12-2021

Examination of Antiviral Resistance in Venezuelan Equine Encephalitis Virus

Jasper Lee
University of Tennessee Health Science Center

Examination of Antiviral Resistance in Venezuelan Equine Encephalitis Virus

Abstract

Venezuelan equine encephalitis virus (VEEV) is a New World Alphavirus that causes Venezuelan equine encephalitis (VEE), which is characterized by a febrile illness that can progress to neurological disease and death. While no major outbreaks of VEE have occurred since 1995, VEEV is a virus of concern as, in addition to its spread through mosquitos, it can be aerosolized and used as a bioweapon. Unfortunately, there are currently no FDA-approved vaccines or antivirals against VEEV. Efforts have been made to discover small molecules with an inhibitory effect on VEEV, but the potential for emergence of antiviral resistance to these compounds will remain a concern because VEEV is an RNA virus with a high mutation rate and grows to high titers. To examine the evolutionary trajectory of antiviral resistance in VEEV, we developed a next-generation sequencing pipeline to examine single-nucleotide polymorphisms that emerged after repeated passaging of the virus with increasing concentrations of antiviral compounds. In addition, we examined the effect of the microenvironment on the evolution of antiviral resistance, both in cell culture and mouse models. We found that VEEV evolves resistance to the compound ML336 and its derivatives through mutations in the nsP2 and nsP4 genes, but the number, timing of emergence, and the extent of penetrance of these SNPs depend on the compound. These mutations emerged more slowly when infecting an astrocyte cell line. We also found that neurons in the mouse brain did not impose a selective pressure on VEEV during an infection. These results demonstrate how the population dynamics of RNA viruses can be tracked over time and the extent to which they are affected by selective pressures, as well as opening questions about how viruses can mutate and adapt at the molecular level.

Document Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Program

Biomedical Sciences

Research Advisor

Colleen B. Jonsson, PhD

Keywords

Antiviral resistance, Venezuelan equine encephalitis virus, Virus evolution

Subject Categories

Analytical, Diagnostic and Therapeutic Techniques and Equipment | Diseases | Investigative Techniques | Medicine and Health Sciences | Virus Diseases

UNIVERSITY OF TENNESSEE HEALTH SCIENCE CENTER

DOCTORAL DISSERTATION

Examination of Antiviral Resistance in Venezuelan Equine Encephalitis Virus

Author:
Jasper Lee

Advisor:
Colleen B. Jonsson, PhD

A Dissertation Presented for The Graduate Studies Council of
The University of Tennessee Health Science Center
in Partial Fulfillment of the Requirements for the Doctor of Philosophy degree from
The University of Tennessee

in

Biomedical Sciences: Microbiology, Immunology, and Biochemistry
College of Graduate Health Sciences

December 2021

ACKNOWLEDGEMENTS

Firstly, I would like to thank Dr. Colleen Jonsson and everyone in the Jonsson lab for all of the help they have given me over the past six years. I would especially like to thank Dr. Jonsson for never giving upon me even as I struggled my way through graduate school, and for helping me grow as both a scientist and as a person.

I would also like to thank my graduate committee for their help and advice over the past four years. Drs. Michael Whitt, Lorraine Albritton, Kui Li, and Dave Rogers were all very happy to answer my questions and keep me thinking in the right directions. I would also like to thank Dr. Elizabeth Fitzpatrick, who while not on my committee was able to answer a lot of questions for me regarding the graduate program and managing the transfer from UTK.

I would also like to especially thank my fellow graduate students who have been with me since the days in Knoxville, Evan Williams and Mariah Taylor. You have become good friends of mine and have helped me so much, and I wish you well in your future careers. I would also like to thank the other graduate students in the Jonsson lab, Briana Spruill-Harrell, Walter Reichard, and Jacob Nichols, for their help and also being good friends in the lab.

Of course, a good portion of the studies that were done for this dissertation would not have been possible without many other people. I would like to thank Dr. Shruti Bansal, Jyothi Parvathareddy, and Dr. Yi Xue, all of whom made significant contributions and helped guide me through many of the procedures and data analysis. The RBL staff were also invaluable in their expertise in the animal studies and helping keep me on track with Biological Safety protocols; I would like to especially thank Dr. Jennifer Stabenow, Dr. Dong Yang, Haley Writt, and Dr. Ying Kong for their help in these regards. In my time at UTK, Gillian Eastwood, Leonardo Valdivieso-Torres, Ashley Yu, and Aubrey Sawyer for all their help and advice.

While my time at UTK ended on an unexpected note, I would like to thank Drs. Elizabeth Fozo, Brian O'Meara, Tim Sparer, and Chunlei Su for their help and advice as part of my initial graduate committee. I would also like to thank them and the rest of the faculty and graduate students at the Department of Microbiology for helping me get through a challenging first two years of graduate school.

Lastly, I would like to thank all of my family and friends who have helped keep me sane and focused over the past six years. I would like to especially thank my parents, who have shown me nothing but love, support, and encouragement throughout the entirety of my time in Tennessee. Thank you for everything.

The work presented here was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health with grant 5R01AI118814.

ABSTRACT

Venezuelan equine encephalitis virus (VEEV) is a New World Alphavirus that causes Venezuelan equine encephalitis (VEE), which is characterized by a febrile illness that can progress to neurological disease and death. While no major outbreaks of VEE have occurred since 1995, VEEV is a virus of concern as, in addition to its spread through mosquitos, it can be aerosolized and used as a bioweapon. Unfortunately, there are currently no FDA-approved vaccines or antivirals against VEEV. Efforts have been made to discover small molecules with an inhibitory effect on VEEV, but the potential for emergence of antiviral resistance to these compounds will remain a concern because VEEV is an RNA virus with a high mutation rate and grows to high titers. To examine the evolutionary trajectory of antiviral resistance in VEEV, we developed a next-generation sequencing pipeline to examine single-nucleotide polymorphisms that emerged after repeated passaging of the virus with increasing concentrations of antiviral compounds. In addition, we examined the effect of the microenvironment on the evolution of antiviral resistance, both in cell culture and mouse models. We found that VEEV evolves resistance to the compound ML336 and its derivatives through mutations in the nsP2 and nsP4 genes, but the number, timing of emergence, and the extent of penetrance of these SNPs depend on the compound. These mutations emerged more slowly when infecting an astrocyte cell line. We also found that neurons in the mouse brain did not impose a selective pressure on VEEV during an infection. These results demonstrate how the population dynamics of RNA viruses can be tracked over time and the extent to which they are affected by selective pressures, as well as opening questions about how viruses can mutate and adapt at the molecular level.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AA	amino acid
ABC	avidin-biotin complex
ADP	adenosine diphosphate
AUD	alphavirus unique domain
BBB	blood-brain barrier
BSA	bovine serum albumin
CHIKV	Chikungunya virus
CNS	central nervous system
Cp	capsid protein
C_T	threshold cycle
CPE	cytopathic effect
CPU	caudate putamen
DAB	3,3'-diaminobenzidine
DMEM	Dulbecco modified Eagle medium
DPBS	Dulbecco's phosphate buffered saline
dpi	day(s) post infection
EDTA	ethylenediarninetetraacetic acid
EEEV	Eastern equine encephalitis virus
EC_{50}	50% effective concentration
FBS	fetal bovine serum
GP	glycoprotein
GTase	guanylyltransferase
H&E	hematoxylin and eosin
Hpi	hours post infection
HTS	high-throughput screening
HVD	hypervariable domain
IFN	interferon
IHC	immunohistochemistry
i.n.	intranasal
i.p.	intraperitoneal
MOI	multiplicity of infection
MTase	methyltransferase
NHP	nonhuman primate
nsP	non-structural protein
nt	nucleotides
ORF	open reading frame
PBS	phosphate buffered saline
PFU	plaque forming unit
RdRp	RNA-dependent RNA polymerase
SINV	Sindbis virus
SLV	Semliki Forest virus
TATase	terminal adenylyltransferase
$TCID_{50}$	median tissue culture infectious dose

TF	transframe
TrD	Trinidad donkey
UTR	untranslated region
VEE	Venezuelan equine encephalitis
VEEV	Venezuelan equine encephalitis virus
WEEV	Western equine encephalitis virus
WT	wild type

CHAPTER 1. INTRODUCTION

VEEV Overview and Classification

The human and equine disease Venezuelan equine encephalitis (VEE) is caused by VEEV, a single-stranded, positive-sense RNA virus belonging to the genus *Alphavirus* and family *Togaviridae*. VEEV is often referred to as a New World Alphavirus alongside its close relatives Eastern equine encephalitis virus (EEEV) and Western equine encephalitis virus (WEEV). VEEV is a complex of strains comprised of multiple antigenic subtypes; of subtype I will be the major focus in this dissertation. The antigenic subtype I is further classified into subtypes AB, C, D, E, and F. Subtypes IAB and IC are the causative agent behind the major epizootic outbreaks and epidemics of VEE, but viruses of all subtypes have the potential to cause human disease [1]. A general idea of the relations between each subtype can be seen in **Figure 1-1**.

Although it is transmitted by mosquitoes in nature, VEEV is infectious through the aerosol route, and was developed as a biological weapon by both the US and former USSR during the Cold War due to this infection route combined with high titer and stability in the environment, leading to its classification as a Category B bioterrorism agent by the Centers for Disease Control and Prevention [2-4]. While not the focus of this review, VEEV outbreaks in equids have the potential to cause serious economic effects, underscoring the importance in understanding the virus and preventing future outbreaks.

VEEV Major Outbreaks and Ecology

VEEV is endemic to South and Central America, where it been responsible for several major outbreaks in the 20th century ranging from Colombia and Venezuela, down as far south as Peru and as far north as Mexico and Texas. The wide majority of outbreaks have been caused by viruses of the IAB and IC subtypes. VEE was first observed in equids in the 1930's, and the causative agent was first isolated from horse brains in 1938 [5]. However, VEEV's ability to cause disease in humans was not confirmed until its isolation from human cases in Colombia and Venezuela in the 1950's and 1960's [6]. One of the largest outbreaks began in 1969 in Central America and reached as far north as Texas by 1971 [7]. The most recent major outbreak occurred in 1995, when over 100,000 human cases were diagnosed in Venezuela and Colombia [5, 8, 9].

In nature, VEEV is classified as epizootic or enzootic depending on the hosts where the virus is circulating, with epizootic strains found in equids and humans while enzootic strains are found in small rodent reservoir hosts. In all cases this intermediate host is an equid, usually a horse or donkey. IAB and IC subtypes are the only epizootic viruses, as ID and IE subtypes do not amplify and disease in equids, instead spreading directly from their rodent reservoir hosts to humans through a mosquito carrier [10].

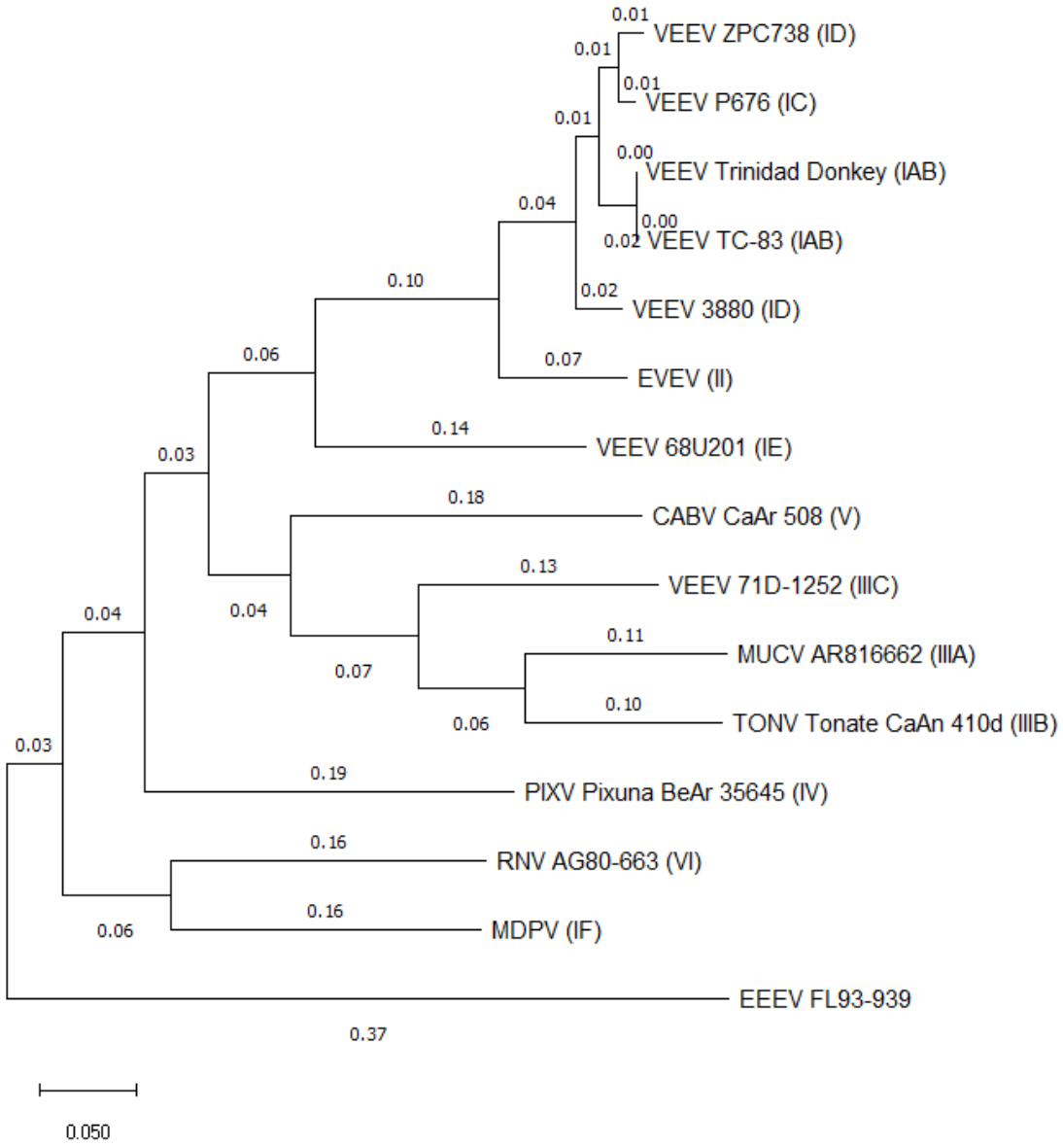


Figure 1-1. Phylogenetic tree of representative strains of VEEV.

Phylogenetic tree showing the relationships between different viruses in the VEEV complex. For the most part, sequences chosen were listed as "exemplar isolate of the species" by the International Committee on Taxonomy of Viruses, with a few additional selections of commonly-used laboratory strains [11-15]. Tree was generated using the maximum likelihood method and Tamura-Nei model with 1000 bootstraps. The numbers shown are branch lengths, measured by number of substitutions per site. EEEV was chosen as an outgroup as it is the closest known relative of VEEV [16]. EVEV, Everglades virus; CABV, Cabassou virus; MUCV, Mucambo virus; TONV, Tonate virus; PIXV, Pixuna virus; RNV, Rio Negro virus; MDPV, Mosso das Pedras virus.

Various mosquito species are involved in both epizootic and enzootic transmission of VEEV, depending on the subtype and reservoir host. Mosquitos pick up VEEV through the viremia of the mammalian host they are blood feeding upon, after which the virus likely enters the hemocoel through midgut epithelium and spreads to the rest of the body. Once the salivary glands are reached, VEEV replicates there and is transferred into a mammalian host during a blood meal. In outbreaks of epizootic VEEV strains, *Aedes taenniorhynchus* appears to be the primary mosquito species for transmission, although others such as *Aedes sollicitans* and *Psorophora* species can also be involved [17, 18].

In sylvatic cycles, mosquito species of the subgenus *Culex (Melanconion)* are the main carriers of enzootic VEEV strains [19]. A number of small rodent genera have been cited as likely reservoirs for endemic VEEV strains, including *Sigmodon* and *Proechimys*, as they are found in areas where outbreaks have occurred and can carry a viremia of high enough titer for transmission to mosquitoes [20, 21]. There is evidence of bats being able to harbor VEEV, although their place in the sylvatic cycle is unclear [22].

VEEV Disease and Pathogenesis

In humans, VEE is characterized by fever and malaise that can progress to neurological disease, causing symptoms including convulsions, stupor, seizures, and comas, which, in rare cases, can eventually lead to fatalities [2, 9, 23]. Symptoms of fever and malaise start to appear 2-5 days after infection and subside after an additional 4-6 days but they can recur. Encephalitis occurs if the virus penetrates into the central nervous system (CNS). VEE can occur in all age groups without sex bias, although children are more likely to develop encephalitis and succumb to disease than adults [5].

In fatal human cases of VEE, the histopathological evidence of infection has been found in the brain, thymus, lymph nodes, lungs, spleen, gastrointestinal tract, kidneys, and liver [23, 24]. Follicle necrosis, congestion and edema, and infiltration of lymphocytes, mononuclear cells, and neutrophils were seen in many of these tissues [24]. However, little to no information is known about the histology of the brain or the mechanisms of neurodegeneration in a fatal human infection [25].

