by PCR and each amplicon was used for deep sequencing analysis and alignment to the MR766 genome. Variants below 1 per cent frequency were considered background and were removed from the analysis. (A) Viral variants present in the MR766 stocks that were used for mosquito feeding and needle infections. (B) Total viral variants present in mice subcutaneously infected with ZIKV MR766 (N = 7). (C) ZIKV variants present in bodies of infected mosquitoes 14 days post feeding. After feeding on mice, mosquitoes were ground, RNA extracted, and the ZIKV genome amplified by PCR. (D) Total viral variants present in mice infected by mosquito bite. Data represent mice from three independent experiments (N = 3). Schematics under A–D depict the ZIKV genome and PCR amplicons used for sequencing. (E) Schematics showing genomic locations of the MR766 minority variants corresponding to changes between MR766 and the Cambodian, Brazilian, and Puerto Rican strains. (F) Frequencies of MR766 minority variants corresponding to changes between MR766 and Cambodian, Brazilian, and Puerto Rican strains. (G) Organ frequency where changes between MR766 and Cambodian, Brazilian, and Puerto Rican strains are found. Gray bar = whole mosquitoes, black bar = needle-inoculated mice, and white bar = mosquito-transmitted mice.
in the bone marrow of the mice (NS5 C10337A, R3411S; NS5 T10340G, and Y3412D). These two groups of mutations suggest that tissue-specific variants are generated within the host over the course of a ZIKV infection.

We then hypothesized that although we did not see changes in viral RNA abundance in these organs, the different inoculation routes and, potentially, the different viral populations could influence infectious virus particle production. We
quantified infectious ZIKV by plaque assay and found that viruses isolated from kidney and liver showed a reduced trend in the number of infectious particles generated, and the inoculation route significantly impacted the number of infectious particles isolated from the reproductive tract (Fig. 3B). In line with these results, we observed organ-specific changes in viral-specific infectivity suggesting that the organ microenvironment and viral diversity may contribute to ZIKV infection (Fig. 3C).

Given the relative changes in organ-specific variants, viral titers, and viral genome levels, we quantified both the genetic diversity and genetic distance of each mosquito, mouse organ, and stock sample. We focused on samples that successfully had the complete coding region amplified and sequenced and used only minority variants that were present at a position with 200× coverage and a frequency of 3% or greater for our calculations. When calculating the pairwise genetic distance of all samples, and performing multidimensional scaling on the distance measurements, we found that samples infected with Stock 1 clustered together, while those from Stock 2 clustered independently, as expected (Fig. 4A), with consensus changes and minority variants that were unique to each stock and experiment. To determine whether the trajectories of the virus were tissue specific, we focused on the degree to which each tissue sample differentiated from its respective stock virus using the calculated Euclidean distance measurements (Fig. 4B). All tissues had viral populations that diversified away from their respective stock virus over the course of the infection, with the virus populations in the brain, spinal cord, and reproductive tract being more similar to the stock virus and liver being the most dissimilar (Fig. 4B). Further, liver samples from mice inoculated with Stock 1 virus tend to cluster more closely regardless of inoculation route as observed in Fig. 4A, which suggests that while the virus populations in the liver are the most dissimilar to the stock virus, they evolve in the same direction, away from the virus populations in other tissues or mosquitoes (Fig. 4A and B). Together, these data suggest that even when infected with a genetically distant stock, organs shape these viral populations towards a genetic equilibrium.

