THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

CHARACTERIZATION OF GENETIC MUTATIONS IN EPIDEMIC STRAINS OF ZIKA
VIRUS

SEAN C. HORAN
SPRING 2019

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree in
Biochemistry and Molecular Biology
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Joyce Jose, PhD
Assistant Professor of Biochemistry and Molecular Biology
Thesis Supervisor

Santhosh Girirajan, PhD
Associate Professor of Genomics
Honors Adviser

* Signatures are on file in the Schreyer Honors College

ABSTRACT

Zika virus (ZIKV) was initially discovered in Uganda nearly three quarters of a century ago. Until recently, the virus caused a few mild infections in humans. Genetic analysis performed on the growing number of sequenced strains has revealed the virus diverged into two distinct lineages, African and Asian. The recent rise in the prevalence of Asian lineage Zika virus infections associated with serious neurological disorders in French Polynesia and Brazil (2013 and 2015 respectively) has raised concerns regarding public health. We hypothesize mutations in pre-epidemic strains have contributed to the increased virulence of the epidemic strains associated with neurological disease. We used reverse genetic analysis to identify candidate mutations that were the most likely to cause virulence on par with epidemic strains. We individually introduced these candidate mutations into the African pre-epidemic MR766 strain (1947) and characterized infection in mosquito vector and human neuronal cells. We also substituted the NS5 polymerase region of the epidemic Senegal strain (1984) into the pre-epidemic strain to see how multiple mutations affected viral infection *in vitro*. We expected the mutant constructs to cause infections that were equally or more severe than the pre-epidemic strain. However, in general, we found mutant constructs caused infections that were less severe than the pre-epidemic wild type virus. Our results suggest regions of the ZIKV genome coevolved to produce the more virulent strains associated with severe neurological disease.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

INTRODUCTION

Initial Discovery of Zika Virus

Zika virus (ZIKV) was initially discovered in the Zika Forest of Uganda in 1947 while researchers were attempting to isolate Yellow Fever from the blood of sentinel rhesus macaques. At the time, ZIKV was not believed to be capable of infecting humans. However, after serosurveys of Ugandans showed the presence of anti-Zika antibodies in a significant number of individuals tested and three confirmed human infections, Zika proved to be a cause of human disease (1, 2). Although ZIKV was shown to cause disease, symptoms were mild (low-grade fever, fatigue, etc.) and subsided relatively quickly without any outstanding maladies. Over the past half a century fewer than 15 cases of Zika virus infection were reported worldwide, prompting little concern for public health (1).

Zika Virus Transmission and Infection

ZIKV is a member of the *Flavivirus* genus and *Flaviviridae* family, similar to well-known West Nile, dengue and yellow ever viruses. These viruses, including Zika, are arboviruses—they are transmitted via arthropod vectors. ZIKV is primarily transmitted to humans via mosquito vectors. *Aedes aegypti* is the primary mosquito vector for Zika virus (3), although other species of mosquitoes have been shown to carry the virus (4). Zika can also be transmitted in a vector-independent way through sexual contact as well as perinatally (5). The transmission of the virus via sexual contact is especially alarming, as it is possible to pass on the

virus after the infection has been cleared by the immune system. A study using a mouse model showed that infectious ZIKV was present in semen of infected male mice up to three weeks post-infection and viral RNA was detectable up to 23 days post infection (5). Asymptomatic infection with the potential for transmission independent of mosquitos has raised concerns for public health, as infected individuals can potentially spread the virus unknowingly.

Growing concerns for public health have been exacerbated by large scale outbreaks seen in countries outside of Africa and Asia, rapidly spreading to Pacific island nations, South and Central Americas and the United States (1). In addition, neurological disorders have been observed in association with Zika virus infection, chief among which have been fetal microcephaly—a condition that limits brain development—and Guillian-Barré syndrome—a condition that causes paralysis. One study characterized the infection in neonates and found microcephaly was only one of several neurological disorders observed (6). Although no causal relationships have been established between ZIKV infection and these conditions, the positive correlation observed between infection and these conditions has prompted the World Health Organization to declare a Public Health Emergency of International concern in 2016 (7). The need for research to gain insight on the virulence of ZIKV in this context, therefore, cannot be overstated.

Genome Structure and Evolution

ZIKV is comprised of a positive-sense single-stranded RNA genome of 10.8 kilobases and three structural proteins. The genome contains a single open reading frame, flanked by 5' and 3' untranslated regions, that encodes a polyprotein, which is then processed into ten

individual viral proteins by the viral and host proteases (8). Three of the resulting proteins are structural: capsid (C), precursor membrane (prM) and envelope (E). The prM (cleaved into pr and M by host furin) and E are envelope glycoproteins arranged on the viral lipid bilayer originating from the ER membrane. The other seven proteins are nonstructural, which include the NS3 protease and NS5 RNA-dependent RNA polymerase among others.

Previous phylogenetic characterization shows the virus diverged into two distinct lineages, African and Asian (9). Each lineage has pre-epidemic and epidemic strains, which are strains not associated with widespread infection and those that are, respectively. It is the epidemic strains that are associated with increases in the number of cases of ZIKV infection as well as infection severity—those associated with neurological conditions, such as African lineage Senegal (1984) and Asian lineage Thailand (2014) (10). Conversely, pre-epidemic Uganda (African lineage, 1947) and Malaysia (Asian lineage, 1966) strains have caused fewer and milder infections, often going unnoticed (11). Caused by the lack of proof-reading activity of the RNA-dependent RNA polymerase, genetic differences between pre-epidemic and epidemic strains of both lineages are of particular interest—understanding how accumulated mutations lead to different virulence capabilities is required to develop antiviral targets and therapeutics to mitigate the severe neurological disorders associated with infection.