Much of the pathogenesis of VEEV comes from its study in rodent models, particularly the mouse [25]. In a subcutaneous (s.c.) mouse infection, VEEV first establishes infection in dermal dendritic cells, which move the virus to lymph nodes, where VEEV replicates before causing viremia [26, 27]. To enter the CNS, VEEV accesses the olfactory neurons through the epithelial layer in the nasal cavity, which provides direct access for the virus to cross from the bloodstream to peripheral neurons. The virus then travels through these neurons to the brain, entering in the olfactory bulb [27, 28]. From there, VEEV spreads to the rest of the brain and CNS; this process can occur as quickly as 36-48 hours post infection (hpi) [27]. In aerosol and intranasal (i.n.) infections, VEEV establishes an infection at the olfactory epithelium, reaches the brain in as little as 16 hpi, causing symptoms to appear more quickly than through other routes of

infection [29]. Once the CNS is penetrated, VEEV spreads through the rest of the brain primarily through neurons, although astrocytes can be infected as well. Areas of the brain that are particularly affected include the piriform cortex, one of the first areas of the brain to be infected, as well as neocortex, thalamus, midbrain, cerebellum, and brainstem [25]. In aerosol and i.n. infections, viremia is not necessary to cause neuroinvasion. Neuronal cell death occurs primarily through apoptosis, although cell death does not appear to occur in the cerebellum [30]. Overexpression of inflammatory cytokines by the host or possibly induced by virus may also contribute to neuronal cell death [31, 32].

VEEV may enter the CNS through other means. One of the more studied locations for possible CNS entry is the blood-brain barrier (BBB). It has been suggested that once viral replication is established in the brain, opening of the BBB can be induced, further allowing entry of virus into the CNS, although whether this is due to the virus directly inducing it or simply because white blood cells such as monocytes are crossing the BBB is still unclear [33, 34]. Interestingly, a study using a recombinant VEEV.3908 (subtype IC) strain with luciferase could penetrate the CNS by entering through circumventricular organs, areas at which the BBB is not present [35]. Markers of viral infection of the CNS can be detected at least three days before symptoms of neurological disease appear [36]. Trigeminal nerves also serve as an alternate pathway to enter the CNS from the bloodstream if the olfactory nerves are not available, with VEEV crossing the dental pulp [27, 28].

VEEV Genome and Protein Functions

Like other alphaviruses, the genome of VEEV is comprised of single-stranded RNA, which is packaged in virions in a manner that allows it to be directly translated into protein using host ribosomes when inside a host cell. The genome is comprised of two open reading frames (ORFs), each of which contain several genes (**Figure 1-2**). The open reading frame closer to the N-terminus contains the genes nsP1, nsP2, nsP3, and nsP4, all of which encode proteins of the same name. Further downstream in the second reading frame are the C, E3, E2, 6K, and E1 genes, which also encode their namesake proteins, except in the case of 6K where a different protein, transframe (TF), can be synthesized instead. Lastly, there are three untranslated regions (UTRs) in the genome, ones at the 5' and 3' ends respectively, and a junction region between the two ORFs. It is worth noting, however, that most of the studies to determine alphaviral genome and protein characteristics and functions have been done in Old World Alphaviruses, particularly Sindbis virus (SINV) and Semliki Forest virus (SFV), which means some information described here may not necessarily be exactly true for VEEV but is presumed as such [37].

The nsP1 gene encodes a methyltransferase and a guanylyltransferase that add a 5' cap on genomic and subgenomic RNA generated during viral replication. 5' capping occurs after the substrate RNA is prepared by the nsP2 [38]. In addition, the nsP1 gene contains a packaging signal at nucleotides 856-1150 that is recognized by the capsid

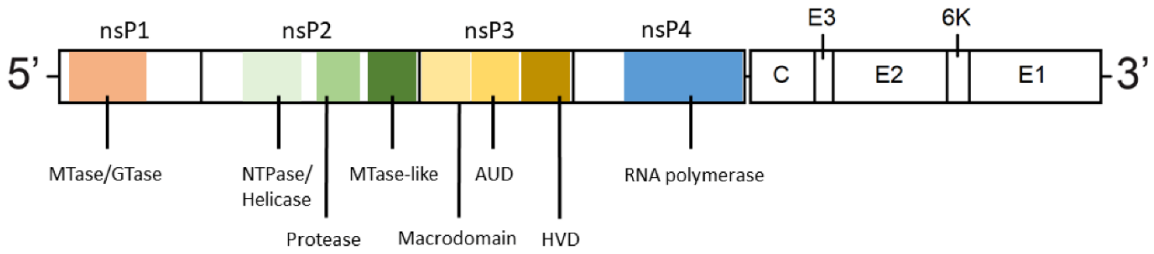


Figure 1-2. Genome structure of VEEV.

All genes and positions in the genome are listed, as well as known domains in the non-structural genes. MTase, methyltransferase; GTase, guanylyltransferase; NTPase, nucleoside triphosphatase; AUD, alphavirus unique domain; HVD, hypervariable domain.

during viral assembly [39]. nsP1 protein may also have a role in minus-strand synthesis as mutations in it rescue viruses with a G83 mutation in nsP4 [40]. Finally, nsP1 also has the ability to associate with host membranes, although this may not be essential depending on the alphavirus [41].

nsP2 mainly serves as a nucleoside triphosphatase (NTPase), helicase, and protease. The first clearly defined domain contains both the NTPase and helicase activities, both of which are essential to viral replication [42]. This domain also appears to contain an RNA 5'-triphosphatase activity that removes a phosphate from the 5' end of newly-generated RNA, allowing for the nsP1 capping reaction to happen [38]. The second domain contains a cysteine protease that cleaves the P1234 and P123 polyproteins at various points in the replication process [43]. At the C-terminal end of nsP2 lies an MTase domain, which does not seem to exhibit MTase activity due to the lack of an active site but may assist in substrate binding to the protease [43, 44]. The N-terminal portion of nsP2 contains two putative domains, the first of which may be a cofactor for the protease [45]. In addition to these functions, nsP2 may also have a role in the encapsidation of RNA into new virions [46]. In SINV and SLV, and likely other Old World Alphaviruses, nsP2 carries out transcriptional shutoff of host protein activity [47].

nsP3 contains a macrodomain, an alphavirus unique domain (AUD), and a hypervariable domain (HVD), all of which are involved in RNA synthesis but whose exact functions in replication have not been well-elucidated. The macrodomain binds to DNA, RNA, and poly(adenosine diphosphate (ADP)-ribose) and contains ADP-phosphatase activity [48]. The AUD, well-conserved among all alphaviruses, is also named the zinc-binding domain (ZBD), as an essential zinc-binding motif is located within it [49]. Despite the lack of conservation between different species, the HVD may be essential for alphavirus replication as deleting it causes virus attenuation, at least in vivo [50]. The HVD is phosphorylated during replication, but this phosphorylation may not be essential, as with SFV the lack of these phosphorylation sites had little effect *in vitro* but attenuated the virus in mice [51]. The VEEV HVD has been found to be nonessential in the baby hamster kidney (BHK)-21 cell line, while its phosphorylation is not required in vertebrate cells but essential in mosquito cells, showing differential virus-host interactions are likely at play with this domain [52]. However, regardless of host, it is still unclear as to what kinases carry out that activity [51, 52].

Once the core polymerase subunit was discovered in nsP4, studies quickly confirmed its presence and role in replication complexes, showing that it indeed carried out RNA-dependent RNA polymerase (RdRp) activity [53]. It was also predicted and confirmed that the alphavirus RdRp contains a palm with GDD motif, a thumb domain, and finger domains [54-56]. The SINV nsP4 also contains terminal adenylyltransferase (TATase) activity that is independent of all of the other nonstructural proteins, and which may maintain the poly(A) tail on synthesized RNA [55, 56]. Through a recombinant bacterial system, SINV nsP4 was shown to be able to carry out TATase activity, but required the other non-structural proteins to carry out de novo RNA synthesis, which suggest optimal replication requires all four proteins [56]. What is less understood is the function of the N-terminal disordered region. This region is essential, as when deleted or

mutated at certain amino acids such as G83, it compromises the ability of the virus to carry out minus-strand and plus-strand synthesis [40, 55]. Unfortunately, a lack of structural information regarding nsP4 leaves much uncertainty over its arrangement of the complexes it is a part of.

The C, E2, and E1 genes encode for the main structural components of alphaviruses, with E3 sometimes also present. The C gene encodes the viral capsid protein (Cp), which assembles into an icosahedral shape and associates with the viral genome to form the nucleocapsid. In New World Alphaviruses, Cp has also been implicated in shutting down host translation through binding to factors such as importin α/β and nucleoporins [57, 58]. The E1 and E2 proteins assemble to form the spike on the envelope of virions, arranged in a trimer of the combined E1 and E2 [59]. E2 is used for host receptor binding, while the primary role of E1 is in membrane fusion [60]. E3 is bound to E2 in the precursor protein pE2 before it is cleaved by furin-like proteases in the Golgi apparatus, and assists in keeping E1 and E2 in the heterodimer to prevent early fusion at low pH within vesicles [61, 62]. In VEEV, E3 appears to still associate with the heterodimer in mature virions [63].

The 6K protein interacts with E2 to help form mature glycoproteins and may also assist in budding [64]. The TF protein is generated from a frameshift of -1 as the RdRp stalls at a UUUUUUA motif in the 6K gene during synthesis of the subgenomic RNA, causing TF to be formed in place of 6K after post-translation modification [65, 66]. TF has fewer amino acids than 6K and has only one transmembrane domain compared to the two that 6K possesses. Removing the ability of alphaviruses to create TF protein impairs replication, and can also decrease the pathogenicity of the virus, although its exact functions are currently not well-understood [67]. E3, 6K, and TF may be incorporated into the envelopes of mature virions, although this is not essential for the virus [68].

Noncoding regions of the genome also serve important functions for VEEV. The 5’ UTR contains an important stem-loop that is necessary for the initiation of RNA synthesis [69]. The 3’ UTR is of varying length depending on the alphavirus; in VEEV, it is roughly 120 nucleotides in length and consists of a conserved RNA sequence element and poly(A) tail [70]. This conserved element is the location of initiation of minus-strand synthesis [71]. The junction region between nsP4 and C is a third UTR that, along with the C-terminal end of nsP4, forms the promoter for the synthesis of subgenomic RNA, so that this promoter sequence is only active on the minus strand [72].

VEEV Replication

VEEV replicates through a series of steps that are summarized in **Figure 1-3**. As an enveloped virus, the life cycle of VEEV begins with the attachment and binding of envelope glycoproteins to host cell receptors, triggering clathrin-mediated. Various host factors have been studied as possible receptors for attachment and binding to VEEV envelope glycoproteins. Following entry, the virion stays within the endosome until its fusion with a lysosome to form the endolysosome. Once the endolysosome is formed, the

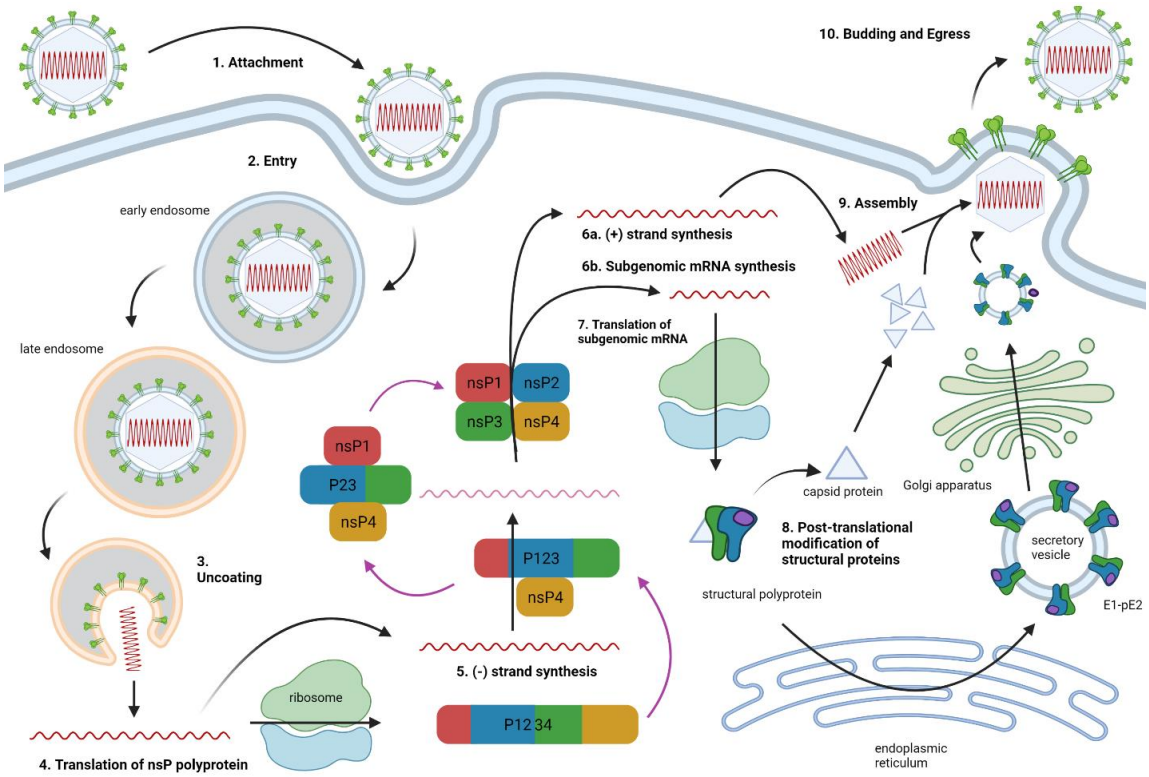


Figure 1-3. Alphavirus replication in a mammalian host cell.

VEEV enters the cell through receptor mediated endocytosis, and the genome is released into the cytoplasm in the late endosome following a pH change. The nonstructural polyprotein is translated from the viral genome, which is used by the partially cleaved polyprotein as the template for minus-strand synthesis. Following full cleavage of the polyprotein into the four individual nonstructural proteins, the replicase complex synthesizes the full-length genome and the subgenomic mRNA, the latter of which is translated into the structural polyprotein. The structural polyprotein is cleaved and modified to form the capsid protein and envelope proteins. During assembly, the nucleocapsid links with the envelope proteins to drive budding and egress from the host cell, producing a mature virion.

ensuing pH drop triggers the fusion of the viral envelope to the endolysosome membrane, allowing the viral genome to enter the host cytoplasm [73]. Once in the cytoplasm, the nsPs are produced directly from the genome through its translation by a host ribosome, and the replication complex is formed. Replication begins with minus-strand synthesis, followed by synthesis of the plus-strand and subgenomic RNA. Subgenomic RNA is translated into a polyprotein that is cleaved during post-translational modification into mature capsid and envelope proteins. These proteins assemble near the host plasma membrane, where mature virions exit the host cell.

Amino acid residues on E2 are responsible for binding to receptors on the host cells and facilitating attachment. A number of host receptors and attachment factors have been proposed for VEEV and other alphaviruses. For VEEV this list includes heparan sulfate, human lamin-binding protein, and low-density lipoprotein receptor class A domain-containing 3 [74-76] although the identity may depend on the type of cell and the host species. For example, while heparan sulfate has been found to be an attachment factor for multiple alphaviruses, it may be an adaptation from repeated passaging in cell culture models [76, 77]. Once attachment occurs, clathrin-coated pits are formed on the host plasma membrane, leading to clathrin-mediated endocytosis endocytosis [78]. VEEV enters the cytoplasm during the late endosome stage, after fusion of the endosome with a lysosome. The resulting acidification is what allows the viral envelope to fuse with the endolysosome membrane [73].

Once in the cytoplasm, the viral genome is directly translated into P123 and P1234 polyproteins, with both products translated due to a leaky opal stop codon (UGA) following the nsP3 gene [79]. These polyproteins are cleaved by nsP2 into the four constituent proteins over the course of replication. Firstly, nsP4 is cleaved from P1234 by the nsP2 protease [80]. The resulting complex of P123 and nsP4 is used to generate a minus-strand RNA template [81]. To mark the switch from minus-strand synthesis to plus-strand synthesis, P123 of the P123/nsP4 complex is cleaved into nsP1 and P23. P23 is then quickly further cleaved into nsP2 and nsP3, after which all four nonstructural proteins form the final replication complex [49]. This replication complex has been shown, at least in Old World alphaviruses, to form on membranes either on the host plasma membrane or on membranes within the cytoplasm, creating invaginations called spherules within which the replication occurs [82, 83]. The RdRp activity of nsP4 as well as the helicase and NTPase activities of nsP2 are used in plus-strand synthesis, generating both genomic and subgenomic RNA. After the nsP2 NTPase prepares the plus-strand RNA, nsP1 adds a 5’ cap to the genomic RNA using its MTase and GTase activities [38]. Unfortunately, it is still unclear as to how each non-structural protein is arranged and organized in the replication complexes, as well as what causes spherule formation [37].