To quantify how the inoculation route impacts the genetic diversity observed in the virus populations of each tissue, we calculated the Shannon entropy from the complete coding region (Fig. 4C-E). We found that the inoculation route did not significantly affect the observed genetic diversity. However, organs collected from mice inoculated via needles typically yielded lower genetic diversity than organs from mice infected via mosquito bite, except for the liver (Fig. 4C and D), suggesting that mosquito bite may influence organ-specific viral diversity in a similar manner to what has been seen previously with rhesus macaques (Dudley et al. 2017). When we looked at each organ individually from mice infected via needle inoculation, we found that the liver produced a highly diverse viral population, followed by the spinal cord and brain (Fig. 4D). Similarly, virus populations in organs from mice infected via a mosquito bite were more diverse in the spinal cord, brain, and liver and the least diverse in the kidney, spleen, and serum (Fig. 4D). This was particularly interesting as we hypothesized that Mouse C, which was infected with the high diversity Stock 2, would maintain this high diversity and be different from those inoculated with Stock 1. However, we found that organs from Mouse C had lower diversity than Stock 2 and had similar levels of diversity as Mouse A and B, inoculated with Stock 1 (Fig. 4C). These data show that independent infections with two genetically distinct inputs result in the same outcome after vector-borne transmission, suggesting that regardless of input, virus diversity is reshaped by the mosquito, transmission bottlenecks, and/or the mouse during vector-borne infection. It has been shown that the read depth and samples with low RNA quantities could impact viral diversity (Zhao and Illingworth 2019). While the concentration of RNA in our samples is well above the quantity shown to affect diversity, we performed rigorous analyses to ensure RNA concentration was not impacting virus diversity. To do this we (1) calculated the nucleotide diversity as outlined in Zhao and Illingworth (2019) and obtained values nearly identical to ones obtained using Shannon entropy (Supplementary Fig. S4); (2) directly compared the viral RNA titers with organ diversity and found that these two parameters do not correlate (Fig. 4E); and (3) subsampled our sequencing reads and found that subsampling did not impact viral diversity dynamics (Supplementary Fig. S5). Finally, when looking at the Shannon entropy of each nucleotide position across the genome, we identified potential diversity hotspots in the capsid, NS1, NS2A/B, and NS5 coding regions (Fig. 4F). Altogether, these data indicate that organ microenvironments are the main drivers for viral diversification and that inoculation route plays a minor role, at least 7 days post infection.

3.4 ZIKV infection generates tissue-specific defective viral genomes

Flavivirus RNA replication can lead to the production of DVGs (Vignuzzi and Lopez 2019; Yang et al. 2019), which can play an important role in viral interference (Brinton 1983) and persistence (Brinton 1982; Lancaster et al. 1998; Li et al. 2011). Given this, we hypothesized that ZIKV also produces DVGs and that different organ microenvironments could influence DVG production. To address this, we mapped the coordinates of gaps within sequencing reads, corresponding to deleted regions in the genome, and identified candidate ZIKV DVGs within the three PCR amplicons of the coding sequence (Figs 5 and 6). To determine if there were specific hotspots implicated in DVG generation within the ZIKV genome, we looked at which regions of the coding sequence harbored the most deletions. Interestingly, we found that capsid, pr, M, and NS4B genes had the most start sites, while NS1, NS2A/B, and NS5 contained predominantly DVG end sites (Fig. 5A and B).

When we looked at DVGs in each sample, we observed that the viral input stock contained a high diversity of different DVG types within PCR amplicons 1 and 3, as well as a high abundance of DVG gap-spanning reads (Fig. 5C and D). While we did not observe statistical differences in the abundance of DVGs by the inoculation route, we did observe organ specific differences in both DVG diversity and abundance. Most organ samples either completely lacked or had a low abundance of DVGs in the PCR 2 amplicon (Fig. 5C and D). Brain, spinal cord, and bone marrow samples had the highest diversity of DVG types. Surprisingly, we did not detect DVGs in most of the liver samples (Fig. 5C and D). These data are in contrast to what was seen for viral SNV diversity (Fig. 4) and may implicate DVGs as playing a role in ZIKV dissemination or pathogenesis. To ensure that read depth was not impacting our results, we performed our DVG analysis on subsampled PCR amplicon libraries (Supplementary Fig. S6). While subsampling of the alignment files did impact the number of unique DVGs identified and the abundance (FPKM) of these DVGs, the tissue-specific DVG dynamics were maintained across the five subsampling frequencies (Supplementary Fig. S6A). Further, we observed strong positive correlations between the subsampled and non-subsampled FPKMs of individual DVG species, indicating that
sampling depth was not changing the tissue-specific DVG dynamics (Supplementary Fig. S6B).

While globally we found a large number of DVGs in both the input stocks and the samples, we were interested in whether unique and tissue-specific DVGs were produced during ZIKV infection. We analyzed each mouse sample for DVGs that were unique to one tissue of one mouse and, therefore, likely to be generated de novo (Fig. 6A). We hypothesized that DVGs that have a potential advantage within an infection are more likely to be found at higher abundances and shared across multiple samples. 