Experimental Design

We hypothesize that accumulated mutations over time are the reason for the difference in infection severity between pre-epidemic and epidemic ZIKV strains. We used reverse genetic analysis intending to identify mutations that confer higher virulence, as well understand how

these identified mutations lead to higher virulence. We expected ZIKV mutants incorporating the epidemic strain mutations to show higher virulence parameters than wild type pre-epidemic strains that do not contain mutations. We tested the virulence parameters by characterizing infection of mammalian and mosquito cells via fluorescence microscopy, quantifying infection via viral titers, and the number of RNA molecules produced via qRT-PCR.

MATERIALS AND METHODS

Sequence Analysis

Multiple amino acid sequence alignments were performed using ClustalOmega to identify genetic differences between pre-epidemic Uganda (MR766) and epidemic Senegal (SEN/1984/41525-DAK) and epidemic Honduras (R103451) strains. To understand the tolerance of identified mutations, these ZIKV strains' sequences were also aligned to other flaviviruses: tick-borne encephalitis virus (TBEV), yellow fever virus (YFV), Dengue virus (DENV1), Saint Louis encephalitis virus (SLEV), Japanese encephalitis virus (JEV), and Murray Valley encephalitis virus (MVEV).

Viruses and Cells

Vero, HEK-293T, and SHSY5 cells were maintained in Dubleco's MEM (Fisher Scientific) with 10% FBS (Hyclone) at 37 °C in the presence of 5% CO_2 in an incubator. C6/36 cells were maintained at 30 °C. ZIKV strains were obtained from BEI Resources. An infectious cDNA clone of MR766 (BEI number NR-50065) served as a pre-epidemic virulence control. This clone was also used as the backbone for substituting the SEN/1984/41525-DAK (BEI number NR-50338) NS5. Senegal NS5 insert was synthesized from viral RNA using RTPCR.

Ligation

Purified cDNA insert 100 ng was combined with 30 ng of purified vector, along with 1 µL of 10X T4 DNA Ligase Buffer and 0.5 µL of T4 DNA Ligase (New England Biolabs). Reactions were incubated at 4 °C overnight and transformed into *E. coli* MC1061 competent cells.

Bacterial Transformation

Two µL of DNA was added to 20 µL of MC1061 competent cells and incubated on ice for 30 minutes. Cells were heat shocked for 90 seconds at 42 °C and incubated on ice for 5 minutes. LB broth (300 µL) was added to the transformation mixture and incubated on the shaker at 180 rpm at 33 °C for 90 minutes. Cells were plated onto LB plates containing ampicillin and incubated at 37 °C overnight.

Verification of Clones by Colony PCR

Reactions (15 µL) were set up in PCR tubes with 7.5 µL of OneTaq 2X Master Mix, 5.5 µL of sterile water and 1 µL of the forward and reverse primers specific to the insert cDNA of interest. Colonies from ligation plates were mixed into each reaction. Reactions were run for 28 cycles of 98 °C melting for 30 seconds, 55 °C annealing for 30 seconds, 68 °C extension for 90 seconds.

Plasmid Extraction by Miniprep

Colonies from transformation plates were inoculated into individual tubes containing 6 mL of Luria Broth and ampicillin (1 µL AMP/mL LB). Cultures were incubated on a shaker at 37 °C overnight. Cells were pelleted in the centrifuge for 10 minutes and the media was poured off. Cells were resuspended in 250 µL of P1 buffer and lysed with the addition of 250 µL of P2 lysis buffer. Lysis was neutralized with 350 µL of N3 neutralizing buffer and each reaction was centrifuged to pellet out cell debris. The supernatant was transferred to a Qiagen Spin Miniprep Column and centrifuged for 30 seconds at 13,000 g. The flow-through was discarded and the column was treated with 500 µL PB binding buffer and centrifuged at 13,000 g for 30 seconds. The flow-through was discarded and the column was washed with 750 µL PE wash buffer and centrifuged at 13,000 g for 30 seconds. The column was transferred to a clean centrifuge tube and DNA was eluted from the column by treating the column with 50 µL of EB elution buffer and centrifuging at 13,000 g for one minute. A QIAprep Spin Miniprep Kit was utilized for plasmid extraction (Qiagen).

Agarose Gel Electrophoresis

Two µL of DNA and 1 µL of 6X loading dye were mixed and loaded into wells on a 0.8% (w/w) agarose gel containing ethidium bromide. Five µL of 1 kb DNA ladder (New England Biolabs) was loaded for MW reference. The gel was placed in 1X TAE buffer and run at 120V for 15 minutes. Gel bands were visualized using a UV transilluminator.

Gel Purification of DNA

Low-melting agarose gel (0.8% w/w) and 1X TAE buffer were cooled to 4 °C and loaded with DNA samples and 1kb DNA ladder (New England Biolabs). The gel was run at 120V for 30 minutes. The gel was viewed with a UV transilluminator and desired bands were excised with a razor blade and transferred to a 1.5 mL Eppendorf tube. Capture buffer (500 µL) was added and the sample was heated at 50 °C for 10 minutes with intermittent mixing. Once the gel piece was fully melted, the resulting solution was loaded onto a GFX column (GE Healthcare) and incubated for 5 minutes. The column was centrifuged at 10,000g for 1 minute and the flow-through was discarded. The column was washed with 500 µL of wash buffer and centrifuged at 10,000g for 1 minute. The flow-through was discarded and the column was centrifuged at 10,000g for 1 minute to dry the column. The column was transferred to a new 1.5 ml microcentrifuge tube and 35 µL of warm elution buffer (50 °C) was added and stood for 1 minute. The DNA was eluted by centrifuging at 10,000 for 1 minute. The eluate was collected and verified by agarose gel electrophoresis.