Subgenomic RNA is translated into a polyprotein that is then separated into Cp, pE2, 6K or TF, and E1 proteins during following translation. Cp self-cleaves from the rest of the polyprotein while it is still in the cytoplasm, and packages with the genomic RNA to form the nucleocapsid [84]. However, it is unclear as to how the nucleocapsid is trafficked to the plasma membrane. The remaining polyprotein undergoes post-

translational modification in the endoplasmic reticulum. pE2, 6K/TF, and E1 are separated through signal peptidase cleavage, then travel to the plasma membrane using the secretory pathway from the ER through the Golgi apparatus. In the ER, pE2 and E1 form a heterodimer and cellular host-mediated furin cleavage separates pE2 into E2 and E3 proteins before they reach the plasma membrane [61, 62]. Correct E1/E2 heterodimer formation is required for assembly and binding. The heterodimers themselves also trimerize to form an icosahedral lattice before assembly [85]. E3 does not dissociate from the heterodimer until the proteins leave the low-pH environment of the vesicles to prevent triggering fusion [62]. Once the nucleocapsid and envelope proteins arrive at the plasma membrane, assembly and budding into mature virions occurs, with an important interaction between a hydrophobic cleft in the Cp protease and the endodomain of E2 [86, 87]. This Cp-E2 interaction is incorporated into the virion after budding. Budding appears to take place at specific locations on the plasma membrane and optimally occurs at a neutral pH [88, 89]. The driving force behind budding is still somewhat unclear; Cp-E2 binding may help drive it, while the 6K and TF proteins may also be involved [64, 90, 91]. In addition to the production of cell-free virus, alphaviruses are also capable of direct cell-to-cell transmission through the formation of long intercellular extensions [92].

RNA Virus Population Dynamics

RNA viruses have a high mutation rate due to its RNA-dependent RNA polymerase lacking proofreading ability, with an estimated mutation rate between 10^{-4} to 10^{-6} errors per round of replication [93, 94]. VEEV is no exception, with an estimated 10-4 mutation rate [95]. In addition, its rapid life cycle allows it to replicate in cell culture in as little as 8 hours, leading to high titers of virus, that can exceed values of 1 x 10^{8} pfu/ml [96]. This means that, in a given population, there will be individual viral genomes that contain 0, 1, or several to many SNPs compared to the consensus sequence of that population within a sequence space containing all possible mutations. Because of the constant generation of individual mutants, the composition of an RNA virus population is always changing and allows for flexibility to evolve and adapt to selective pressures in its environment. Molecular clones will also eventually diversify into such a population after replicating [97].

The possibilities for a virus population as it replicates and generates mutations can be represented by a fitness landscape, on which these possibilities take up a certain "sequence space" that peaks where fitness is highest for that population [97]. Selective pressures from the environment cause these sequence spaces to shift as the virus population adapts to those pressures. Many of the non-consensus viruses may have a lower fitness relative to the rest of the population, but these less-fit genomes are kept without much detriment to the overall population in what is termed mutational robustness [98]. Robustness is what allows less-fit genomes to reside within the population while expanding the population's sequence space and giving it a better chance to adapt to selective pressures.

Bottlenecks are also a significant factor in the population dynamics of RNA viruses, particularly for those that infect multiple hosts. Mosquito and vertebrate hosts present a variety of different environments for viruses to replicate in, and this causes bottlenecks when viruses move from, for example, one organ within a mosquito, mammal, or bird to another. One location where this bottleneck occurs is at the mosquito midgut, where it has been shown through haplotype tracking that the virus population undergoes a decrease in population and a reduction in genetic diversity after crossing the barrier [99, 100].

RNA viruses have adapted to the strategy of generating a diverse intra-population well, to the point that altering the fidelity of the RdRp in these viruses, whether to a higher or lower mutational frequency, causes the wide majority to be attenuated, particularly *in vivo* [101]. In VEEV, this tendency holds true as RdRp mutants of VEEV generally showed attenuation and inability to overcome bottlenecks, although there was still the chance of reversion to WT levels of virulence [102, 103].

The ability of RNA viruses to rapidly mutate has major implications for antiviral development, as drugs can be made ineffective if the target virus rapidly evolves resistance to those drugs, and drug resistance can easily emerge and persist even for FDA-approved antivirals [104, 105]. Therefore, understanding the dynamics of the evolution of resistance can aid in developing strategies to counter this property of RNA viruses.

Discovery and Development of Antivirals to Treat VEE

Antiviral development against alphaviruses is of high importance, as currently there are no FDA-approved antivirals or vaccines against any alphavirus available to the public. In addition, due to the unpredictable nature of epidemics and the underdiagnosis of cases in high-risk locations, a broad-ranging vaccine program would not be currently feasible, necessitating the use of antivirals to halt outbreaks [1]. In the past and present, studies have been undertaken to find compounds that cause a strong inhibitory effect on VEEV, a summary of which is given here and in **Table 1-1**.

One method of searching for these compounds is high-throughput screening of a library of antivirals, measuring the reduction in cytopathic effect (CPE) when each compound is added in cell culture. Each hit compound discovered is then assayed in a dose-response experiment to determine their effectiveness while assayed for cytotoxicity, determining either the IC_{50} or EC_{50} as well as their CC_{50}. Lead compounds are then selected using this data alongside other factors such as feasibility of synthesis and compound stability, after which the compound is characterized in its mechanism of action and effectiveness *in vivo*. This process was used in the discovery of a quinazolinone compound designated CID15997213, which had antiviral activity against all strains of VEEV tested [106]. Using this quinazolinone as the basis for a chemical scaffold, synthesis of derivatives led to the discovery of amidines that also showed anti-VEEV activity, including ML336 [107].

Table 1-1. List of antivirals tested for treatment of VEE.

Compound Name	Year Published	Cell Systems Used for Assays	Mode of Action	EC50	IC50	CC50	Source
Ribavirin	2000	Vero	Nucleoside analogue	N/A	> 500 µM	446 µM	[108]
VX-497 (merimepodib)	2000	Vero	IMP dehydrogenase inhibitor	N/A	12.4 µM	> 31 µM	[108]
(-)-carbodine	2008	Vero	Inhibition of cytidine triphosphate synthetase	0.2-0.3 µg/mL	> 100 µg/mL	N/A	[109]
BIOder	2012	Vero	Glycogen synthase kinase-3β inhibitor	N/A	< 0.1 µM	> 500 µM	[110]
acriflavine	2014	U87MG	Ago2 inhibitor	N/A	0.20 µM	7.22 µM	[111]
CID15997213	2014	Vero 76	Replicase complex inhibitor	N/A	1.3 µM	> 25 µM	[106]
favipiravir	2014	Vero-A	Nucleoside analogue	11 µM	N/A	N/A	[112]
β-D-N^4-hydroxycytidine	2018	Vero	Nucleoside analogue	0.426 µM	N/A	> 200 µM	[113]
sorafenib	2018	Vero	Possibly phosphorylation of eIF4E	3.7 µM	N/A	> 80 µM	[114]
RU486 (mifepristone)	2019	Vero	Possibly inhibition of nuclear import of Cp	19.9 µM	N/A	165 µM	[115, 116]

Table lists a representative collection of compounds that have been tested on VEEV in vitro. Compounds are ordered by year of first publication.

Broad spectrum antiviral compounds have also been tested on VEEV with varied success. A number of these compounds interfere with viral replication, often acting as nucleoside analogues. Along with other alphaviruses, VEEV was inhibited by the nucleoside analog favipiravir and its defluorinated analogue T-1105, with the EC_{50} being 11±4.3 and 13±2.9 μM respectively [112]. A different nucleoside analog, β-D-N^4-hydroxycytidine, was found to both decrease viral titer and increase the number of mutations in the viral genomes [113]. In contrast, ribavirin, another commonly-used nucleoside analog, does not appear to inhibit the replication of VEEV [108].

Another category of compounds act on host factors that are used by viruses during their life cycle. IMP dehydrogenase, an enzyme found in a guanine nucleotide synthesis pathway, is inhibited by compounds such as merimepodib (VX-497) to limit the amount of nucleotide substrate available within the host cell for synthesis of viral RNA or DNA. Merimepodib was shown to significantly inhibit VEEV replication *in vitro* on its own and when combined with IFNα treatment [108]. RU486, also known as mifepristone, was originally developed as an anti-HIV drug that prevented HIV integrase from binding to host importin α/β1, but was found to also alter the cellular distribution and decrease nuclear import of VEEV Cp [115]. Synthesis and investigation of mifepristone derivatives on inhibition of VEEV is ongoing [116].

CHAPTER 2. EMERGENCE AND MAGNITUDE OF ML336 RESISTANCE IN VENEZUELAN EQUINE ENCEPHALITIS VIRUS DEPEND ON THE MICROENVIRONMENT*

Introduction

Venezuelan equine encephalitis virus (VEEV) is an arbovirus belonging to the genus Alphavirus that causes a febrile illness in humans that may progress to life-threatening encephalitis but is often accompanied with symptoms including stupor, convulsions, and comas [2, 9]. The largest and most recent major VEEV outbreak occurred in 1995 in northern Colombia and Venezuela and resulted in 85% of infected children showing neurological symptoms, as compared to 15% of infected adults. In addition, the disease had a 5% case fatality rate [9]. VEEV remains infectious when aerosolized and was developed as a biological weapon by the USA and USSR during the Cold War. Despite the ratifications of the Biological Weapons Convention in 1972 that bans such activities, VEEV remains a potential biological threat with high priority [2, 4]. Despite its potential threat to public health, there are currently no FDA- approved antivirals or vaccines for VEEV. Antivirals are of high importance as epidemics of VEE are unpredictable and broad-ranging vaccine programs may not be feasible in all situations [1].

In recent years, we have worked on the discovery and subsequent development of promising chemical scaffolds which have potent antiviral activity against VEEV both *in vitro* and *in vivo*. In the first studies, we reported the antiviral activity of the hit compound, CID15997213, which showed a 50% effective concentration (EC_{50}) of 0.8 μM against the VEEV strain TC-83 [106, 107]. Later, we described compounds ML336 and BDGR-4, both of which showed increased potency in inhibiting TC-83 with respective EC_{50} values of 32 and 47 nM [107, 117]. *In vitro* studies of CID15997213 and ML336 revealed mutations in VEEV TC-83 viruses when exposed to very high concentrations of each respective compound. These mutants were plaque purified and sequenced using the Sanger method. The identification of these mutations, Y102C and D116N in nonstructural protein 2 (nsP2) and the Q210K mutation in nsP4, along with our measurements demonstrating inhibition by ML336 in an *in vitro* assay using purified replication complexes, suggests the mechanism of action of these compounds interferes with the functions of those proteins in replication [118]. *In vitro* studies show that these mutations confer a 600-fold (Y102C) to a greater than 1600-fold (Q210K) loss in the EC_{50} when tested with VEEV TC-83 and ML336 [107]. Here, we explore the emergence and magnitude of resistance-conferring mutations in two distinct microenvironments, nonhuman primate (NHP) kidney epithelial cells which lack interferon-alpha (IFN-α) (Vero 76) and a human astrocyte cell line, SVGA [119].

---

*Adapted with permission from final submission. Lee J, Parvathareddy J, Yang D, Bansal S, O'Connell K, Golden JE, Jonsson CB. Emergence and Magnitude of ML336 Resistance in Venezuelan Equine Encephalitis Virus Depend on the Microenvironment. J Virol. 2020 Oct 27;94(22):e00317-20. doi: https://doi.org/10.1128/JVI.00317-20 [120]

The trajectory and dynamics of the emergence of antiviral resistance as a function of the compound concentration are an aspect of acute viral infections that has not been explored for many RNA viruses. This is due in part to the fact that there are only a handful of effective antivirals for treatment of acute RNA virus infections, such as influenza A viruses, poliovirus, and respiratory syncytial virus [121]. RNA viruses, whether they cause acute or persistent infections, constantly generate random mutations in the genome during replication due to the high error rate of their polymerase and the lack of a proofreading function. On average, one error is made per every 10,000 nucleotides replicated [122]. This leads to the creation of viral populations that contain many related, yet individually distinct, sequences within each of the virus particles [97]. Whether these random mutations are positively or negatively selected during an infection clearly depends on the microenvironment in which a virus establishes an infection and replicates. Potential selective pressures include the host cell type (neuronal versus epithelial cell) and immune response (and type of immune cell present) as potential drivers of the evolution of these populations and fixing of specific advantageous mutations. The resistance to antivirals is also considered an evolutionary, adaptive process [123].

Of note, the maintenance of resistance mutations in the virus population is often eliminated once drug treatment stops due to the tradeoffs inherent to the cost of resistance. However, this is not always the case, as observed with the seasonal H1N1 influenza virus for which a positive relationship was measured for fitness and resistance of the H274Y mutation with the neuraminidase inhibitor oseltamivir. Interestingly, serial passage of the N9 neuraminidase in a PR8 versus H1N1 A/WSN/33 (WSN) backbone with increasing oseltamivir concentration differed from that at passage 10 (1,954 μM), no mutants were observed in the N9 in the WSN backbone [124]. This suggests cis-acting genetic factors may also influence the adaptive trajectory, which underscores the importance of assessment of each antiviral drug against the strains circulating. Lastly, sofosbuvir, a hepatitis C virus (HCV) NS5B polymerase inhibitor, was the first antiviral that included resistance data using NGS [125, 126]. Establishment of robust NGS pipelines and examination of mutational variants, low or high, will be critical to interpretation of the relevance of variants over time in different cell types within an organ that may influence the metabolism and stability of the drug.

Here, we developed and deployed an NGS pipeline for whole-genome sequencing of VEEV TC-83 viral populations. In this study, we used this pipeline to define viral populations from mock-treated TC-83 and ML336-treated TC-83 over several escalating concentrations or passaging after removing ML336 entirely. The objective was to provide population-level assessments of information into the dynamics of resistance of ML336 over escalating doses. In agreement with reports of others, we show that virus populations change with respect to drug concentration, which is an important consideration in the design of effective antiviral treatment protocols [127, 128]. In other words, in selection of the effective dose, one must consider not only the safety and elimination of the virus, but the treatment regimen that obviates emergence of resistance. Additionally, we examined the VEEV TC-83 resistance dynamics to ML336 in two distinct biological systems: a nonhuman primate epithelial cell line (Vero 76) and a

human neuronal, astrocyte cell line (SVGA). We showed that the microenvironment influenced the timing of appearance, the constellation and the penetration of SNPs associated with ML336 resistance. We also showed that the nsp4 Q210 mutations are stable in the absence of compound over seven passages. As we design and test new small molecule derivatives for the alphaviruses, the pipeline developed and the results of these studies provide a framework and benchmark for the selection of compounds to advance.

Materials and Methods

Cells and Viral Culture

Vero 76 cells (ATCC CCL-131) and SVGA cells (gift from Kui Li, UTHSC) were maintained at 37°C and 5% CO_2 in complete medium containing Dulbecco modified Eagle medium (DMEM) with Hi-glucose and L-glutamine supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin. VEEV TC-83 (lyophilized vaccine) was obtained as a gift from Connie Schmaljohn (United States Army Medical Research Institute of Infectious Diseases, Maryland). Virus seed stocks were amplified in Vero 76 cells in infection medium comprising minimum essential medium with Earle's salts (MEM) with 2% FBS and 1% penicillin-streptomycin. All cell culture reagents were purchased from Thermo Fisher Scientific unless otherwise specified.

Viral Quantification

We used the median tissue culture infectious dose ($TCID_{50}$) to quantify virus in a 96-well-plate format. Plates were seeded with Vero 76 cells overnight. Eleven 10-fold dilutions of virus samples were made, and 100 µl of each dilution was added to each well in triplicate and incubated at 37°C and 5% CO_2 for 48 h. Following incubation, plates were read visually for cytopathic effect and scored. To confirm visual results, medium was removed, and 12 nM MTT [3-(4,5-dimethylthiazol-2-yl)- 2,5-diphenyltetrazolium bromide] (Acros Organics) was added to each well. Plates were incubated for 4h, after which 10% sodium dodecyl sulfate in 0.01 HCl solution was added to each well. Plates were read at an absorbance wavelength of 570 nm. Wells were scored as positive or negative, and the $TCID_{50}$ was calculated based on the Reed-Muench method [129].

Plaque assays were also used to determine viral titer. Twelve-well plates were seeded with Vero 76 cells overnight. Ten series of 10-fold dilutions of virus samples were made in infection medium, and 200 l of each dilution was added to each well in duplicate and incubated at 37°C and 5% CO_2 for 1 h. Following incubation, overlay medium consisting of a 1:1 mix of 2% carboxymethylcellulose and 2X MEM supplemented with 5% FBS, 1% L-glutamine, and 2% penicillin-streptomycin was added to each well. Plates were incubated at 37°C and 5% CO_2 for 48 h. Wells were fixed with

10% formalin, stained with crystal violet, and washed with Dulbecco's phosphate-buffered saline (DPBS), and plaques were counted.