Figure 4. Organ-specific diversity during ZIKV infection. Euclidean distance and Shannon entropy were calculated using sequence data from the complete coding sequence. A frequency threshold of 3 per cent was used for all analyses. (A) Multi-dimensional scaling plot showing all-versus-all Euclidean distance measurements of each organ sample. (B) Genetic distance of each sample from its corresponding stock control separated by tissue type. Data represent the median with inter-quartile range. Black dots represent outliers. Kruskal–Wallis with Dunn’s post-test, considering a type-I error. (C) Organ-specific Shannon entropy by inoculation route. Data represent the average and standard deviation. Kruskal–Wallis with Dunn’s post-test, considering a type-I error. (D) ZIKV genetic diversity by transmission route. Data represent the median with inter-quartile range. Kruskal–Wallis with Dunn’s post-test, considering a type-I error. (E) Correlation between ZIKV RNA genome copies and Shannon entropy. Color code and legend correspond to all figures. (F) Shannon entropy calculated for each nucleotide across the complete ZIKV coding region. All statistical analyses can be found in Supplementary Tables S6–S8.
Therefore, we selected the most abundant DVG type within each tissue of each mouse and examined whether it was unique or shared with other samples. Out of the sixty-three identified abundant DVGs, twenty-three of these were present in multiple mice and across different organ tissues (Fig. 6B and C) and varied in their abundance (Fig. 6D). The high variability in DVG populations in each tissue of the individual mice suggests that the generation and abundance of DVGs is stochastic within each
host. Taken together, these data show that organ microenvironments impact not only ZIKV SNV diversity but also the production of ZIKV DVGs, which could be acting as regulators in the viral life cycle.

4. Discussion

Viral infections are dynamic, involving multiple cell types, organ systems, and transmission routes. In particular, arboviruses are transmitted by insect vectors to mammalian hosts, yet we understand little of these complex infections or how vector-borne transmission influences viral diversity or emergence. In this work, we aimed to study how inoculation route and different organ microenvironments defined ZIKV pathogenesis in mice through the modulation of viral diversity, DVG production, and the emergence of new viral variants. To address this, we infected Ifnar−/− mice with MR766 ZIKV either by the bite of a ZIKV-infected Ae.aegypti mosquito or by needle inoculation. We found that the inoculation route had little impact on the course of the disease or on viral RNA levels in isolated organs. When we analyzed the viral populations present in our controls, mosquitoes, and individual mice during ZIKV infection, we were surprised that only a handful of high frequency (>5%) and consensus changes (>50%) could be found. These data are in contrast to what was seen in rhesus macaques where a number of high-frequency variants were identified in both mosquitoes and the plasma of monkeys (Dudley et al. 2017). There are many explanations for this discrepancy, such as vector and mammalian host specificity, the viral strain used and its initial diversity, and the use of immunocompromised animals. However, in this case, it would suggest that the immune response in the monkeys would contribute to selection pressure, increasing viral diversity and leading to changes in the viral populations. It will be interesting to use immunocompetent mouse models to look specifically at the role of interferon and other host pathways on viral populations, as well as continuously passage these viruses through mosquito and mouse to watch how viral populations evolve.

Figure 6. Identification of abundant DVGs during ZIKV infection in mice. (A) Distribution of DVGs found to be shared or unique within mice. (B) Number of different mice, mosquito, and stock samples with the given DVG annotated on the x-axis at the bottom of panel D. (C) Number of different tissues with given DVGs. Size of the circle indicates the number of tissue samples with the identified DVG. (D) Abundance of unique DVGs in gap-spanning reads per kilobase per million with a given sample.
Nonetheless, one aspect particularly intriguing in this study is that our data resembled data from natural infections with the identification of a limited number of variants, and most minority variants present at around 30 per cent of the population, or below (Metsky et al. 2017). Moreover, the MR766 African lineage of ZIKV generated minority variants that corresponded to exact nucleotide changes found in currently circulating Asian and American strains. In addition, we identified minority variants that were found in human isolates, suggesting that ZIKV may be taking a similar evolutionary trajectory in mice and providing evidence that viral infections in mouse models can be used to study the evolution of the virus in human infections. In particular, as this study was conducted at 7 days post infection, where there is a limited number of strain-level variants in the blood of mice, it will be necessary to look at earlier days post transmission to understand how these strain-level changes may be impacting ZIKV infection.