Site-Directed Mutagenesis

Mutations were introduced into the cDNA clone using site directed mutagenesis (SDM) using mutagenic primers. Reactions (20 µL) were set up in PCR tubes as follows: 4 µL of Phusion 5X buffer (New England Biolabs), 0.2 µL of Phusion Polymerase enzyme (New England Biolabs), 0.4 µL of dNTPs, 0.6 µL of DMSO, 11.8 µL of sterile water, 1.0 µL of template DNA, and 1.0 µL of each designed primer. The reaction used a 10 second denaturation

at 98 °C, 30 second annealing at 55 °C, 8-minute extension at 72 °C for 17 cycles. Reaction

products were verified via agarose gel electrophoresis. For samples that had amplification, 9 µL

were digested with 1.0 µL of DpnI (New England Biolabs). Digestion products were transformed

into MC1061 competent cells. Cultures were prepared of colonies from transformation plates and

miniprepped to yield target DNA.

Primers Used

The following primers were used to introduce mutations into MR766 via SDM as well as

perform qRT-PCR:

Primer	Length	Orientation	Purpose	Sequence
602	22	ZKV-MR766 (+)	change M422 to L in MR766 E	GCCAAGAGACTGGCAGTCCTGG
603	28	ZKV-MR766 (-)	Change M422 to L in MR766 E	AGGACTGCCAGTCTCTTGGCGCCTCTCA
604	28	ZKV-MR766 (+)	change T487 to M in MR766 E	TATCTCCCTCATGTGCTTGGCCCTGGCC
605	30	ZKV-MR766 (-)	change T487 to M in MR766 E	GCCAAGCACATGAGGGAGATAGATCCATTC
606	26	ZKV-MR766 (+)	change S139 to N in MR766 prM	CTTGGATAGGAACGATGCCGGGAAGG
607	25	ZKV-MR766 (-)	change S139 to N in MR766 prM	GGCATCGTTCCTATCCAAGTACATG
608	24	ZKV-MR766 (+)	change K143 to E in MR766 prM	AGCGATGCCGGAGAGGCCATTTCG
609	28	ZKV-MR766 (-)	change K143 to E in MR766 prM	AAATGGCCTCTCCGGCATCGCTCCTATC
610	27	ZKV-MR766 (+)	change A148 to P in MR766 prM	ATTTCGTTTCCTACCACATTGGGAGTG
611	26	ZKV-MR766 (-)	change A148 to P in MR766 prM	ATGTGGTAGGAAACGAAATGGCCTTC
612	35	ZKV-MR766 (+)	change V153 to M in MR766 prM	ACATTGGGAATGAACAAGTGCCACGTACAGATCAT
613	34	ZKV-MR766 (-)	change V153 to M in MR766 prM	ACTTGTTCATTCCCAATGTGGTAGCAAACGAAAT
756	23	ZIKV (+) MEG3	prM qRT-PCR	TTGGTCATGATACTGCTGATTGC
757	22	ZIKV (-) MR766	E qRT-PCR	CCCTCCACGAAGTCTCTATTGC

Table 1. List of Primers Used in This Project

Sequencing

Samples were sequenced (Sanger sequencing method) at the Genomic Core Facility at the Pennsylvania State University.

Plaque Assay

Plaque assays were done on monolaters of Vero cells. Ten-fold dilutions of P1 virus stocks were prepared with PBS, Ca^{2+}, Mg^{2+}, and 1% FBS. The media was removed from 6-well plates containing Vero cells. 250 µL of dilutions 10^{-3} through 10^{-8} were dispensed into individual wells. Plates were incubated on a rocker at room temperature for one hour. Wells were overlaid with 3 mL of 1X MEM, 5% FBS and 1% agarose. Plates were incubated at 35 °C for 48 hours. Plaques were counted and titers were calculated and recorded.

Transfection of Viral cDNA *in vitro*

Confluent HEK-293T cells were prepared in six-well plates. Five µL of lipofectamine (Fisher Scientific) was mixed with 45 µL of OptiMEM (Fisher Scientific) and allowed to incubate at room temperature for 5 minutes. Five µL of cDNA was mixed with 45 µL of OptiMEM. The two 50 µL mixtures were combined and incubated at room temperature for 20 minutes. One hundred µL of the solution was added to each well. Plates were incubated overnight at 37 °C. The media was removed the following day and the cells were washed with 2 mL of PBS. The PBS was removed and 2 mL of DMEM/10% FBS was added to each well. Plates were then returned to 37 °C and allowed to incubate overnight.

RNA Extraction

RNA was extracted from Huh7.5 cells using a RNAeasy Kit (Qiagen). Each centrifugation occurred at 10°C for 1 minute at 13,000g. Media was aspirated from each well and 250 µL of RLT buffer was added to each well and incubated at room temperature for 5 minutes on a rocker. Ethanol (350 µL, 70% v/v) was added to each well and the solution in each well was transferred to an individual spin column and centrifuged. The flow-through was discarded and 500 µL of RW1 buffer was added to the column and was centrifuged. The flow-through was discarded and 400 µL of RPE buffer was added to the columns and centrifuged. The flow-through was discarded and the RPE buffer treatment was repeated. The columns were transferred to a Qiagen 2 mL collection tube and centrifuged to dry the column. The columns were transferred to 1.5 mL RNase-free centrifuge tubes and 50 µL of RNase-free water was added to each column. Columns were centrifuged to elute the RNA, which was subsequently stored in the -80°C freezer.

cDNA Synthesis

Total RNA was extracted from virus-infected cells as described above and used as the template for cDNA synthesis. Reverse transcriptase PCR was used to synthesize cDNA of the Senegal NS5 region for the NS5 chimera virus. Reactions were set up using 25 µL of 2X one step reaction buffer (New England Biolabs), 1.5 µL of reverse transcriptase enzyme (New England Biolabs), 15 µL of water, 5 µL of RNA template and 2 µL of each primer. The reaction

was performed at 48 °C for 30 minutes (reverse transcription) and 40 cycles of 94 °C melting, 55 °C melting and 68 °C extension.