RT-qPCR was used to determine the amount of total viral RNA in tissue samples. cDNA of each sample was amplified using 2 SYBR green Master Mix (Applied Biosystems) and the primers 5'-AGGAGCGCT GAACACTGATGAAGA-3' and 5'-AGGCGAATTCATGGAAGGGAGGAT-3'. Reactions were run on the QuantStudio 6 platform (Applied Biosystems) with an initial denaturation step at 95°C for 10 min and 40 cycles of denaturation at 95°C for 15 sec followed by annealing/extension at 60°C for 1 min. Viral RNA copy number was quantified from threshold cycle (C_T) values using a standard curve generated from a linearized plasmid standard containing a partial 250-bp sequence of the nsP2 gene.

Selection of Resistant Mutants

Prior to infection, cells were seeded onto 6-well plates and grown to 90% confluence. Cells were infected with VEEV TC-83 at an initial MOI of 1 or 0.1 as specified in the text and figure legends. Resistant viruses were selected for by passaging wild-type (WT) VEEV TC-83 in 2-fold increasing concentrations of ML336, starting at 50 nM and increasing to 100, 200, 400, 800, 1,600, and lastly, 3,200 nM, for a total of 7 passages. We used 250 µl of supernatant from each passage to initiate the subsequent blind passage. Four distinct replicates were passaged in each experiment. As a control group, VEEV TC-83 was also passaged without compound in triplicate at the same time and in the same manner as the virus was passaged with compound.

Next-Generation Sequencing

To sequence virus genomes from cell culture, 0.5 to 1 ml of supernatant from each passage was harvested and twice-clarified by centrifugation, and RNA was isolated using TRIzol LS reagent (Invitrogen). cDNA was synthesized from purified RNA using random hexamer primers and Superscript IV. cDNA was amplified through 7 or 30 rounds of PCR, using the Phusion high-fidelity kit and a set of 25 primer pairs (**Table 2-1**), using a tiling approach that doubly covered the whole genome, making PCR fragments either 500 bp or 1 kb in length. PCR products were purified using the Wizard SV gel and PCR clean-up system (Promega). Libraries for NGS were prepared using the Nextera XT kit (Illumina) and then sequenced on either the Illumina MiSeq or NextSeq 500. Data were analyzed using CLC Genomics Workbench v.12 (Qiagen). To process the sequencing data, reads less than 50 nucleotides in length and with a Phred score of 20 were removed. The remaining reads were mapped to the seed stock consensus sequence, and variants were identified either with a minimum frequency above 1% in all reads at positions with a minimum coverage of 400 or, if the quality and quantity of reads were sufficient, with a minimum frequency above 0.5% in all reads at positions with minimum coverage of 1,000 [130, 131].

Table 2-1. Primer pairs used to amplify the full genome of VEEV.

Primer Set Number	Direction	Sequence
1	F	ATAGGCGGCGCATGAG
	R	CCCTTCGTAGCGACACG
2	F	ATAGGCGGCGCATGAG
	R	GCACAAGAATCCCTCGC
3	F	CAAGTCGCTGTTTACCAGG
	R	GGTAATGAGAGGTGACGGC
4	F	TATGGGAAGCCTTCAGG
	R	TACCTGTTTACGAACTCACG
5	F	CAGGAAAATGTTAGAGGAGC
	R	GAGCGCTCTGAGAGTACC
6	F	CCACCATTGTGTACAACG
	R	TTATCCATGGGTCGCC
7	F	GCTTTTGCTTGTCATGC
	R	CATAATTGCGCAGTGTACC
8	F	GTGGAAAACACTAGCCGG
	R	AACTTCCGTCTCTTCAAGTG
9	F	AGTCTATGACATGAACACTGG
	R	AAATGGTTCAATGATTGGG
10	F	GGTATGCAAACCGAAATCC
	R	TGTTCTGGACGTGAGGC
11	F	GAACAAAGATCGACTAACCC
	R	GAAACTCCATACTCTTTGCG
12	F	ATCCATGCCATGACTCC
	R	TTTCTTGGTCGAGGCG
13	F	CCGAGACTAACTCTTACTTCGC
	R	GCATTCCACATTAAAGGCC
14	F	TTGGAGATTTCGTATGCC
	R	GACGCGATGTCAGTTTCC
15	F	GAGAATTGCCCGTATTGG
	R	TGTCATCATCATGTTCATCG
16	F	TTCTGGAAACTGACATCGC
	R	GGTTTCTTAGCGGATGGC
17	F	CTGTTTAAGCTTGGCAAACC
	R	TCCCAGCACAATAGCG
18	F	GCTAACCTGACGTTCAAGC
	R	ATCCAGGCCATACTGCG
19	F	ACGACCCATTCTGGATAACC
	R	CCTTCTTTGTGCACTGGC

Table 2-1. Continued.

Primer Set Number	Direction	Sequence
20	F	TGTTAGACTTCAGACTTCCTCG
	R	GGCGGCACAAATTGAC
21	F	GCAGTGCAGAGCATATCG
	R	GGCCTGTAAGTTGGAGTGC
22	F	ACCACAGATACCCTATGTCC
	R	CCAAGCTGTGGAAAGCGG
23	F	AGGAATGGATTCACCAGCC
	R	TGACAATCACCTTTGCACG
24	F	CTCCTGTGAATTTCAATGG
	R	AACAAAATCCGATTCGG
25	F	CCATCAGGGACTGCTACC
	R	AACAAAATCCGATTCGG

In vivo Assessment of VEEV TC-83 in the Presence and Absence of ML336 by NGS

All mouse studies were reviewed by and approved by the UTHSC IACUC in animal care and use protocol number 17.057. Five- to six-week-old, female, C3H/HeN mice were obtained from Charles River Laboratories and randomly assigned to one of three treatment groups as follows: group 1, uninfected control with vehicle; group 2, TC-83 and vehicle only; group 3, TC-83 and ML336 at dosing concentration of 24 mg/kg/day. Mice were dosed via intraperitoneal (i.p.) injection twice per day with vehicle (25% polyethylene glycol 400, 10% Kolliphor RH40, 65% water) or with vehicle containing 12 mg/kg ML336, starting 2 h before infection for 8 days. Mice were infected intranasally (i.n.) with VEEV TC-83 diluted to a concentration of 1 × 107 PFU/30 µl, 15 µl per naris in PBS. PBS was used in place of virus for the vehicle-only control group. Mice were weighed daily and checked twice daily for mortality and morbidity, and subgroups of the VEEV TC-83 infected mice (four mice in each subgroup) were humanely sacrificed on day 4, 7, or 14 post infection (dpi) or when they either reached >30% weight loss of the starting body weight, or close to 30%, and a clinical score of 0.8 on a euthanasia score sheet approved by the IACUC.

Following euthanasia, the whole brain, including the olfactory bulbs, was removed from each mouse. One half of the brain was homogenized in 1 ml of DPBS using the Omni bead mill. We added an equal volume of TRIzol reagent (Invitrogen) to extract total RNA from brain tissue. Because brain tissue has a high lipid content, the RNA from TRIzol was further purified; 2.5 µl of the total RNA, equivalent to 5 mg of homogenized brain originally extracted in TRIzol reagent, was purified with the MagMAX mirVana using the KingFisher Flex system (Thermo Scientific). RNA concentration was measured using a Qubit fluorometer 4 (Invitrogen) running the Qubit RNA HS assay kit (Invitrogen) and stored at -20°C for RT-qPCR to be carried out the following day. Approximately 20 µg of total RNA was obtained from 5 mg of tissue. RT-qPCR was conducted to quantitate viral RNA in mouse brain specimens. Then, 400 ng RNA per sample was taken for assessment using the Superscript III Platinum one-step qRT-PCR kit (Invitrogen) using a TaqMan probe (5'-6-FAM-ACTGGGCTAGGGCTCACAGGCGAG-[TAMRA]-3') (Applied Biosystems) and primers to the VEEV TC83 nsp2 gene (forward, 5'-AGGAGCGCTGAACACTGATGAAGA-3'; reverse, 5'-AG GCGAATTCATGGAAGGGAGGAT-3'). The RT-qPCR was run on a Quantstudio 6 Flex system for 40 cycles. Viral copy number in the brain samples was calculated using a standard curve made from 10-fold dilutions of VEEV TC83 RNA from the master seed stock. For NGS, the cDNA, PCR amplification and NGS were performed as described above. In brief, mouse brains were normalized as half a brain per ml of DPBS. We measured the C_T values from the brain for the cDNA amplification step. We also normalized the amplicons with our pooling strategy to obtain even coverage. Lastly, we normalized each sample to the same concentration of amplicon prior to library production.

Network Analysis

The whole-genome consensus sequences from each sample were obtained using CLC Genomics Workbench v.12 (Qiagen). Sequences were aligned using MUSCLE and mapped by the program PopART v.1.7, utilizing a minimum spanning tree to plot each sequence.

Data Availability

All the NGS data from the studies here are publicly available under BioProject accession number PRJNA655204.

Results

Development of a Whole-Genome Next-Generation Sequencing Approach to Assess Single Nucleotide Variation in VEEV TC-83

In order to accurately identify single nucleotide polymorphisms in VEEV TC-83, we needed a method to generate enough sequences to provide sufficient coverage and depth across the entire genome to detect low-frequency variants. To achieve this goal, we developed a tiling method for whole-genome amplification in which we designed primers to create amplicons of approximately 1,000 or 500 bp long while ensuring the entire VEEV TC-83 genome was covered twice (**Figure 2-1A**). In total, 25 primer pairs were identified as providing an efficient, and a similar level of, amplicon product using RNA isolated from our seed stock of wild-type (WT) VEEV TC-83 (**Table 2-1**). To ensure that the correct sequences were amplified, amplicons were purified by agarose gel electrophoresis and Sanger sequencing. This also provided a reference for the VEEV TC-83 genome based on Sanger sequencing. Primer sequences were optimized over time to produce equivalent amounts of product from the PCR. Initially, all amplicons were evaluated separately to ensure evenness in coverage. Second, we evaluated various approaches to pooling amplicons for sequencing. The optimal pooling noted from assessment of NGS data was one in which we sorted each primer pair into five different groups for PCR amplification (**Figure 2-1A**). We pooled the PCR products from those groups together before purifying, preparing libraries, and conducting NGS.

We performed NGS on the pooled products to assess the coverage and depth of sequencing data we could obtain. We examined the data on low-frequency variants at a confidence level of 99.999%, according to standards outlined by Ladner et al. [130, 131]. In the alignment step, we used a full genome sequence published on GenBank of VEEV TC-83 (accession number L01443.1, GI: 323714) to provide an initial reference [11]. Thereafter, we used our consensus sequence as the reference. As shown in **Table 2-2**, there were two differences between the GenBank reference sequence and the consensus sequence of our seed stock virus. We noted nucleotide changes of T7208C and C11386T.

Figure 2-1. Methods for amplifying viral genomes for sequencing and detecting small nucleotide variations, and for passaging TC-83 through increasing concentrations of ML336.

(A) Schematic for the tiling PCR method. Each colored bar represents an amplicon generated by a singular primer pair and is placed relative to the portion of the VEEV TC-83 genome it covers. Primer pairs of the same color were pooled together during PCR setup. Pools were combined prior to DNA purification. (B) Sample coverage graphs from applying the tiling method and deep sequencing the resulting amplicons. WT VEEV TC-83 was used to produce these graphs, using a total of 622,583 mapped reads. The top graph shows the depth and coverage throughout the genome, while the bottom graph shows coverage using a cutoff of 1,000 depth. (C) Schematic for the passaging of VEEV TC-83 in the presence of 2-fold increasing concentrations of ML336. The starting concentration was 50 nM. Vero 76 cells were grown for 2 days in a six-well plate and infected at the start of the experiment with 1 MOI. For each experiment, replicates are marked as R1, R2, R3, and R4.

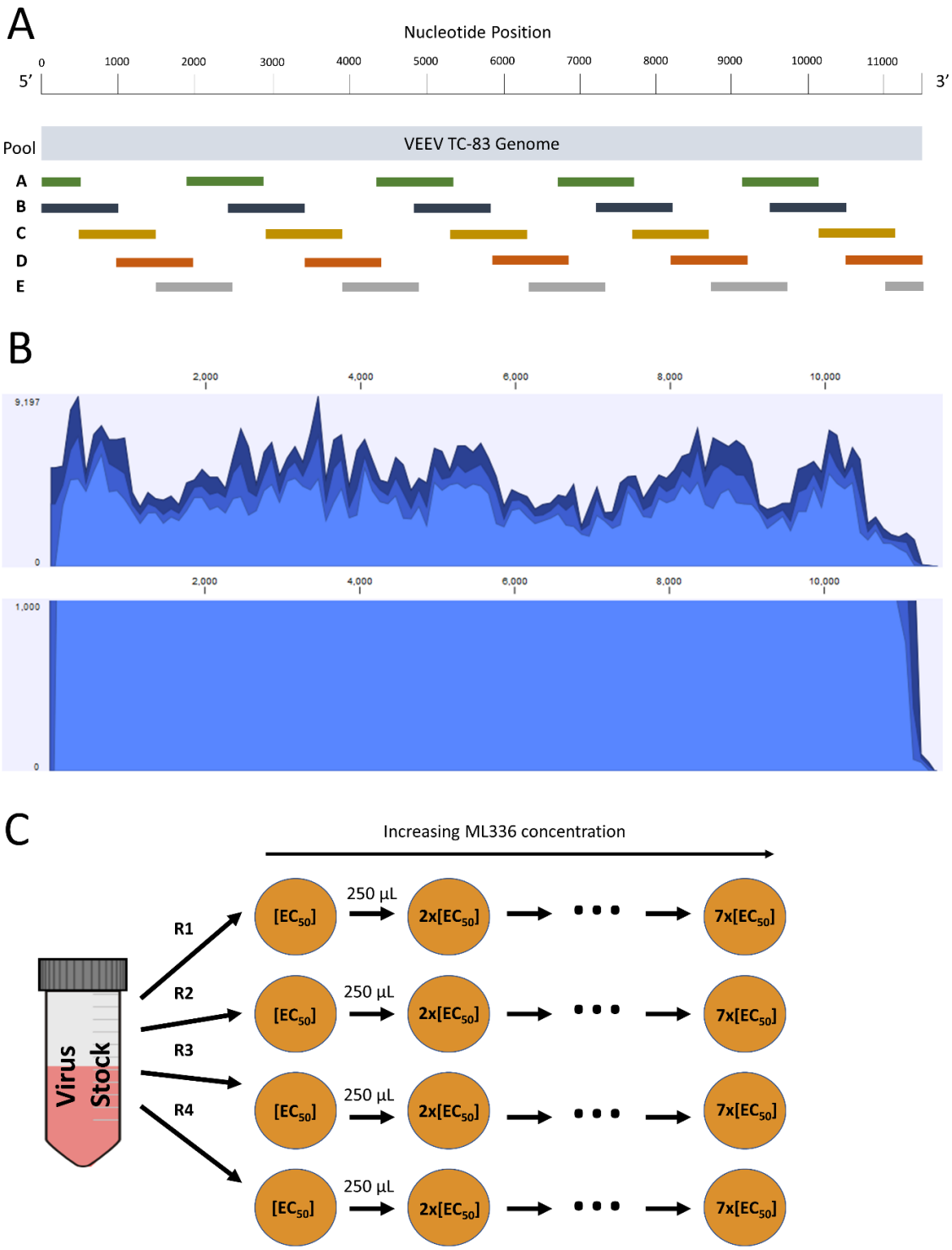

A
Nucleotide Position
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
11000
5'
3'
Pool
VEEV TC-83 Genome
A
B
C
D
E
B
2,000
4,000
6,000
8,000
10,000
9,197
0
1,000
C
Increasing ML336 concentration
R1
R2
R3
R4
Virus Stock
[EC50]
250 µL
2x[EC50]
7x[EC50]

Table 2-2. Summary of SNPs and amino acid changes between five samples of WT VEEV TC-83 and reference sequence.

Type	Nucleotide Change	Amino Acid Change	Gene	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
SNP	C1151T	N369N	nsP1	7.26	5.11	5.14	5.92	6.34
Insertion	-2253A	N/A	nsP2	1.86	3.83	2.42	-	-
Insertion	-2501T	N/A	nsP2	3.17	3.12	3.44	-	-
SNP	C2627A	D326E	nsP2	-	1.51	1.08	-	-
Insertion	-2628A	N/A	nsP2	4.61	6.36	5.49	5.01	2.78
SNP	T2873A	H408Q	nsP2	-	1.04	-	-	-
SNP	A5151C	I374L	nsP3	-	1.4	1.58	-	-
SNP	A5573G	S514S	nsP3	-	1.05	-	-	-
SNP	T7208C	P502P	nsP4	99.86	99.74	99.95	99.94	99.88
SNP	G7787A	G76R	C	1.36	-	-	-	-
Insertion	-9964T	N/A	6K	-	2.98	3.24	-	-
SNP	C9987T	A52V	6K	-	1.79	1.7	-	-
SNP	T10759C	F253F	E1	1.82	-	-	-	-
Insertion	-10768T	N/A	E1	-	-	-	1.52	-
SNP	G11149T	P383P	E1	1.57	1.05	-	-	-
SNP	C11386T	noncoding	N/A	99.75	100	100	100	100

All SNPs appearing over 1% in at least one sample are shown. Numbers under sample columns indicate the percentage of SNPs present in that sample compared to the reference sequence. A dash indicates the SNP did was not observed above 1% in that sample. Reference sequence used was GenBank Accession L01443.1.