To understand how individual organs impact virus diversity and evolution, we characterized the viral populations in specific organ microenvironments. Each organ had distinct populations of viral variants, suggesting that viral populations may be impacted by organ and cell-specific mechanisms. This has been seen previously with multiple viruses where organ-specific bottlenecks have been studied extensively (Pfeiffer and Kirkegaard 2006; Kuss, Etheredge, and Pfeiffer 2008; Forrest et al. 2012; Coffey et al. 2013; Xiao et al. 2017; Riemersma et al. 2019; Warnbrood et al. 2019), supporting that this may be the case in our model too. Interestingly, we observed high genetic diversity in the brain and spinal cord, sites of viral replication with the highest viral RNA and infectious particle loads. Moreover, the liver was an organ that generated the highest viral diversity yet had the lowest viral RNA and infectious viral titers, suggesting that this diversity may be influencing viral replication. Interestingly, we found a significant difference in viral titers in the reproductive tract between needle and mosquito inoculated mice, suggesting that transmission route can impact infection in certain organs. We also found that the serum, which would contain the transmittable population of virus, lacked high levels of diversity, in contrast to what was seen in monkeys (Dudley et al. 2017). However, similar to what was found by Dudley et al., our data support that vector-borne transmission does impact viral populations and can influence selection. The question remains as to why vector-borne transmission would specifically influence viral diversity 7 days after a mosquito bite. One explanation could be that early and late inflammatory responses from the mosquito bite are exerting selection pressure impacting diversity and, as our data suggest, each individual organ may possess intrinsic mechanisms to regulate viral populations. How this diversity contributes to ZIKV fitness, pathogenesis, and transmission remains to be explored.

In addition to viral diversity, we also addressed the production of DVGs during ZIKV infection, something that, to our knowledge, had not yet been investigated. However, our analysis of ZIKV DVGs was limited by the fact that the genome was amplified in three separate amplicons such that DVGs with large deletions spanning across the genome would not be captured by our approach. Nevertheless, we found that the brain and spinal cord appeared to generate more DVGs—and the most diverse—as compared to other organs such as the liver, which produced the smallest number of DVGs. These data are intriguing and suggest that different organs may influence viral diversity and DVG production as a way to control the virus, or that ZIKV may take advantage of these organ-specific processes for its benefit. Interestingly, DVG production was not impacted by transmission route as it was for viral diversity. This could imply that DVG production and genetic drift are influenced by distinct factors during infection. A detailed analysis of ZIKV DVGs produced in humans and other animal models will be essential to understand the role of these truncated genomes in the viral life cycle.

In summary, these studies begin to define the evolutionary landscape and trajectories of ZIKV during vector transmission and mammalian infections. In particular, we have established an animal transmission model that can recapitulate aspects of ZIKV infection seen in humans, thus providing a powerful system to study arbovirus evolution in the lab. Using this system, we have shown that vector-borne transmission influences viral diversity and that ZIKV diversity and DVG production are impacted by the infection of various organs in immunocompromised mice. These studies of arbovirus evolution during vector-borne transmission are essential to understand the molecular mechanisms ZIKV and other arboviruses use for infection and disease, and provide a framework to study the transmission of virus variants and organ-specific forces that drive viral diversity and DVG production.

Acknowledgements

We thank all members of the Stapleford Lab for helpful comments on this manuscript. M.G.N was supported by a Jan Vilcek/David Goldfarb Fellowship from the New York University Department of Microbiology, New York University. M.V.R. and K.E.E.J were supported in part by the Public Health Service Institutional Research Training (award T32 AI007180).

Data availability

Sequencing data that support the findings of this study have been deposited in the Sequence Read Archive (SRA) under BioProject ID PRJNA589089.

Supplementary data

Supplementary data are available at Virus Evolution online.

Conflict of interest: None declared.

References

Agarwal, A. et al. (2016) ‘Mosquito Saliva Induced Cutaneous Events Augment Chikungunya Virus Replication and Disease Progression’, Infection, Genetics and Evolution, 40: 126–35.

Aldunate, F. et al. (2017) ‘Evidence of Increasing Diversification of Zika Virus Strains Isolated in the American Continent’, Journal of Medical Virology, 89: 2059–63.

Aliota, M. T. et al. (2018) ‘Molecularly Barcoded Zika Virus Libraries to Probe In Vivo Evolutionary Dynamics’, PLoS Pathogens, 14: e1006964.