Quantitative RT-PCR

Real time quantitative reverse transcription PCR was conducted with a QuantStudio 3 thermocycler, using PowerSYBR Green RNA-CT 1-step kit (Applied Biosystems). A standard curve was constructed using pre-quantified purified ZIKV cDNA and primers 756 and 757 for absolute quantitation. qRT-PCR reactions were set up using 0.16 µL of reverse transcriptase enzyme, 10 µL of RTPCR mix, 0.5 µL of each forward and reverse primers, 4 µL of water and 5 µL of extracted RNA. Reactions were performed. using the following settings:

Stage	Step	Temperature (°C)	Duration (mm:ss)	Number of Cycles
Hold	1	40	30:00	
	2	95	10:00	
PCR	1	95	00:15	40
	2	60	01:00	
Melt Curve	1	95	00:15	
	2	60	01:00	
	3	95	00:15	

Table 2. qRT-PCR Thermocycler Settings

QuantStudio Design and Analysis software was used to analyze the results. Reactions were performed in triplicate.

Microscopy

Cells were visualized 4 days post transfection/infection using a Nikon A1R microscope with fluorescence settings at 10X and 40X magnification and 300ms exposure with a filter for green fluorescent protein.

RESULTS

Sequence Analysis Reveals Candidate Epidemic Strain Mutations

Comparison of amino acid sequence between Uganda (GenBank number KU963573), Senegal (GenBank number KX601166) and Honduras (GenBank number KX262887) strains yielded multiple amino acid differences per protein, most of which were not extreme—the amino acids differed, but their properties were similar so that the resulting proteins' structures and functions were not likely affected. Thus, analysis was focused on amino acid substitutions that were extreme and, therefore, the most likely to affect structure and/or function in the resulting proteins. Amino acid sequences of Uganda, Senegal and Honduras ZIKV strains were aligned to other amino acid sequences of viruses within the flavivirus genus—these viruses are known to cause relatively severe infections in humans (12)—to identify amino acid changes that had the potential to greatly affect structure and function. Six mutations in envelope (E) and pre-membrane (prM) proteins were selected for further characterization, as these amino acid differences were suspected to affect the structure and function of the resulting proteins the most (Table 3). Amino acid sequences of African lineage Uganda and Senegal strains did not differ among the selected mutations but differed from the Asian lineage Honduras epidemic strain. The amino acid properties between the African and Asian lineages differed greatly, generally changing size and/or hydrophobicity—these properties were also observed in various flavivirus strains. A notable exception was the lysine-to-glutamic acid change, in which positively charged lysine was substituted with a negatively charged glutamic acid. The serine-asparagine difference

in prM was of particular interest, as Yuan et al implicated this mutation in fetal microcephaly (13).

Protein	MR766 AA #	Zika Strain			Flavivirus Strain						
		Uganda	Senegal	Honduras	TBEV	YFV	DENV 1	SLEV	WNV	JEV	MVEV
E	422	M	M	L	L	L	M	L	L	L	L
	487	T	T	M	G	S	T	T	T	A	A
prM	139	S	S	N	K	E	Q	T	T	T	T
	143	K	K	E	------	K	K	S	D	D	D
	148	A	A	P	Q	-----	K	P	P	P	P

Table 3. Summary of Amino Acid Comparisons

Amino acid sequence alignments were constructed using ClustalOmega. These alignments highlight mutations of interest for this experiment, although there were many other mutations present.

Mutations in Envelope and Pre-membrane Proteins Affect Viral Replication

The six selected mutations were individually introduced into the MR766 cDNA clone via site-directed mutagenesis, and the resulting single-mutation plasmids were transfected into HEK-293T cells and the supernatant with the progeny virus was collected (P0) after four days of incubation. Wild type ZIKV MR766 plasmid was likewise transfected into HEK-293T cells and the collected P0 virus stock was used as a wild type control. The efficiency of these viruses to infect mosquito cells and human neuronal cells were tested by infecting C6/36 (*Aedes albopictus*) and differentiated SHSY5 cells respectively at a multiplicity of infection (MOI) of one. The cells were visualized four days post infection using fluorescent microscopy to see how these mutations affected virus production *in vitro*. Transfection of plasmid containing the

V153M mutation did not produce virus for any cell line, so this mutation was removed from further analysis.

MR766 wild type produced infectious virus (Figure 1A) consistent with previous images captured in the Jose lab (data not shown). Transfection of the M422L mutation in the envelope protein produced infectious virus but not to the same extent as the MR766 wild type. The M422L transfection yielded localized infection, only spreading to a few distinct clusters of cells (Figure 1B).
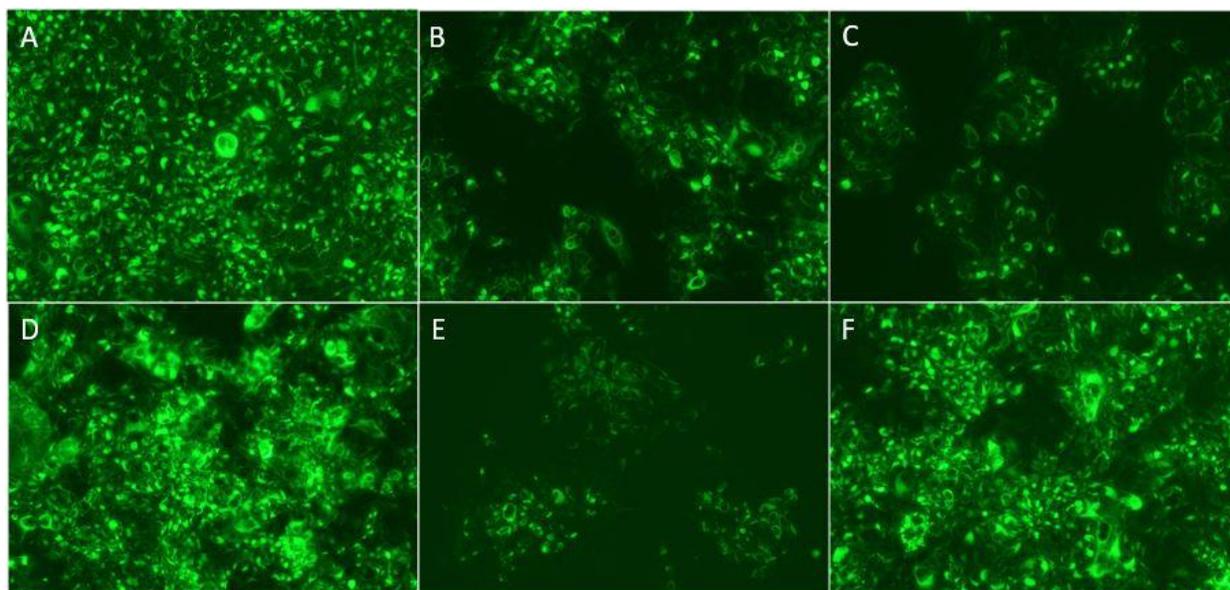