These mutations are located in the nsP4 gene and the 3' untranslated region (UTR) of the genome, respectively,and as such, neither causes an amino acid change. The mapped reads showed 100% coverage throughout the genome (**Figure 2-1B**). In addition, a more than 1,000× depth of coverage was obtained except at the very end of the 3'UTR, surpassing the benchmarks necessary for confidence in identification of variants. In summary, the tiling method was sufficient in providing us the amount of sequencing reads necessary to detect and call variants.

Emergence and Magnitude of VEEV TC-83 Resistance to ML336 at Increasing Concentrations

To determine how VEEV TC-83 evolves and adapts to ML336 at different concentrations, we infected VEEV TC-83 at a multiplicity of infection (MOI) of 1.0 in Vero 76 cells in a 6-well plate in which each well had 600,000 cells (**Figure 2-1C**). This provided a sufficient population size to determine if any mutants were preexisting in the mock-treated group and for sampling the standing genetic variation/sequence space of the seed stock (approximately 10^9/ml). VEEV-TC83 grows rapidly to very high titers, and hence, at this MOI, we were sampling approximately 0.6% of the available seed stock virions. All virus was removed following infection by washing with Dulbecco's phosphate-buffered saline (DPBS), after which fresh medium or medium with compound was added. We started the first passage at 50 nM, a concentration slightly above the EC50 (30 nM) and increased the concentration of ML336 2-fold at each passage, namely, 100-, 200-, 400-, 800-, 1,600-, and 3,200-nM concentrations [107]. The experiment included four biological replicates (R1 to R4). As a control, VEEV TC-83 was passaged in the absence of ML336 in each experiment in an identical manner as the virus with compound. RNA was isolated from twice-clarified supernatant at each passage, and we sequenced each sample as described in Materials and Methods. We also assessed the amount of virus after each of the passages R1 and R2 (**Figure 2-2**).

Each replicate at each concentration was evaluated for SNPs (**Figure 2-3**, **Table 2-3**). We only considered SNPs that reached a percentage of over 10% of the population, as those that emerged between 1 and 10% were usually not stable and did not penetrate over time. Replicate 1 (R1) produced the highest number of SNPs, with 12 in total. In R2, four SNPs became dominant in the population by the third passage. Replicates R3 and R4 only saw two major SNPs emerge. The identities of the SNPs detected were distinct; however, R1, R2, and R4 all showed a Q210R mutation in nsP4. R1 and R2 shared an E97Q SNP in nsP2 and a Y87F SNP in nsP4, as well as a synonymous mutation at position A461 in nsP3. R1 was the only replicate where a significant percentage of SNPs in structural genes emerged, all within in the E1 gene and all synonymous mutations. A second nsP4 Q210 mutation, with an amino acid change into histidine, was observed in R1 at a lower frequency than Q210R. The emergence of E118V and G786G in nsP2, and A201V and H414Y in nsP4 were also unique to R1. A nucleotide change at R305 in nsP4 was found in R2 only. The I268I mutation only appeared in R3, while the only nsP1 mutation, R220C, was found in nsP4. No SNPs above our variant calling cutoff were detected in virus-only controls where TC-83

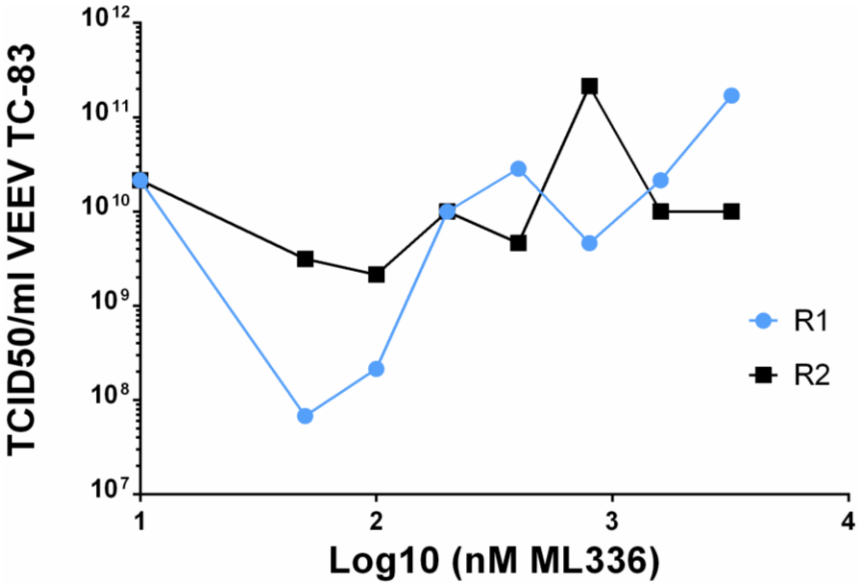


Figure 2-2. $TCID_{50}$ data for R1 and R2 from passaging VEEV TC-83 in Vero 76 cells through increasing concentrations of ML336.

For the biological replicates R1 and R2, we measured the titer of the virus after each passage at the time the supernatant was collected (48 hours post infection [hpi]) and clarified. All supernatant was stored at –80°C until virus titers were assessed using the TCID50 assay, which is reported in this figure as $TCID_{50}$/ml.

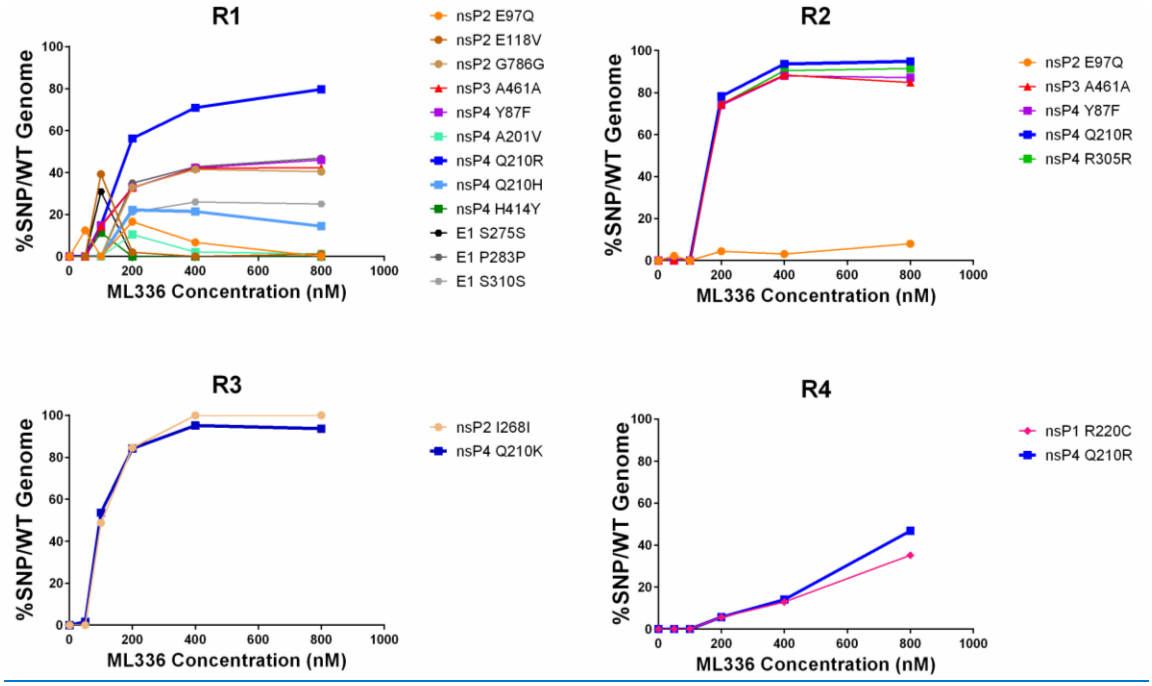


Figure 2-3. Four biological replicates of VEEV TC-83 passaged in Vero 76 cells over escalating doses of ML336 reveal four distinct paths to a common solution of resistance.

Each well with approximately 600,000 cells was infected with an initial MOI of 1, washed to remove any unbound virus, and then passaged in the presence of 50-, 100-, 200-, 400-, 800-, 1,600-, or 3,200-nM concentrations. The first five passages (50, 100, 200, 400, 800) of each replicate, R1, R2, R3, and R4, are plotted. Each SNP that was found in at least 10% of all reads in at least one passage was defined and graphed for each passage/ML336 concentration. Each unique SNP is represented by an individual color.

Table 2-3. Summary of SNPs and amino acid changes found at increasing concentrations of ML336 in Vero 76.

Nucleotide Change	Amino Acid Change	Gene	Replicate	50 nM	100 nM	200 nM	400 nM	800 nM	1600 nM	3200 nM
C692T	R220C	nsP1	R4	-	-	5.65	12.95	35.18	50.61	33.17
C920T	I292I	nsP1	R3	-	9.59	3.59	-	-	-	-
G1938C	E97Q	nsP2	R1	12.42	-	16.55	6.72	-	20.34	-
			R2	2.25	-	4.42	3.08	8.01	15.57	20.6
A2002T	E118V	nsP2	R1	-	39.24	2.06	-	1.22	1.63	2.16
T2453C	I268I	nsP2	R3	-	48.76	84.54	>99.99	>99.99	>99.99	>99.99
T4007C	G786G	nsP2	R1	-	-	32.99	41.52	40.43	39.84	31.60
G5414A	A461A	nsP3	R1	-	14.24	32.72	42.12	42.28	43.81	38.16
			R2	-	-	74.60	88.35	84.82	82.66	66.62
A5962T	Y87F	nsP4	R1	-	14.9	32.93	42.26	45.95	48.88	51.23
			R2	-	-	73.84	88.01	87.14	88.04	84.43
C6330A	Q210K	nsP4	R3	1.67	53.61	84.16	95.15	93.62	95.72	95.59
A6331G	Q210R	nsP4	R1	-	14.9	56.25	70.89	79.65	85.50	89.94
			R2	-	-	78.22	93.61	94.89	96.37	95.06
			R4	-	-	5.72	14.04	46.75	85.10	85.12
A6332T	Q210H	nsP4	R1	-	-	22.20	21.48	14.46	9.73	5.14
G6617A	R305R	nsP4	R2	-	-	74.32	90.42	91.54	92.68	92.59
C6942T	H414Y	nsP4	R1	-	11.4	-	-	-	-	-
C7367T	R555R	nsP4	R1	1.03	-	10.00	2.15	-	-	-
A10825T	S275S	E1	R1	-	30.90	-	-	-	-	-
C10849T	P283P	E1	R1	-	-	34.99	42.79	46.83	48.27	45.84
C10930T	S310S	E1	R1	-	-	21.08	25.95	25.00	25.91	24.81

was passaged without the presence of ML336 (data not shown).

In R1, R2, and R4, SNPs appeared at a concentration of 100 nM or 200 nM ML336 and increased rapidly up to 400 nM, after which the magnitude of SNP penetration appeared to plateau, especially if the percentage of the SNP in the population reached over 90%. However, in R1, a few SNPs, including H414Y (nsP4) and S275S (E1), appeared at 100 nM and then immediately disappeared from the population in subsequent passages. The E97Q mutation in R1 and R2 showed two distinct trajectories, where it would either disappear and reappear or slowly increment in the population. Both SNPs that reached consensus in R3 appeared at significantly high percentages at 100 nM. Lastly, both major SNPs found in nsP4 exhibited trajectories where both SNPs increased in percentage at a much slower rate, with the Q210R mutation only peaking at 3,200 nM (**Table 2-3**). In summary, these results show that TC-83 evolved resistance to ML336 through a variety of paths but converged in a resistance solution at the Q210 amino acid position.

Mutant Populations Stayed Consistent When Passaging in the Absence of ML336

To determine whether the mutations generated in the previous experiment were stable, each replicate was passaged an additional seven times in Vero 76 cells without the presence of ML336. The starting virus sample used was from the 3,200-nM supernatant from each replicate in the previous experiment (referred to as VEEV TC-83/3200), which was inoculated at an MOI of 1. At each passage, virus supernatant was collected and sequenced in the same manner as before. As shown in **Figure 2-4** and **Table 2-4**, the major mutations in each population remained similar through each passage, although there were two interesting cases. First, the nsP2 E97Q mutation in R2 disappeared from the population after one passage and did not reappear. Second, the R220C amino acid change in nsP1 in R4 fluctuated in percentage through the seven passages, but we observed that all the percentages were higher than when the virus was passaged in the presence of ML336. In conclusion, these combinations of mutations were stable and did not cause a loss in fitness to the viral population in the absence of ML336, since they were maintained well through repeated passages.

Passaging TC-83 Through Increasing Concentrations of ML336 in the SVGA Cell Line

The Vero 76 cell line is deficient in type-I IFN production, thus creating a cellular environment that allows viruses to easily proliferate [119]. In order to test the emergence of resistance to ML336 in an environment that more closely resembles that of a natural infection and creates a higher barrier to replication, we selected the SVGA cell line as our model. The SVGA cell line is derived from human astrocyte cells, and VEEV has been shown to replicate in astroglia in the central nervous system [30, 31]. To show SVGA was a viable model for VEEV TC-83 infection, we performed a time course infection of SVGA cells.TC-83 reached a titer of over 1×10^9 median tissue culture infectious dose

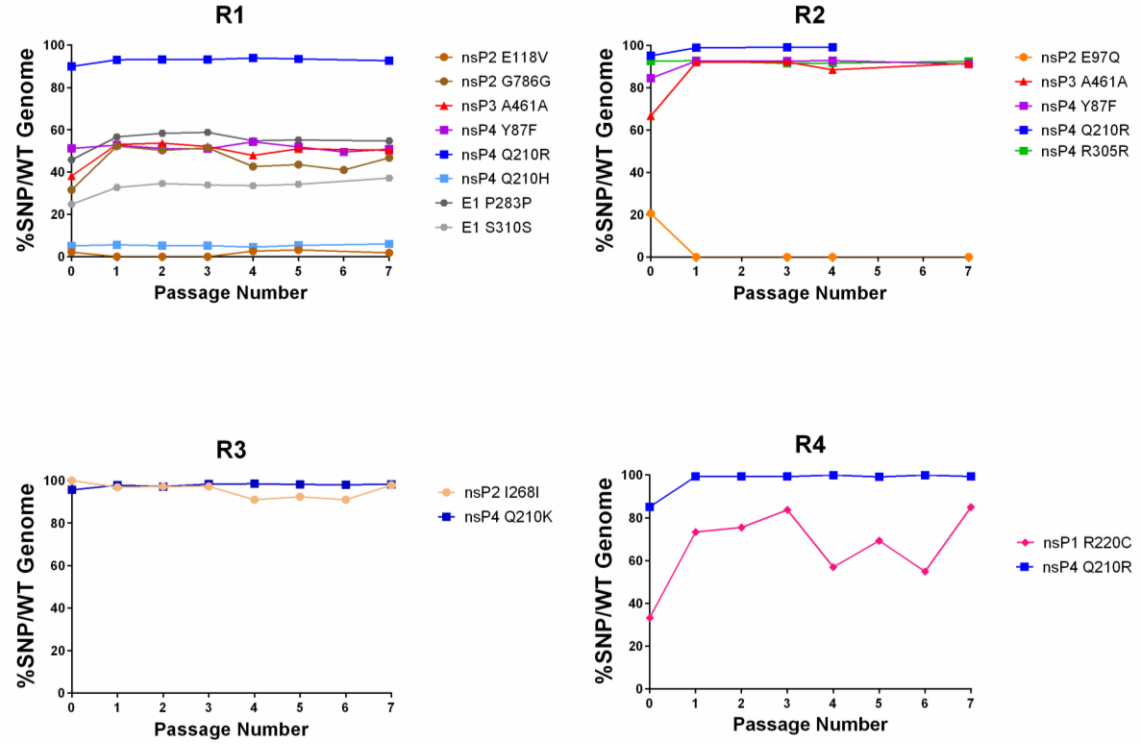


Figure 2-4. Most SNPs show stability over seven passages in the absence of ML336.

Supernatant from VEEV TC-83/3200 was used to infect a six-well plate seeded with Vero 76 cells for 1 h at an initial MOI of 1. After 1 h, medium was removed and replaced with fresh complete MEM. Supernatant from each passage was collected at 24 h and subjected to NGS as described in Materials and Methods. Then, 250 µl of supernatant was used to infect new cells (passage 2). Passages were repeated 5 additional times. The y axis shows the percentage of SNPs relative to the WT population.

Table 2-4. Summary of the SNPs and amino acid changes for VEEV TC-83/3200 passaged in the absence of ML336. P1, P2, etc. represent each successive passage.