Brinton, M. A. (1982) ‘Characterization of West Nile Virus Persistent Infections in Genetically Resistant and Susceptible Mouse Cells. I. Generation of Defective Nonplaquing Virus Particles’, Virology, 116: 84–98.

(1983) ‘Analysis of Extracellular West Nile Virus Particles Produced by Cell Cultures from Genetically Resistant and Susceptible Mice Indicates Enhanced Amplification of Defective Interfering Particles by Resistant Cultures’, Journal of Virology, 46: 860–70.
Collins, N. D. et al. (2019) ‘Inter- and Intra-Lineage Genetic Diversity of Wild-Type Zika Viruses Reveals Both Common and Distinctive Nucleotide Variants and Clusters of Genomic Diversity’, Emerging Microbes and Infections, 8: 1126–38.

Coyne, C. B., and Lazear, H. M. (2016) ‘Zika Virus - Reigniting the TORCH’, Nature Reviews Microbiology, 14: 707–15.

Delatorre, E., Mir, D., and Bello, G. (2017) ‘Tracking the Origin of the NS1 A188V Substitution Responsible for Recent Enhancement of Zika Virus Asian Genotype Infectivity’, Mem Inst Oswaldo Cruz, 112: 793–79.

Dobin, A. et al. (2013) ‘STAR: Ultrafast Universal RNA-Seq Aligner’, Bioinformatics, 29: 15–21.

Dudley, D. M. et al. (2017) ‘Infection via Mosquito Bite Alters Zika Virus Tissue Tropism and Replication Kinetics in Rhesus Macaques’, Nature Communications, 8: 2096.

Fauci, A. S., and Morens, D. M. (2016) ‘Zika Virus in the Americas–Yet Another Arbovirus Threat’, New England Journal of Medicine, 374: 601–4.

Forrester, N. L. et al. (2012) ‘Vector-Borne Transmission Imposes a Severe Bottleneck on an RNA Virus Population’, PLoS Pathogens, 8: e1002897.

Gill, C. M. et al. (2019) ‘Five Emerging Neuroinvasive Arboviral Diseases: Cache Valley, Eastern Equine Encephalitis, Jamestown Canyon, Powassan, and Usutu’, Seminars in Neurology, 39: 419–27.

Hastings, A. K., and Fikrig, E. (2017) ‘Zika Virus and Sexual Transmission: A New Route of Transmission for Mosquito-Borne Flaviviruses’, The Yale Journal of Biology and Medicine, 90: 325–30.

Hermance, M. E., and Thangamani, S. (2017) ‘Powassan Virus: An Emerging Arbovirus of Public Health Concern in North America’, Vector-Borne and Zoonotic Diseases, 17: 453–62.

Kim, D. et al. (2013) ‘TopHat2: Accurate Alignment of Transcriptomes in the Presence of Insertions, Deletions and Gene Fusions’, Genome Biology, 14: R36.

Kuss, S. K., Etheredge, C. A., and Pfeiffer, J. K. (2008) ‘Multiple Host Barriers Restrict Poliovirus Trafficking in Mice’, PLoS Pathogens, 4: e1000882.

Lancaster, M. U. et al. (1998) ‘Characterization of Defective Viral RNA Produced during Persistent Infection of Vero Cells with Murray Valley Encephalitis Virus’, Journal of Virology, 72: 2474–82.

Langmead, B., and Salzberg, S. L. (2012) ‘Fast Gapped-Read Alignment with Bowtie 2’, Nature Methods, 9: 357–9.

Le Coupene, A. et al. (2013) ‘Aedes Mosquito Saliva Modulates Rift Valley Fever Virus Pathogenicity’, PLoS Neglected Tropical Diseases, 7: e2237.

Li, D. et al. (2011) ‘Defective Interfering Viral Particles in Acute Dengue Infections’, PLoS One, 6: e19447.

Liu, Y. et al. (2017) ‘Evolutionary Enhancement of Zika Virus Infectivity in Aedes aegypti Mosquitoes’, Nature, 545: 482–6.

Lucchesi, G., and Kanduc, D. (2016) ‘Zika Virus and Autoimmunity: From Microcephaly to Guillain-Barre Syndrome, and Beyond’, Autoimmunity Reviews, 15: 801–8.