Figure 1. Transfection of HEK-293T Cells by ZIKV Mutants

Pre-epidemic MR766 served as the wild type control (A). M422L (B), T487M (C), S139N (D), K143E (E), and A148P (F) epidemic strain mutations were transfected for comparison. Cells were visualized under 10X magnification.

Transfection of the other mutation in the envelope protein, T487M, also produced infectious virus with similar characteristics as the M422L mutant (Figure 1C). Transfection of the S139N mutation of the pre-membrane protein produced infectious virus that was most similar to the wild type control (Figure 1D). Transfection of the K143E mutation of the pre-membrane protein

produced infectious virus to a lesser extent than did the wild type. Like the envelope protein

mutation transfections, the infection was localized and less wide-spread but showed the least

severe infection (Figure 1E). Transfection of the A148P mutation of the pre-membrane protein

produced infectious virus that resembled the wild type widespread infection but to a slightly

lesser extent (Figure 1F). We expected the mutants to yield infections to the same or higher

degree as the pre-epidemic wild type, as these mutations were characteristic of the epidemic

ZIKV strain; however, all epidemic strain mutants caused infections that were less widespread

than the pre-epidemic wild type virus and were thus less virulent. Other concurrent mutations in

epidemic strains and/or cellular factors must play a role in the infectivity of epidemic ZIKV in
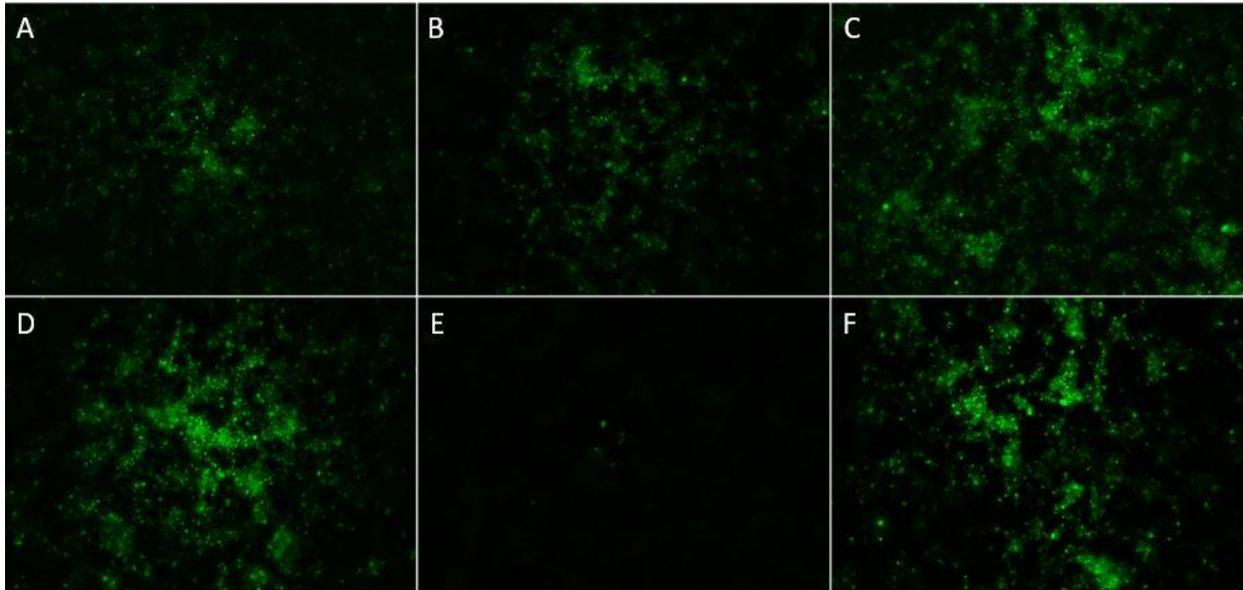
HEK-293T cells.

Figure 2. Infection of C6/36 Cells by ZIKV Mutants

Pre-epidemic MR766 served as the wild type control (A). M422L (B), T487M (C), S139N (D), K143E (E), and A148P (F) epidemic strain mutations were infected with P0 viruses for comparison. Cells were visualized under 10X magnification.

Pre-epidemic wild type and epidemic mutant P0 viruses were added to C6/36 cells to see how epidemic strain mutations affected virus infection in mosquitos, the vector necessary in the transmission of ZIKV to humans (the virus must be able to infect and replicate in the vector for successful transmission). Addition of MR766 pre-epidemic wild type virus yielded infectious virus (Figure 2A) consistent with previous transfections (data not shown). The M422L (envelope protein) mutant virus produced infectious virus and yielded a widespread infection similar to the wild type control (Figure 2B). The T487M (envelope protein) mutant virus also yielded infectious virus with an infection displaying slightly more severe than the wild type control and M422L mutant (Figure 2C). Addition of the S139N (pre-membrane protein) mutant yielded infectious virus and an infection more severe than the wild type control (Figure 2D). The K143E (pre-membrane) mutant yielded non-infectious virus, as virtually no cells showed signs of infection (Figure 2E). The A148P (pre-membrane protein) mutant yielded infectious virus that showed similar infection characteristics to the S139N mutant (Figure 2F).