Replicate	Nucleotide Change	Amino Acid Change	Gene	3200 nM (Original)*	P1	P2	P3	P4	P5	P6	P7
R1	A2002T	E118V	nsP2	2.16	-	-	-	2.69	3.21		1.83
	T4007C	G786G	nsP2	31.60	52.29	50.26	51.48	42.64	43.57	41.04	46.78
	G5414A	A461A	nsP3	38.16	53.1	53.7	52.11	47.9	51.01		50.37
	A5962T	Y87F	nsP4	51.23	52.8	51.19	51.04	54.33	51.9	49.53	50.85
	A6331G	Q210R	nsP4	89.94	93.13	93.3	93.24	93.93	93.53		92.69
	A6332T	Q210H	nsP4	5.14	5.62	5.25	5.23	4.57	5.35		6.07
	C10849T	P283P	E1	45.84	56.7	58.35	58.8	54.84	55.24		54.76
	C10930T	S310S	E1	24.81	32.72	34.64	33.95	33.61	34.25		37.17
R2	G1938C	E97Q	nsP2	20.6	-		-	-			
	G5414A	A461A	nsP3	66.62	91.92	92.47	92.13	88.45			91.48
	A5962T	Y87F	nsP4	84.43	92.56		92.59	92.76			91.96
	A6331G	Q210R	nsP4	95.06	98.85		99.1	99.07			
	G6617A	R305R	nsP4	92.59	92.71		91.46	91.55			92.46
R3	T2453C	I268I	nsP2	-	96.81	97.14	97.22	90.92	92.32	90.92	97.84
	C6330A	Q210K	nsP4	95.59	97.8	97.13	98.25	98.45	98.12	98.02	98.16
R4	C692T	R220C	nsP1	33.17	73.31	75.47	83.65	56.93	69.25	54.86	84.95
	A6331G	Q210R	nsP4	85.12	99.25	99.30	99.26	99.99	99.13	99.99	99.37

*Values under the 3200 nM column are the same as the 3200 nM column in **Table 2-3**. Blanks indicate depth of coverage was too low to analyze at that nucleotide position.

($TCID_{50}$)/ml by 24 h, which is in line with titers from infection of Vero 76 cells (data not shown). Thus, we showed that TC-83 was able to productively infect the cell line.

With this result, we passaged and sequenced TC-83 in SVGA in the same manner as with Vero 76, using the same ML336 concentrations as before. To serve as a control, we passaged the virus without the addition of ML336. We sequenced all four replicates at each concentration (**Figure 2-5A**, **Table 2-5**) and three replicates of virus without ML336 (**Figure 2-5B**, **Table 2-5**). In all four replicates passaged with ML336, the Q210H mutation arose to become the most dominant mutation, again appearing within 100 to 200 nM of ML336. In addition, the Q210R mutation also appeared in all replicates at a similar timing, though in each replicate it followed a trajectory where it was less prominent than the Q210H mutation. Interestingly, in R1, we discovered a second nucleotide mutation in the Q210H mutation in which cytosine instead of the usual thymine arose at nucleotide position 6332, and the cytosine mutant became more dominant than the thymine as ML336 concentration increased in subsequent passages. The trajectories of each replicate compared to each other were slightly more uniform, as the A461 synonymous mutation in nsP3 and the Y87F and R305 mutations in nsP4 appeared in all replicates, starting at 100 nM. The A425V mutation in nsP4 was shared between R1 and R4, the E213D mutation in nsP4 was shared between R1 and R2, and the E118V mutation was shared between all replicates except R2. SNPs unique to each replicate mostly occurred in nsP2, where the P644S (R1), D116N (R2), and E87A (R4) mutations appeared. R2 also contained a D253D mutation in nsP1, and mutations in E1 appeared in R1.

Lastly, in the virus-only controls, the only SNPs that emerged were located in genes coding for structural proteins (**Figure 2-5B**, **Table 2-5**). Notably, the E3K mutation in the E2 gene was found in both R1 and virus control 3 (VC3), while the P78S mutation was found in both VC2 and VC3. A unique Y107Y mutation also emerged in VC1. Taken together, these results show that the mutation trajectory of TC-83 when subjected to ML336 diverges in SVGA cells compared to Vero 76 cells.

NGS of VEEV TC-83 in Brains from a Lethal Mouse Model Treated with ML336 Did Not Detect Resistance Variants

Next, we aimed to examine the potential of ML336 resistance in vivo, using the mouse strain C3H/HeN, in which TC-83 can cause a lethal infection at 1×10^7 PFU [132]. Previously, we reported that intraperitoneal (i.p.) administration of ML336 at 25 mg/kg/day, twice per day (BID), gave complete prophylactic protection (n = 8) when challenged with VEEV TC83 [132]. Based on these data, we selected time points at days 4, 7, 8, and 14 to examine brains for the presence of viral RNA by MiSeq NGS using the aforementioned approach developed for our *in vitro* studies. Groups of C3H/HeN mice were infected with 1×10^7 PFU TC-83 and treated with either 24 mg/kg/day ML336 or vehicle, starting 2 h before infection for 8 days. Mice were

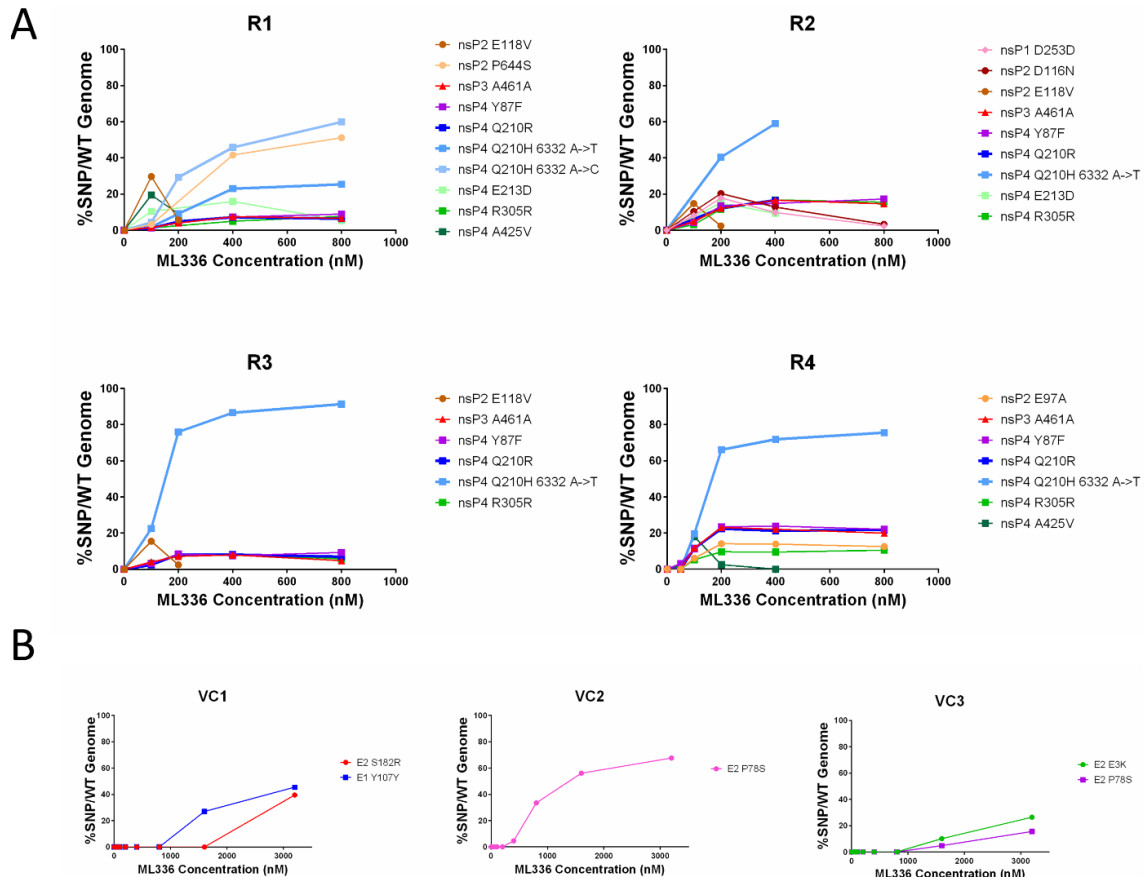


Figure 2-5. Passaging TC-83 in SVGA cells through increasing concentrations of ML336 showed that ML336 exerted a stronger environmental pressure.

Each well was infected with an initial MOI of 0.1. All mutations appearing in at least 10% of all sequence reads found in at least one passage were included, plotting their percentage in the population over each ML336 concentration. (A) The first five passages of each experimental replicate were plotted. A distinction has been made between the two Q210H mutations to show that the A to T mutation is more common. (B) Graphs of three virus-only controls, showing all passages. All mutations occurred in genes coding for envelope proteins.

Table 2-5. Summary of SNPs and amino acid changes found at increasing concentrations of ML336 in SVGA.

Nucleotide Change	Amino Acid Change	Gene	Replicate	50 nM	100 nM	200 nM	400 nM	800 nM	1600 nM	3200 nM
C803T	D253D	nsP1	R2	-	8.16	17.91	9.82	2.33	-	-
A1939C	E97A	nsP2	R4	-	6.02	14.06	13.86		14.62	18.18
G1995A	D116N	nsP2	R2	-	10.34	20.31	12.81	3.28	-	-
A2002T	E118V	nsP2	R1	-	29.64	5.62	-	-	-	-
			R2	-	14.83	2.34	-	-	-	-
			R3	-	15.53	2.46	-	-	-	-
C3579T	P644S	nsP2	P1	-	3.19	-	41.57	51.17	55.36	45.54
G5414A	A461A	nsP3	R1	-	1.56	4.02	7.48	6.6	10.36	11.32
			R2	-	4.9	12.04	16.63	14.59	22.94	33.26
			R3	-	4.09	7.17	7.91	4.82	7.53	12.01
			R4	-	11.74	23.06	22.09		21.6	35.88
A5962T	Y87F	nsP4	R1	-	1.27	3.95	7.28	8.94	7.55	9.99
			R2	-	4.15	13.55	14.82	17.75	25.65	34.43
			R3	-	3.21	8.42	7.52	9.32	9.72	13.65
			R4	3.19	11.29	23.38	23.81		25.35	34.12
A6331G	Q210R	nsP4	R1	-	1.19	4.84	7.15	6.45	8.01	10.61
			R2	-	-	12.28	16.56	-	19.12	36.27
			R3	-	2.44	8.17	8.11	6.99	9.34	13.56
			R4	-	11.51	22.38	21.26		24.85	30.37
A6332T	Q210H	nsP4	R1	-	1.71	9.18	22.94	25.32	35.36	35.74
			R2	-	-	40.32	58.91	-	77.3	61.54
			R3	-	22.55	75.97	86.54	91.33	89.24	84.69
			R4	-	19.6	66.07	71.83		71.46	66.67

Table 2-5. Continued.

Nucleotide Change	Amino Acid Change	Gene	Replicate	50 nM	100 nM	200 nM	400 nM	800 nM	1600 nM	3200 nM
A6332C	Q210H	nsP4	R1	-	4.38	29.28	45.75	59.85	55.26	51.95
A6341T	E213D	nsP4	R1	-	10.52	-	15.88	5.34	-	-
			R2	-	-	15.08	9.31	-	-	-
G6617A	R305R	nsP4	R1	-	-	-	4.95	7.69	6.87	12.16
			R2	-	3.03	11.55	16.77	15.51	21.85	32.26
			R3	-	3.27	7.79	8.04	5.82	9.31	12.64
			R4	-	5.24	9.63	9.49		11.18	13.26
C6976T	A425V	nsP4	R1	-	19.45	9.03	-	-	-	-
			R4	1.81	17.93	2.47	-		-	-
G8570A	E3K	E2	R1	-	-	-	-	-	4.5	13.11
			VC3	-	-	-	-	-	10.09	26.4
C8795T	P87S	E2	VC2	-	-	-	4.63	33.49	56.07	67.62
			VC3	-	-	-	-	-	4.81	15.54
A9107C	S182R	E2	VC1	-	-	-	-	-	-	39.52
A9210G	Q216R	E2	R1	-	-	-	-	-	5.31	11.83
C10321T	Y107Y	E1	VC1	-	-	-	-	-	26.96	45.45

monitored for survival and weight change (**Figure 2-6A** and **B**).

As shown previously, treatment of mice with ML336 at 24 mg/kg/day over 8 days provided protection (n = 7/8) through day 14, with only a single mouse succumbing to disease at 8 dpi (**Figure 2-6A**) [107, 132]. In the group with VEEV TC83, a sharp decline in body weight was observed beginning at 5 dpi, whereas the ML336-treated mice continued to have a similar level of weight gain as the noninfected mice (**Figure 2-6B**). Brains from sacrificed mice at 4, 7, 8, or 14 dpi were homogenized, and RNA was extracted and processed for quantification through reverse transcriptase quantitative PCR (RT-qPCR) to confirm the amount of virus present before NGS (**Figure 2-6C**). SNPs detected in the brains of TC-83-infected mice at percentages above 10% are shown in **Table 2-6**. The ML336- treated mouse brains did not result in a sufficient number of sequencing reads that allowed for analysis and featured uneven depth of coverage; however, we were able to pull SNPs from the 6K portion of the genome for one sample where SNPs were observed above 10% penetration. None of the SNPs found in this experiment had been previously found to confer resistance to ML336.

Network Analysis of TC-83 in Different Microenvironments Reveals Different Trajectories of Resistance to ML336

After determining the major SNPs that emerged in the previous in vitro and in vivo experiments, we aimed to analyze these data using network analysis. Consensus sequences for the full-length genome for each of the samples from each of the experiments described above, along with the original reference sequence (accession number L01443.1) and sequences from the seed stocks used for passaging, were aligned with the MUSCLE algorithm and plotted using PopART v.1.7, which allowed us to visualize sequence differences [133]. We used a minimum spanning tree to plot the sequences, which are shown in **Figure 2-7**. The resulting graph shows that TC-83 lineages diverged based on the host (Vero 76, SVGA, or mouse brain) from which they were isolated. We observed no overlap in trajectory between virus sequences from Vero 76-derived and SVGA-derived samples. In addition, samples in which TC-83 was passaged through SVGA or used to infect C3H/HeN mice, but without addition of ML336, were included, and these samples also resulted in distinct network branches. The network analysis shows the random and ML336-influenced genetic diversity of the VEEV TC-83 virus population in each environment. Of note, the SNPs detected from the wild-type virus diverged from the original consensus in one SVGA-passaged virus and in TC-83 isolated from mouse brains.

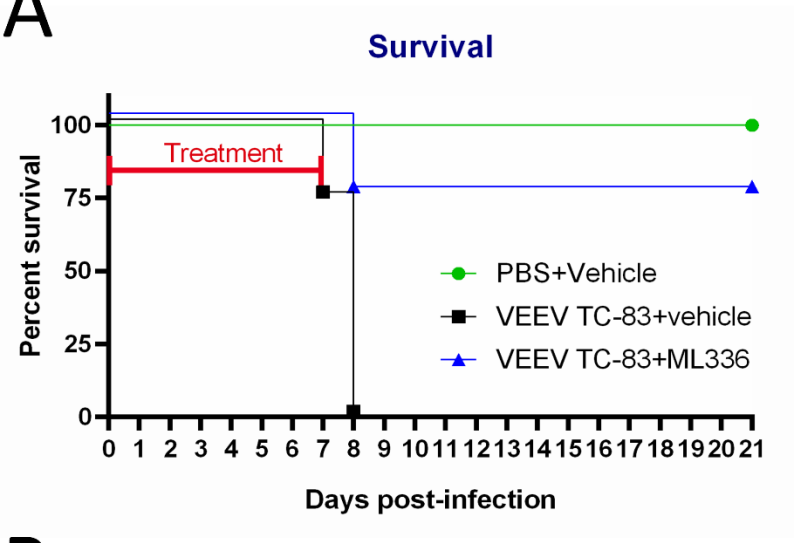
Discussion

In this study, we examined the dynamics of the evolution of resistance to ML336 in VEEV strain TC-83 in three distinct environments. Our experiments were designed to address how different environments for the virus would affect these trajectories. Our network analysis shows that the microenvironment does indeed have an impact on

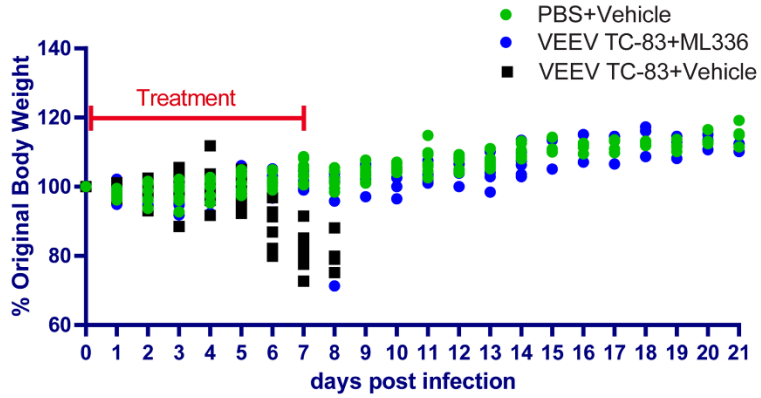
Figure 2-6. Time points chosen for NGS. In vivo efficacy of ML336 tested in C3H/HeN mice challenged with VEEV TC-83.