Magalhaes, T. et al. (2018) ‘Mosquito-Borne and Sexual Transmission of Zika Virus: Recent Developments and Future Directions’, Virus Research, 254: 1–9.

Mersky, H. C. et al. (2017) ‘Zika Virus Evolution and Spread in the Americas’, Nature, 546: 411–5.

Metsky, J. J., and Diamond, M. S. (2017) ‘Zika Virus Pathogenesis and Tissue Tropism’, Cell Host & Microbe, 21: 134–42.

Noval, M. G. et al. (2019) ‘Evolution-Driven Attenuation of Alphaviruses Highlights Key Glycoprotein Determinants Regulating Viral Infectivity and Dissemination’, Cell Reports, 28: 460–71.e465.

Parvez, M. K., and Parveen, S. (2017) ‘Evolution and Emergence of Pathogenic Viruses: Past, Present, and Future’, Interception, 60: 1–7.

Pfeiffer, J. K., and Kirkegaard, K. (2006) ‘Bottleneck-Mediated Quasispecies Restriction during Spread of an RNA Virus from Inoculation Site to Brain’, Proceedings of the National Academy of Sciences of the United States of America, 103: 5520–5.

Pingen, M. et al. (2016) ‘Host Inflammatory Response to Mosquito Bites Enhances the Severity of Arbovirus Infection’, Immunity, 44: 1455–69.

Parvez, M. K., and Parveen, S. (2017) ‘Evolution and Emergence of Pathogenic Viruses: Past, Present, and Future’, Interception, 60: 1–7.

Pfeiffer, J. K., and Kirkegaard, K. (2006) ‘Bottleneck-Mediated Quasispecies Restriction during Spread of an RNA Virus from Inoculation Site to Brain’, Proceedings of the National Academy of Sciences of the United States of America, 103: 5520–5.

Pingen, M. et al. (2016) ‘Host Inflammatory Response to Mosquito Bites Enhances the Severity of Arbovirus Infection’, Immunity, 44: 1455–69.

Schneider, B. S. et al. (2006) ‘Potentiation of West Nile Encephalitis by Mosquito Feeding’, Viral Immunology, 19: 74–82.

et al. (2010) ‘Aedes aegypti Saliva Alters Leukocyte Recruitment and Cytokine Signaling by Antigen-Presenting Cells during West Nile Virus Infection’, PLoS One, 5: e11704.

Schwarz, M. C. et al. (2016) ‘Rescue of the 1947 Zika Virus Prototype Strain with a Cytomegalovirus Promoter-Driven cDNA Clone’, mSphere, 1.

Shi, W. et al. (2016) ‘Increasing Genetic Diversity of Zika Virus in the Latin American Outbreak’, Emerging Microbes & Infections, 5: e68.

Vignuzzi, M., and Lopez, C. B. (2019) ‘Defective Viral Genomes Are Key Drivers of the Virus-Host Interaction’, Nature Microbiology, 4: 1075–87.

Warmbrod, K. L. et al. (2019) ‘Viral RNA-Dependent RNA Polymerase Mutants Display an Altered Mutation Spectrum Resulting in Attenuation in Both Mosquito and Vertebrate Hosts’, PLoS Pathogens, 15: e1007610.

Weaver, S. C. (2006) ‘Evolutionary Influences in Arboviral Pathogenesis: Past, Present, and Future’, Current Topics in Microbiology and Immunology, 299: 285–314.

et al. (2018) ‘Zika, Chikungunya, and Other Emerging Vector-Borne Viral Diseases’, Annual Review of Medicine, 69: 395–408.
Xiao, Y. et al. (2017) 'Poliovirus Intrahost Evolution Is Required to Overcome Tissue-Specific Innate Immune Responses', Nature Communications, 8: 375.

Yang, Y. et al. (2019) 'The Antiviral and Antitumor Effects of Defective Interfering Particles/Genomes and Their Mechanisms', Frontiers in Microbiology, 10: 1852.

Yuan, L. et al. (2017) 'A Single Mutation in the prM Protein of Zika Virus Contributes to Fetal Microcephaly', Science, 358: 933–6.

Zhao, L., and Illingworth, C. J. R. (2019) 'Measurements of Intrahost Viral Diversity Require an Unbiased Diversity Metric', Virus Evolution, 5: vey041.