This infection was also more severe than that of the wild type control. As with HEK-293T, we expected the viruses containing epidemic strain mutations to show similar or more severe infection characteristics than the pre-epidemic wild type. This was indeed the case for all but the K143E mutant, suggesting epidemic strains containing this mutation must contain compensatory mutations elsewhere in the genome in order to infect mosquito vectors.

The wild type pre-epidemic and mutant P0 viruses were added to neuronal cells to see how epidemic strain mutations affect the virus's ability to infect cells involved in the severe neurological conditions associated with ZIKV infection in humans. While severe infection *in vitro* may not necessarily establish a causative link between infection and neurological disease, it

is highly likely that the virus's ability to infect neuronal cells is related to the manifestation of

these severe conditions. The pre-epidemic MR766 strain is not associated with the neurological

diseases that are associated with epidemic strains, so we expected to observe poor infective

ability for this strain. This was the case, as the infection afflicted very few cells (Figure 3A). We

expected to observe more severe infections for epidemic strain mutants, as epidemic strains are

associated with neurological diseases. This was only the case for the A148P (prM) mutant. This

mutant yielded infection more severe than the non-neurotropic pre-epidemic wild type and other

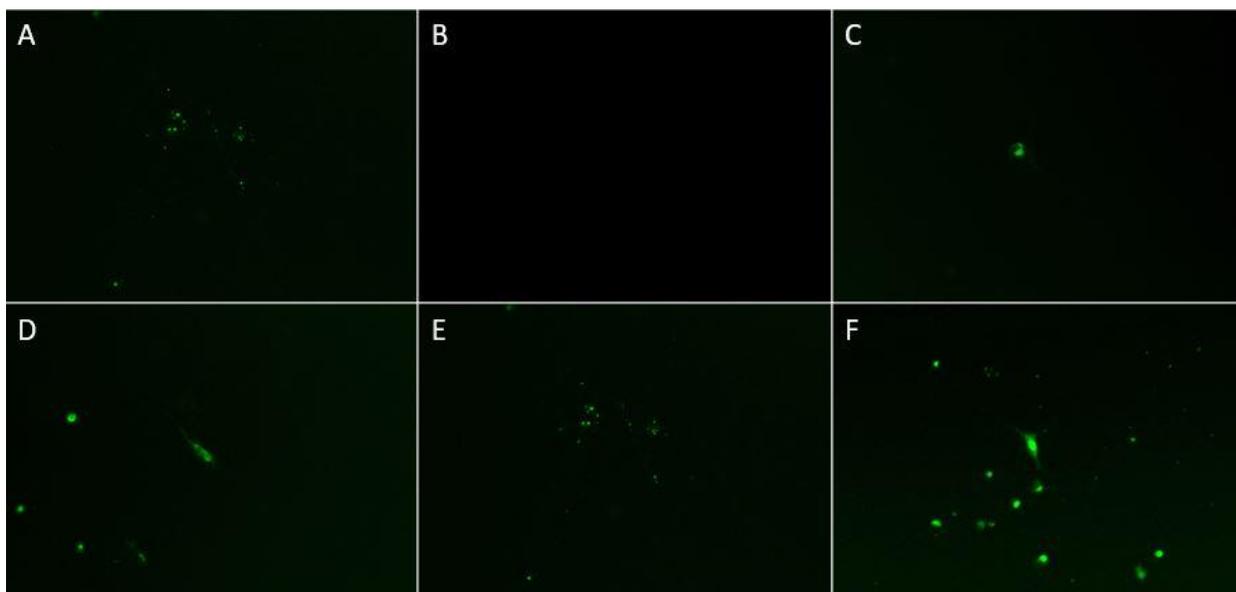mutant viruses, afflicting several neuronal cells (Figure 3F).

Figure 3. Infection of SHS5Y Neuronal Cells by Epidemic Mutants

Pre-epidemic MR766 served as the wild type control (A). M422L (B), T487M (C), S139N (D), K143E (E), and A148P (F) epidemic strain mutations were infected with P0 viruses for comparison. Cells were visualized under 40X magnification.

The other epidemic mutants—except for the M422L mutant—yielded infections similar to the

pre-epidemic wild type virus (Figure 3C, 3D and 3E), suggesting these mutations most likely do

not contribute to the neurotropism in epidemic strains and that other mutations and/or cellular

factors at play. The M422L (envelope protein) mutant produced no virus, suggesting residue at this position may play a role in neurotropism (Figure 3B).

Plaque assays were performed in Vero cells to quantify infections caused by the epidemic strain mutants. Epidemic strain mutant titers were expected to be higher than that of the pre-epidemic wild type control. This was not the case, as all titers were lower than the wild type control (Table 4). However, titers of the pre-epidemic wild type and epidemic mutant viruses were of the same magnitude except for the K143E mutant, which had a titer three orders of magnitude less. Additionally, the size of the plaques for each mutant were approximately the same size as the wild type except for the K143E mutant, which yielded very small plaques. The lower titer and smaller plaque size are consistent with observations of minimal infection in C6/36 cells as well as reduced infection in HEK-293T cells.

Mutation	Plaque Size (mm)	Titer (pfu)
Wild Type	2	8×10^5
M422L	2.5	1.2×10^5
T487M	1.5-2.0	2.4×10^5
S139N	3	1.6×10^5
K143E	<1	8×10^2
A148P	2	1.6×10^5

Table 4. Plaque Assays of Epidemic Mutants

Vero cells were infected with dilutions of MR766 pre-epidemic wild type and epidemic mutant viruses and overlaid with agarose to allow for plaque formation. Plaques were counted to determine the titer. Viral titer is represented as plaque forming units (pfu) per mL.

Most epidemic mutant strains did not match the microscopy results in HEK-293T or C6/36 cells, and none of the titers nor plaque sizes of the mutants corresponded to the observed infections of neuronal cells. This may be due to different growth characteristics of ZIKV in Vero cells.