Mice were i.n. infected with TC-83 and treated with either ML336 (i.p.; 24 mg/kg/day) or vehicle for 8 days starting from 0 dpi. (A and B) Survival (A) and weight change (B) were monitored over 14 days. The survival rates were compared between the ML336-treated and placebo-treated groups using Log-rank (Mantel-Cox) test. (C) RT-qPCR data for each individual mouse. *** represents $P < 0.001$. Mouse brains were collected on days 4, 7, 8, and 14 for RNA extraction.

A


Survival
Percent survival
100
75
50
25
0
Treatment
PBS+Vehicle
VEEV TC-83+vehicle
VEEV TC-83+ML336
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
Days post-infection

B


PBS+Vehicle
VEEV TC-83+ML336
VEEV TC-83+Vehicle
% Original Body Weight
140
120
100
80
60
Treatment
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
days post infection

C

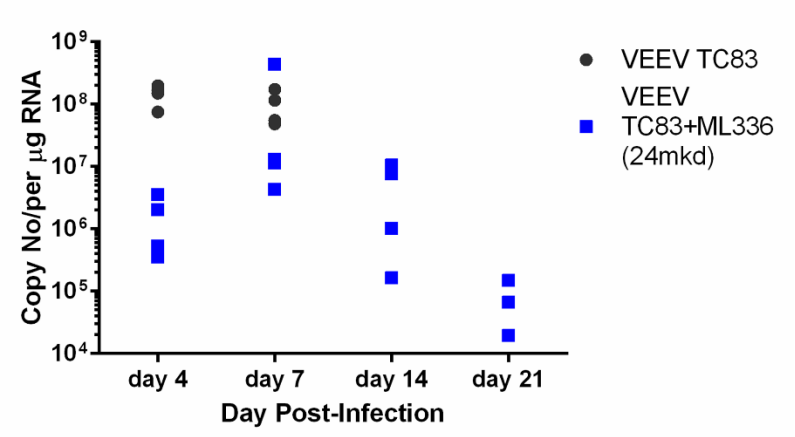

Copy No/per μg RNA
10⁹
10⁸
10⁷
10⁶
10⁵
10⁴
VEEV TC83
VEEV TC83+ML336 (24mkd)
day 4
day 7
day 14
day 21
Day Post-Infection

Table 2-6. Summary of point mutations and amino acid changes found in TC-83 extracted from C3H/HeN mouse brains.

Nucleotide Change	Amino Acid Change	Gene	Treatment	Sample	Day Euthanized	SNP Percentage
T1657C	V3A	nsP2	PBS	B1	7	13.68
G3033A	E462K	nsP2	PBS	C4	7	11.68
C3070T	A474V	nsP2	PBS	D3	8	84.68
C3458T	N603N	nsP2	PBS	C1	8	17.81
G3831A	A728T	nsP2	PBS	D2	7	65.92
C4024T	A792V	nsP2	PBS	D1	8	23.38
A4894G	Q288R	nsP3	PBS	B2	7	52.4
C5606G	S525R	nsP3	PBS	D1	8	13.13
G7282A	R527K	nsP4	PBS	C2	7	13.14
G8042-	N/A	C	PBS	C4	7	12.52
A8108-	N/A	C	PBS	D1	8	15.42
C9481T	Y306Y	E2	PBS	C1	8	16.78
C9850T	S6S	6K	ML336	C4	14	84.78
A9862G	L10L	6K	PBS	C2	7	10.74
			ML336	C4	14	11.56
C10861T	F287F	E1	PBS	A2	4	81.27
			PBS	D3	8	9.52

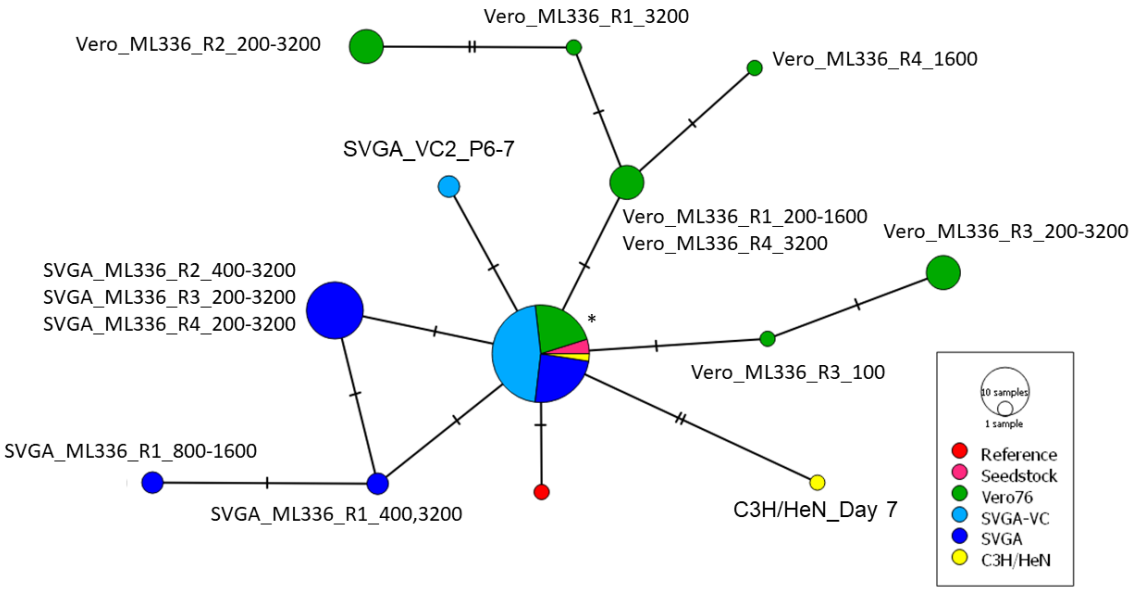


Figure 2-7. Whole-genome consensus sequence network shows that VEEV TC-83 evolves resistance through different trajectories in distinct microenvironments.

Consensus sequences were obtained from NGS sequencing data analysis and then aligned using MUSCLE and plotted based on sequence differences. Samples were grouped and labeled based on virus source. Each tick represents the number of nucleotide differences between sequences. The reference used was taken from GenBank (accession number L10443.1). Samples in the Seedstock, SVGA-VC, and C3H/HeN groups were not treated with ML336. In vitro samples are named by cell type, compound or control, and ML336 concentration or passage number. *, Samples include all seed stocks, C3H/HeN day 8, ML336 R1 50-100, ML336 R2 50, ML336 R3 50, ML336 R4 50-800, SVGA R1 50-200, SVGA R2 50-200, SVGA R3 50-100, SVGA R4 50-100, SVGA VC1 P1-7, SVGA VC2 P1-5, and SVGA VC3 1-7.

the evolution of drug resistance. We observed similarities between the ML336 resistance-conferring SNPs that emerged in each tested microenvironment and in the timing of emergence; however, the extent of virus genomes containing these SNPs as well as the timing varied between these groups. Moreover, the sites of random SNPs were distinct for each population. As far as we are aware, using network analysis to compare and map the trajectory of viral drug resistance from whole-genome sequencing has not been previously reported.

In previous studies, three mutations that conferred resistance to ML336 were discovered, namely, the Y102C and D116N mutations in the nsP2 gene and the Q210K mutation in the nsP4 gene, with the Q210K mutation causing the strongest loss of potency (5, 6). Because of those prior results, we had expected to find these mutations emerging in the TC-83 population when challenged with ML336 in Vero 76 cells. Notably, while the Q210K mutation was detected in R3, distinct mutations at that amino acid position were found in the other replicates, namely, Q210R and Q210H. In R1, multiple Q210 mutations appeared simultaneously. The changes at Q210 to basic amino acids suggests that the mutation from glutamine to a positively charged amino acid, whether it is lysine, histidine, or arginine, is particularly important in conferring resistance to ML336. The previously discovered Y102C and D116N nsP2 mutations were not found at any passage in any sample when passaging through Vero 76 cells, but when passaging through SVGA cells, the D116N mutation did appear for four passages in one replicate, starting at 100 nM, before disappearing from the population entirely. Interestingly, a mutation at E118V, also in nsP2, appeared in both experiments and followed a similar trajectory of peaking at 100 nM and eventually disappearing. The relative lack of prominent nsP2 SNPs past the first two passages may suggest that D116N and E118V are transitionary mutations that confer resistance but come with a higher fitness cost than the mutations at Q210 in nsP4.

When comparing the SNPs found between the Vero 76 and SVGA passaging experiments, we found that there was not a large difference in terms of the SNPs that arose to significant percentages, as mutations at A461 in nsP3, and Y87, Q210, and R305 appeared frequently throughout all samples, and few new SNPs in the nonstructural genes were found in the SVGA-passaged viruses. However, SNPs from SVGA-passaged TC-83 emerged more slowly than those from Vero-passaged TC-83; with the exception of Q210, most SNPs took three or four passages to reach a consensus in Vero 76 cells, while in SVGA, very few SNPs were able to reach a consensus, even Y87F and R305R. This result suggests that SVGA may provide a modest cellular immune response challenge to TC-83, but it is possible that in both systems the virus overwhelms the cells once the SNPs conferring ML336 resistance emerge and are selected. Remarkably, most of these mutations were not lost in the absence of ML336, which suggests that the fitness cost is not very high, at least for the Q210 mutations where they stayed at near-consensus level regardless of the identity of the other SNPs present. It was difficult to assess the SNPs in C3H/HeN mouse brains due to the low depth of coverage, particularly in genes where SNPs known to confer resistance to ML336 are found. However, the variety of SNPs that were detected between individual untreated mice underscores the genetic diversity of each viral population that is generated during infections. In addition, there was little to no

overlap between SNPs found in vitro and in vivo, again highlighting the differences in microenvironments.

For most of the passaging experiments, we did not confirm whether each major mutation occurred on the same individual viral genomes. However, for each replicate of the Vero 76-ML336 passaging experiment (**Figure 2-3**), we isolated RNA from the final passage, converted it to cDNA, and then ran a PCR using primers that covered the portion of the genome containing all of the major nsP3 and nsP4 mutations, ligated the products into a plasmid system, and ran plasmids through Sanger sequencing. For the most part, each individual sequence contained all of the SNPs that were located in the particular replicate tested (data not shown). This result seems to suggest that, while Q210 mutations may be necessary for resistance to ML336, other SNPs such as Y87F, are often found together on the same viral genomes and may play a role in sustaining viral fitness. Further experiments such as using mutagenesis to create mutant TC-83 clones containing these SNPs or sequencing samples using an Oxford Nanopore MinION sequencer would better determine the relationships between each individual SNP.

The data presented here align with data in our previous studies that suggest that ML336 and its analogs directly target the polymerase and disrupt interactions between nsP4 and nsP2 that, in turn, inhibit the synthesis of the positive and negative strands and the subgenomic RNA [106, 107]. However, it is still unclear what exact effect these mutations have on nsP4, as they all fall under a region where there is no known activity or function throughout all alphaviruses [56].

Since the discovery of ML336 and BDGR-4, we have continued to work on the development of new compounds. Lead candidates will be assessed for resistance in the same manner. Testing the trajectory of resistance in more complex systems, such as in primary neuronal cells, combination therapy, and multiple-cell systems resembling the central nervous system and blood-brain barrier, or in more mouse models, will allow us to examine the effects of more complex environments on antiviral resistance in VEEV in more detail.

CHAPTER 3. DIVERGENT EVOLUTIONARY TRAJECTORIES OF VEEV TC-83 OCCUR WHEN PASSAGED WITH ANTIVIRAL COMPOUNDS OF SIMILAR STRUCTURE

Introduction

Venezuelan equine encephalitis (VEE) disease in humans and equines is caused by VEE virus (VEEV), a member of the genus *Alphavirus* and family *Togaviridae*. VEE is characterized by a febrile illness with symptoms including headache, nausea, chills, and myalgia, but in some patients this can progress further to encephalitis, marked by symptoms including convulsions and behavioral change, as well as more severe symptoms such as stupor and coma, and possibly death [2, 9]. The recurrence of outbreaks of VEEV in its origins of South and Central America and its potential use as a bioweapon have led to its consideration as a public health and bioterrorism concern [1, 2]. Currently, there are no FDA-approved antiviral drugs or vaccines for human use against VEEV, which underscores the need for research to develop drugs and vaccines in order to be prepared against future outbreaks and epidemics.

In previous work, we discovered and characterized CID15997213, an antiviral compound capable of inhibiting the replication of VEEV that was discovered through high-throughput screening [106]. This work led to the discovery of a rearrangement of a benzamidine scaffold. These benzamidines include ML336, which has been characterized in its effect on viral replication and *in vivo* efficacy [107]. The compound structures of CID15997213 and ML336 are shown in **Figure 3-1**. Since then, many other compounds in the two chemotypes have been synthesized and tested for efficacy against VEEV, with the best chosen for *in vivo* efficacy and drug resistance studies.

In this study, four of the most promising compounds, BDGR-4, BDGR-49, BDGR-163, and BDGR-164 (**Figure 3-1**), were used to evaluate the emergence of resistance in VEEV strain TC-83 in repeated *in vitro* passaging experiments. We show that the virus populations evolve with various patterns in response to each compound and to the compound's concentration. We also show the convergence of specific mutations conferring resistance to each compound, particularly those in nsP2 and nsP4. These studies will help further understanding of viral population dynamics in response to selective pressures such as antivirals, as well as help inform and predict responses to *in vivo* treatments with these compounds.

Materials and Methods

Cells and Viral Culture

Vero 76 cells (ATCC CCL-131) were maintained at 37°C and 5% CO_2 in complete media (DMEM with hi-glucose and L-glutamine supplemented with 10% fetal

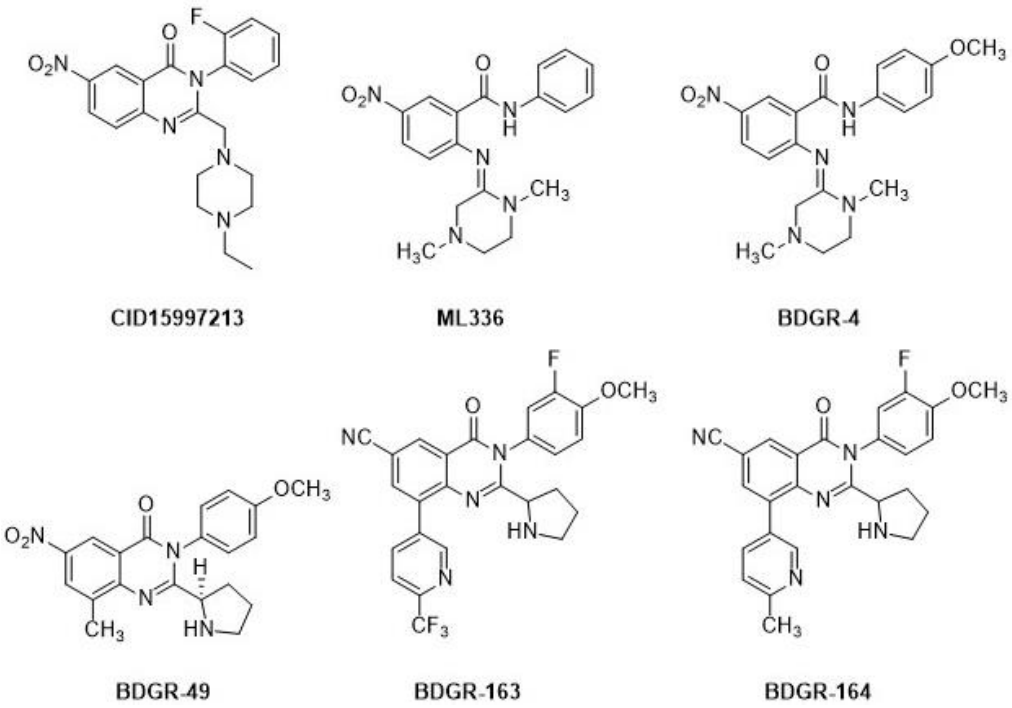


Figure 3-1. Compounds used in this study along with predecessor compounds CID15997213 and ML336.

bovine serum (FBS) and 1% penicillin-streptomycin). VEEV TC-83 (lyophilized vaccine) was obtained as a gift from Dr. Connie Schmaljohn (United States Army Medical Research Institute of Infectious Diseases, MD). Virus seed stocks were amplified in Vero 76 cells in infection media comprising of Minimum Essential Media with Earle's salts (MEM) with 2% FBS and 1% penicillin-streptomycin. All reagents and cell culture reagents were purchased from Thermo Fisher Scientific unless specified.