Substitution of the NS5 Region Hinders Viral Replication

The entire NS5 region of the epidemic Senegal strain was substituted into the pre-epidemic MR766 strain to understand how the many accumulated mutations within the polymerase protein affected viral replication. HEK-293T cells were transfected with the chimeric virus, and the supernatant containing P0 progeny viruses was collected and used to infect Vero cells (MOI = 1) to characterize epidemic polymerase-mediated infection. We expected the epidemic strain NS5 would enhance replication, leading to a more severe infection *in vitro* compared to the pre-epidemic wild type control. Infection of the pre-epidemic wild type strain was consistent with observations made via microscopy during plaque assays (data not shown), with the observation of a widespread infection (Figure 4A). Both chimeras yielded localized infections significantly less severe than the wild type strain (Figure 4B, 4C). The disparity between infection of chimeric and wild type viruses suggest virulence mechanisms are more complex than hypothesized.

Mutations and NS5 Substitution Cause Reduction in Viral RNA Replication

To ensure the validity of previous results, viral RNA was extracted from separate transfections of wild type, mutant and chimeric viruses for real-time quantitative reverse transcription PCR—a reliable method to determine viral replication ability—performed in triplicate. We expected modified viruses to produce RNA in quantities equal to or higher than the pre-epidemic wild type. This was the case for all modified viruses, except for the K143E mutant and both chimeras, as they produced RNA quantities approximately equal to the pre-epidemic wild type. The K143E mutant and both NS5 chimeras produced significantly less RNA than the wild type and other mutants (Figure 4). In general, qRT-PCR data was consistent with plaque assay data and microscopy observations.
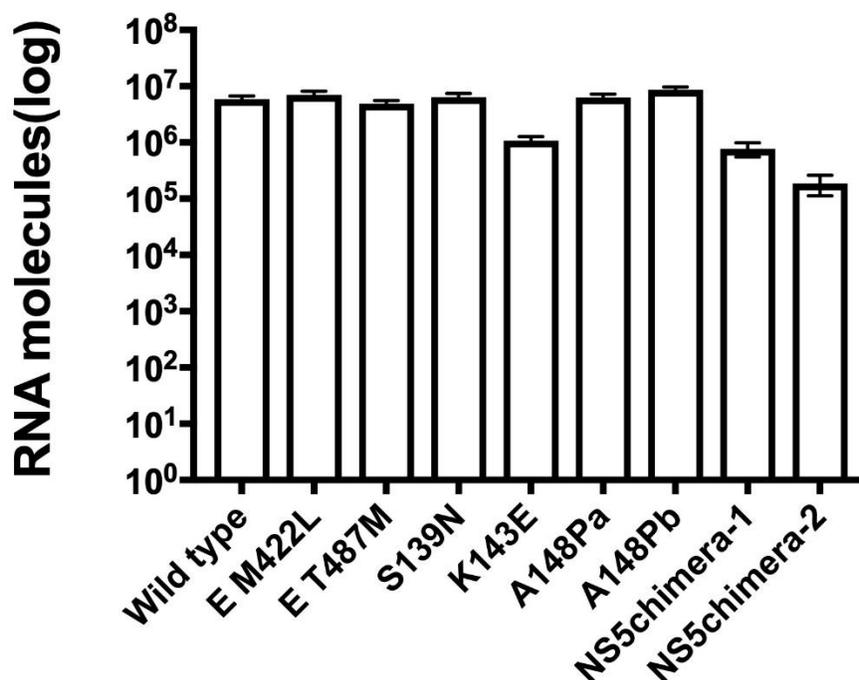

Figure 4. RNA Quantities Produced by Modified Viruses

Viral RNA was extracted from transfections of Huh7.5 cells and used as template in qRT-PCR. Reactions were performed in triplicate. Values are presented as mean ± standard deviation.

DISCUSSION

Sequence analysis performed via multiple sequence alignment of multiple ZIKV strains showed many mutations in epidemic strains compared to the pre-epidemic Uganda strain, most of which yielded significant amino acid substitutions. The number of significant amino acid substitutions reveals the rapid rate at which the virus has evolved in the past 72 years from a relatively benign to a dangerous, prompting the need for further research. ZIKV has recently joined its flavivirus counterparts as yet another mosquito-borne virus with severe public health implications.

In vitro growth characterization via fluorescence microscopy and plaque assay largely refuted our hypothesis, as most infections by the epidemic single mutants yielded infections significantly less severe than that of the pre-epidemic wild type. The only exceptions were the M422L E and S139N prM mutants in mosquito cells and the A148P prM mutant in neuronal cells. However, the more virulent characteristics observed in the mosquito cells for these mutants does not necessarily translate to severe infection in humans, as high virulence in mosquito vectors doesn't necessarily translate to high virulence in a human host. The epidemic mutants' growth *in vitro* suggested epidemic ZIKV virulence is not wholly ascribed to a single mutation. Rather, it is more likely that multiple mutations account for the observed virulence of epidemic strains. Further study is required to determine the mechanisms by which mutations affect virulence, specifically mutations that are related to infections associated with neurological disease. Although we did not study neurotropism exhaustively, this conclusion is in stark contrast to a study conducted by Yuan et al that suggested the S139N prM mutation is responsible for fetal microcephaly (13). Even though we found single mutations are not the cause

of increased virulence, the A148P prM mutation is of particular interest, as it yielded a more severe infection in neuronal cells than any other mutant or wild type. This mutation should be pursued in further studies to understand how it contributes to neurotropism, as neurotropism is an integral factor in the pathogenesis of ZIKV-associated neurological disease. Likewise, the K143E prM mutation should be further assessed to determine its role in viral replication, as this mutant yielded significantly reduced infection in all cell types compared to all other strains characterized in this study.

Viral RNA quantities determined via qRT-PCR corroborated observations made with fluorescent microscopy, showing neither epidemic strain mutants nor NS5 chimeras produced more RNA than the pre-epidemic wild type. The K143E prM mutant produced significantly fewer RNA molecules than the wild type control and other mutants, suggesting the residue at this position may play a role in genome replication via interaction with the polymerase. Furthermore, substitution of the entire NS5 polymerase region lead to significantly reduced RNA quantities, suggesting the epidemic polymerase does not favorably interact with pre-epidemic proteins—most likely the NS5 region—in viral replication. We propose convergent evolution of viral proteins has produced epidemic strains—each region of the polyprotein evolved simultaneously to produce strains whose proteins are able to favorably interact to cause high virulence. Thus, viral proteins from the epidemic strain presumably are not compatible with those of the pre-epidemic strain due to unfavorable protein-protein interactions via introduced mutations. Interactions of NS5 with other viral and host proteins can be determined in further studies based on results from this study.

Further study of ZIKV pathogenesis will lead to a better understanding of how its high virulence associated with epidemic strains causes severe neurological disorders and potential preventative measures against infection. Determining NS5 polymerase interactions with other viral proteins can potentially lead to antiviral target candidates to be used in development of antiviral therapeutics. Likewise, further characterization of S139N and A148P mutations (prM) is necessary in determining factors that lead to ZIKV neurotropism.

BIBLIOGRAPHY

1. Petersen, L. R., Jamieson, D. J., Powers, A. M., and Honein, M. A. (2016) Zika Virus. *The New England Journal of Medicine*. **374**, 1552–1560

2. Musso D, Gubler DJ (2016) Zika virus. *Clin Microbiol Rev*. **29**, 487–524

3. 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., and 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

4. Diallo, D., Sall, A. A., Diagne, C. T., Faye, O., Faye, O., Ba, Y., Hanley, K. A., Buenemann, M., Weaver, S. C., and Diallo, M. (2014) Zika Virus Emergence in Mosquitoes in Southeastern Senegal, 2011. *PLoS ONE*. 10.1371/journal.pone.0109442

5. Duggal, N. K., Ritter, J. M., Pestorius, S. E., Zaki, S. R., Davis, B. S., Chang, G.-J. J., Bowen, R. A., and Brault, A. C. (2017) Frequent Zika Virus Sexual Transmission and Prolonged Viral RNA Shedding in an Immunodeficient Mouse Model. *Cell Reports*. **18**, 1751–1760

6. Suely de Oliveira Melo, A., Santana Aguiar, R., Maria Ramos Amorim, M., and Arruda, M. B. (2016) Congenital Zika Virus Infection Beyond Neonatal Microcephaly. *JAMA Neurology*. **73**, 1407–1416

7. Broutet, N., Krauer, F., Riesen, M., Khalakdina, A., Almiron, M., Aldighieri, S., Espinal, M., Low, N., and Dye, C. (2016) Zika Virus as a Cause of Neurologic Disorders. *The New England Journal of Medicine*. **374**, 1506–1509

8. Wang, A., Thurmond, S., Islas, L., Hui, K., and Hai, R. (2017) Zika virus genome biology and molecular pathogenesis. *Emerging Microbes & Infections*. **6**, 1–6

9. Haddow, A. D., Schuh, A. J., Yasuda, C. Y., Kasper, M. R., Heang, V., Huy, R., Guzman, H., Tesh, R. B., and Weaver, S. C. (2012) Genetic Characterization of Zika Virus Strains: Geographic Expansion of the Asian Lineage. *PLoS Neglected Tropical Diseases*. **6**, 1–7

10. Smith, D. R., Kulesh, D. A., Bellanca, S. A., Haddow, A. D., Sprague, T. R., Buathong, R., Minogue, T. D., Miller, L. J., Alera, M. T. P., Valdez, S. M., Nasar, F., Coyne, S. R., Pitt, M. L., Jarman, R. G., Linthicum, K. J., Weaver, S. C., Golden, J. W., Koehler, J. W., Hollidge, B. S., Gromowski, G. D., Kane, C. D., Tesh, R. B., Lowen, R. G., Yoon, I.-K., Bavari, S., and Padilla, S. L. (2018) African and Asian Zika Virus Isolates Display Phenotypic Differences Both In Vitro and In Vivo. *The American Journal of Tropical Medicine and Hygiene*. **98**, 432–444

11. Zhu, Z., Chan, J. F.-W., Tee, K.-M., Choi, G. K.-Y., Lau, S. K.-P., Woo, P. C.-Y., Tse, H., and Yuen, K.-Y. (2016) Comparative genomic analysis of pre-epidemic and epidemic Zika virus strains for virological factors potentially associated with the rapidly expanding epidemic. *Emerging Microbes & Infections*. **5**, 1–12

12. Gould, E. A., and Solomon, T. (2008) Pathogenic flaviviruses. *Lancet*. **371**, 500–509

13. Yuan, L., Huang, X.-Y., Zhong-Yu Liu, Zhang, F., Zhu, X.-L., Yu, J.-Y., Ji, X., Xu, Y.-P., Li, G., Li, C., and Wang, H.-J. (2017) A single mutation in the prM protein of Zika virus contributes to fetal microcephaly. *Science*. **358**, 933–936

Education
Major: Biochemistry and Molecular Biology
Honors: Biochemistry and Molecular Biology

Thesis Title: Characterization of Genetic Mutations in Outbreak Strains of Zika Virus

Thesis Supervisor: Joyce Jose

Work Experience
Date: May 2013 – Present
Title: Emergency Medical Technician
Description: Responsible for emergency medical care of patients as part of advanced life support ambulance crew. This included mentoring volunteers to develop their patient care skills.
Institution/Company (including location): Warrington Community Ambulance Corps, Warrington, Pennsylvania.
Supervisor's Name: Ian Libertore

Community Service Involvement: Penn State Dance Marathon Captain—served as Stage Performance Coordinator and DJ Coordinator in THON 2018 and THON 2019 respectively. In these roles I was part of a group that provided music and performers during THON Weekend and at more than 50 events. These events helped raise over $10 million for Four Diamonds.