Viral Quantification

Plaque assays were used to measure viral titer. Twelve-well plates were seeded with Vero 76 cells and grown overnight. Ten series of 10-fold dilutions of virus samples were made in infection media, and 200 μL of each dilution was added to each well in duplicate and incubated at 37 °C, 5% CO_2 for 1 h. Following incubation, overlay media consisting of a 1:1 mix of 2XMEM supplemented with 5% fetal bovine serum (FBS), 1% L-glutamine, 2% penicillin-streptomycin and 2% carboxymethylcellulose was added to each well. Plates were incubated at 37 °C, 5% CO_2 for 48 h. Wells were fixed with 10% formalin, stained with crystal violet, washed with Dulbecco's phosphate buffered saline, and plaques were counted.

Blind Passage Evaluation for Selection of Resistant Mutants at Increasing Compound Concentrations

Prior to infection, cells were seeded onto 6-well plates and grown to 90% confluency. Cells were infected with VEEV TC-83 at an initial MOI of 1 or 0.1 as specified in text and figure legends. Resistant viruses were selected for by passaging wild-type (WT) VEEV TC-83 in two-fold increasing concentrations of BDGR-4, BDGR-49, BDGR-163, or BDGR-164, starting at 50 nM, and increasing to 100, 200, 400, 800, 1600, and lastly 3200 nM, for a total of 7 concentrations. We used 250 μL of supernatant from each passage to initiate the subsequent passage. Three distinct replicates were passaged in each experiment. VEEV TC-83 was passaged without compound in triplicate as a control.

Blind Passage Evaluation of Resistant Mutants at Constant Concentrations

BDGR-49 was repeatedly passaged at either 20 nM or 200 nM, starting at an MOI of 0.1, using 6-well plates with Vero 76 cells. 250 μL of supernatant from each passage was transferred to another seeded 6-well plate initiate the subsequent passage.

Whole Genome Sequencing

To sequence virus genomes, 0.5 μL of supernatant from each passage was harvested and twice-clarified by centrifugation. RNA was isolated using Trizol LS

reagent using the manufacturer's protocol (Invitrogen). cDNA was synthesized from purified RNA using random hexamer primers and Superscript IV (Invitrogen). Each cDNA sample was amplified through 7 or 30 rounds of PCR, using the Phusion High-Fidelity Kit (Thermo Scientific) and a set of 25 primer pairs (Chapter 2, **Table 2-1**), making PCR fragments either 500 bp or 1 kb in length using a tiling approach that doubly covered the whole genome as described. PCR products were purified using the Wizard SV Gel and PCR Clean-Up System (Promega). Libraries for NGS were prepared using the Nextera XT kit (Illumina), then sequenced on the Illumina MiSeq. Data was analyzed using CLC Genomics Workbench v. 12 (Qiagen). To process the sequencing data, reads less than 50 nucleotides in length and with a Phred score < 20 were removed. The remaining reads were mapped to the seed stock consensus sequence, and variants were identified either with a minimum frequency above 1% in all reads at positions with a minimum coverage of 400, or, if the quality and quantity of reads were sufficient, with a minimum frequency above 0.5% in all reads at positions with minimum coverage of 1000.

Network Analysis

Consensus sequences from each sample were obtained using CLC Genomics Workbench v.12 (Qiagen). Sequences were aligned using MUSCLE and mapped by the program PopArt v. 1.7 [133], utilizing a Minimum Spanning Tree to plot each sequence.

Results

Emergence of VEEV TC-83 resistance to BDGR-4 required concentrations 6.7-fold over the EC_{50}

BDGR-4 is a benzamidine derivative of ML336. Given the similarities in structure between the two molecules, it was expected that the same nsP2 and nsP4 SNPs that conferred resistance to ML336 would appear when VEEV TC-83 was passaged in the presence of BDGR-4. Passaging started at 50 nM, a concentration slightly above the EC_{50} (30 nM); we increased the concentration of ML336 two-fold at each passage, i.e. 100, 200, 400, 800, 1600, and 3200 nM concentrations. RNA was isolated from twice-clarified supernatant at each passage and sequenced each sample as described in the Materials and Methods.

For each sample, the SNPs that appeared in at least 10% of reads for a particular nucleotide at each concentration of BDGR-4 were evaluated. The summary of these SNPs are shown in **Figure 3-2** and **Table 3-1**. Remarkably, no SNPs emerged in any replicate until the third passage at the 200 nM concentration. By the last passage at 3200 nM BDGR-4, mutations at genome positions 5414, 5962, 6331, and 6617 reached a majority in the population in all three replicates. Of these, the A5962T mutation resulted

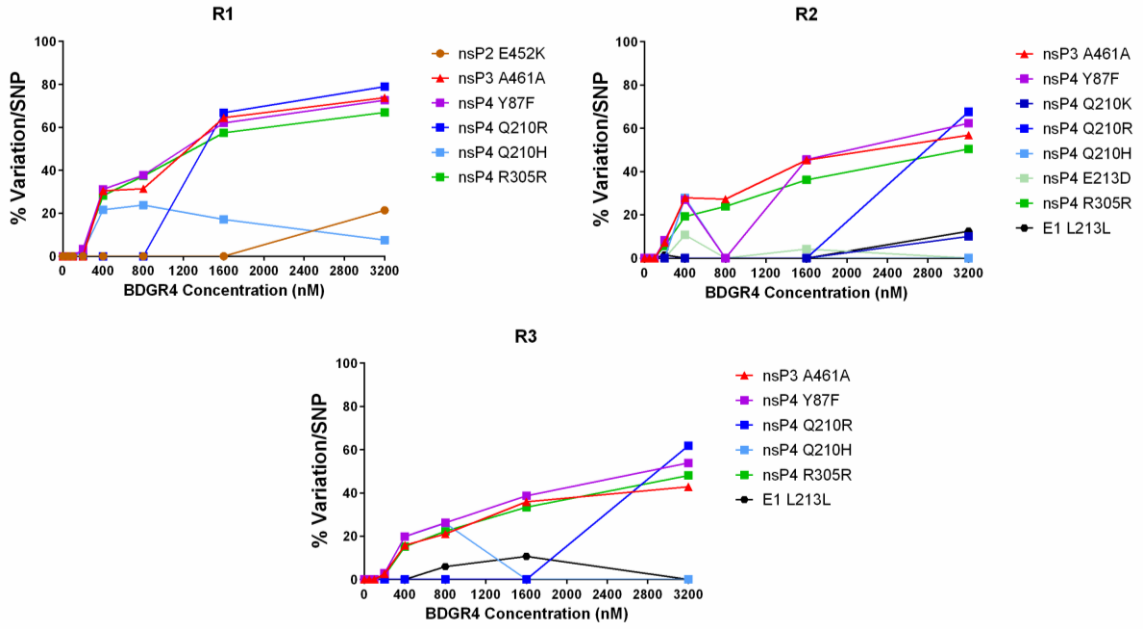


Figure 3-2. Resistance to BDGR-4 only arises after multiple passages.

Each well of a 6-well plate was seeded with approximately 600,000 cells and infected with an MOI of 1, and then passaged in the presence of two-fold increasing concentrations of BDGR-4: 50-, 100-, 200-, 400-, 800-, 1600-, and 3200-nM. All seven passages are plotted. Each SNP in the graphs appeared in at least 10% of all reads in at least one passage and are represented by an individual color. Reprinted from final submission with open access permission. Jonsson, C.B., Cao, X.,Lee, J.Gabbard, J., D.Chu, Y. K., Fitzpatrick, E. A., Julander, J., Chung, D. H., Stabenow, J., Golden, J. E. Efficacy of a ML336 derivative against Venezuelan and eastern equine encephalitis viruses. Antiviral Res, 2019. 167: p. 25-34.DOI: https://doi.org/10.1016/j.antiviral.2019.04.004 [117].

Table 3-1. Summary of point mutations and amino acid changes in TC-83 at increasing concentrations of BDGR-4.

Nucleotide Change	Amino Acid Change	Gene	Replicate	50 nM	100 nM	200 nM	400 nM	800 nM	1600 nM	3200 nM
G3033A	**E452K**	nsP2	R1	-	-	-	-	-	-	21.4
G5414A	**A461A**	nsP3	R1	-	-	-	30.51	31.43	64.53	73.86
			R2	-	-	7.15	27.91	27.26	45.33	56.89
			R3	-	-	2.63	15.85	21.07	35.91	42.91
A5962T	**Y87F**	nsP4	R1	-	-	3.34	31.17	37.73	62.04	72.66
			R2	-	-	8.37	26.95	-	45.58	62.36
			R3	-	-	2.95	19.91	26.23	38.7	53.9
C6330A	**Q210K**	nsP4	R2	-	-	-	-	-	-	10.07
A6331G	**Q210R**	nsP4	R1	-	-	-	-	-	66.8	79.01
			R2	-	-	-	-	-	-	67.65
			R3	-	-	-	-	-	-	61.86
A6332T	**Q210H**	nsP4	R1	-	-	1.96	21.62	23.81	17.14	7.58
			R2	-	-	-	27.89	-	-	-
			R3	-	-	2.77	19.8	26.01	-	-
A6341T	**E213D**	nsP4	R2	-	-	-	10.74	-	4.21	-
			R3	-	-	0.65	4.34	4.06	1.67	-
G6617A	**R305R**	nsP4	R1	-	-	2.52	28.32	37.32	57.49	66.97
			R2	-	-	5.66	19.26	23.88	36.21	50.47
			R3	-	-	1.87	15.21	22.36	33.4	48.06
G10639A	**L213L**	E1	R2	-	-	1.53	-	-	-	12.43
			R3	-	-	-	-	5.97	10.71	-

Table 3-1. Continued.

Reprinted from final submission with open access permission. Jonsson, C.B., Cao, X.,Lee, J.Gabbard, J., D.Chu, Y. K., Fitzpatrick, E. A., Julander, J., Chung, D. H., Stabenow, J., Golden, J. E. Efficacy of a ML336 derivative against Venezuelan and eastern equine encephalitis viruses. Antiviral Res, 2019. 167: p. 25-34.DOI: https://doi.org/10.1016/j.antiviral.2019.04.004 [117].

in an amino acid change, a Y87F mutation in the nsP4 gene. The G5414A and G6617A mutations left the amino acids A461 of nsP3 and R305 of nsP4 respectively unchanged. Interestingly, all three replicates had multiple SNPs causing basic amino acid changes at nsP4 Q210 –Q210K, Q210R, and Q210H. All three mutations were caused by nucleotide substitutions at different positions on the codon. However, their prevalence in the virus populations differed; Q210H appeared first in all three replicates after three or four passages but did not penetrate past 28% before eventually disappearing entirely from two replicates. Q210R only appeared at 1600 nM, but penetrated at 61-79% in all three replicates when it did so. Q210K only appeared at approximately 10% in one replicate. Overall, SNPs emerged in a similar manner across all three replicates, with none emerging until the third passage and the major SNPs only becoming the majority in the final two passages.*

VEEV TC-83 Evolves Resistance to BDGR-49 Rapidly

Unlike BDGR-4, BDGR-49 is a quinazolinone compound, and thus we hypothesized that, like the hit compound CID15997213, the SNP profile would feature mutations in nsP2. Like before, VEEV TC-83 was passaged in increasing concentrations of BDGR-49, starting at 50 nM, and each passage was sequenced. Results from the three experimental replicates are shown in **Figure 3-3** and **Table 3-2**. Unexpectedly, in all three replicates SNPs appeared at high percentages at 50nM; unlike compound CID15997213, high percentage SNPs were not found in nsP2, but instead were four identified from ML336 and BDGR4 experiments, namely the G5414A, A5962T, A6331G, and G6617A mutations. Like before, only A5962T and A6331G mutations caused amino acid changes, of Y87F and Q210R respectively. All four mutations reached a majority in the population by the second passage, then for the most part slowly increased in percentage through each successive passage. The Q210R mutation reached a near consensus by the final passage in all three replicates.

To further test BDGR-49 resistance, VEEV TC-83 was passaged with the compound at set concentrations of 20 and 200 nM seven times. Two biological replicates were tested at each BDGR-49 concentration. Like before, the virus population was sequenced at each passage, with the results shown in **Figure 3-4**. As expected, SNPs such as Q210R and Y87F appeared in all four samples, although they were initially much less penetrant in the 20 nM replicates than in the 200 nM replicates. Interestingly, in these samples more SNPs in total emerged, including a few from nsP2. These results showed that VEEVTC-83 populations obtained BDGR-49 resistance mutations very quickly, but with a wider range of SNPs that emerged.

---

*Paragraph reprinted from final submission with open access permission. Jonsson, C.B., Cao, X.,Lee, J.Gabbard, J., D.Chu, Y. K., Fitzpatrick, E. A., Julander, J., Chung, D. H., Stabenow, J., Golden, J. E. Efficacy of a ML336 derivative against Venezuelan and eastern equine encephalitis viruses. Antiviral Res, 2019. 167: p. 25-34.DOI: https://doi.org/10.1016/j.antiviral.2019.04.004 [117].

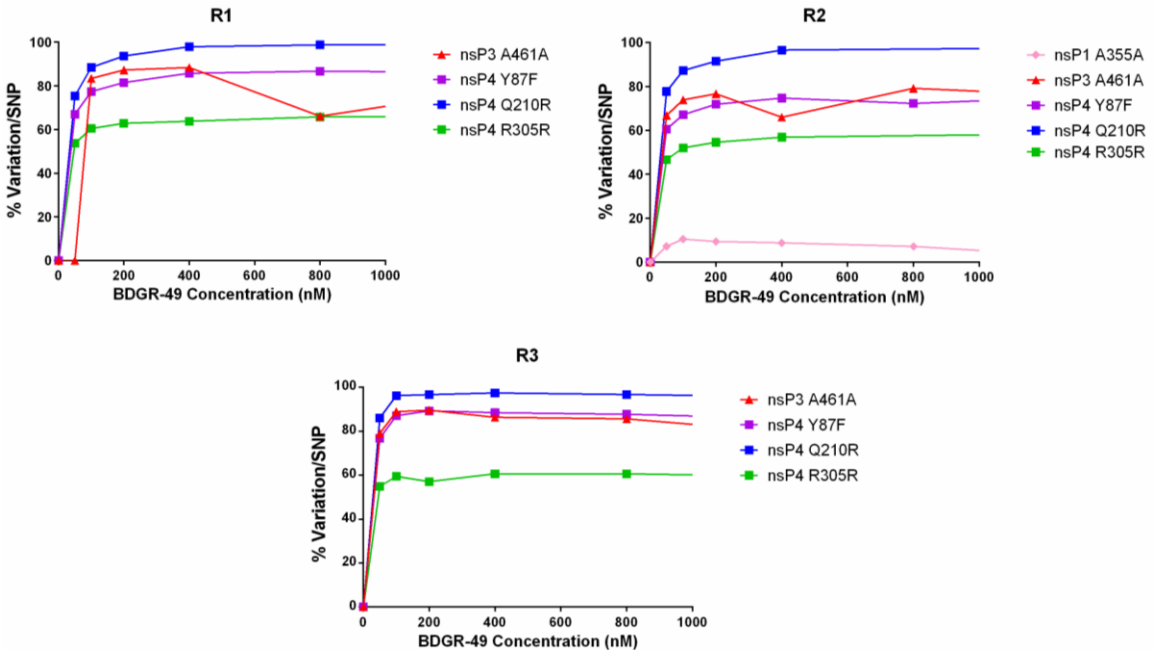


Figure 3-3. SNPs conferring resistance to BDGR-49 emerge after a single passage.

Each well of a 6-well plate was seeded with approximately 600,000 cells and infected with an MOI of 1, then passaged in the presence of two-fold increasing concetrations of BDGR-49: 50-, 100-, 200-, 400-, 800-, 1600-, and 3200-nM. Only the first five passages are plotted here. Each SNP in the graphs appeared in at least 10% of all reads in at least one passage and are represented by an individual color.

Table 3-2. Summary of point mutations and amino acid changes made at various concentrations of BDGR-49.

Nucleotide Change	Amino Acid Change	Gene	Replicate	50 nM	100 nM	200 nM	400 nM	800 nM	1600 nM	3200 nM
G1100A	**E452K**	nsP1	R2	7.21	10.49	9.45	8.81	7.16	-	-
G5414A	**A461A**	nsP3	R1	-	83.41	87.33	88.41	66.15	84.22	75.92
			R2	66.83	73.9	76.66	66	79.18	74.07	73.06
			R3	79.02	88.86	89.58	86.33	85.59	75.8	73.83
A5962T	**Y87F**	nsP4	R1	66.98	77.44	81.57	85.81	86.74	86.27	84.32
			R2	60.53	67.16	71.88	74.71	72.26	77.12	76.98
			R3	76.76	87.08	89.23	88.43	87.67	84.91	83.54
A6331G	**Q210R**	nsP4	R1	75.4	88.53	93.68	97.98	98.78	99.09	98.49
			R2	77.72	87.26	91.53	96.58	>99.9	98.02	98.37
			R3	86.02	96.24	96.64	97.38	96.72	95.06	93.08
G6617A	**R305R**	nsP4	R1	53.69	60.5	62.85	63.79	65.8	66.62	68.71
			R2	46.68	52.05	54.6	56.92		59.06	61.65
			R3	54.84	59.42	57.04	60.63	60.61	58.87	59.44

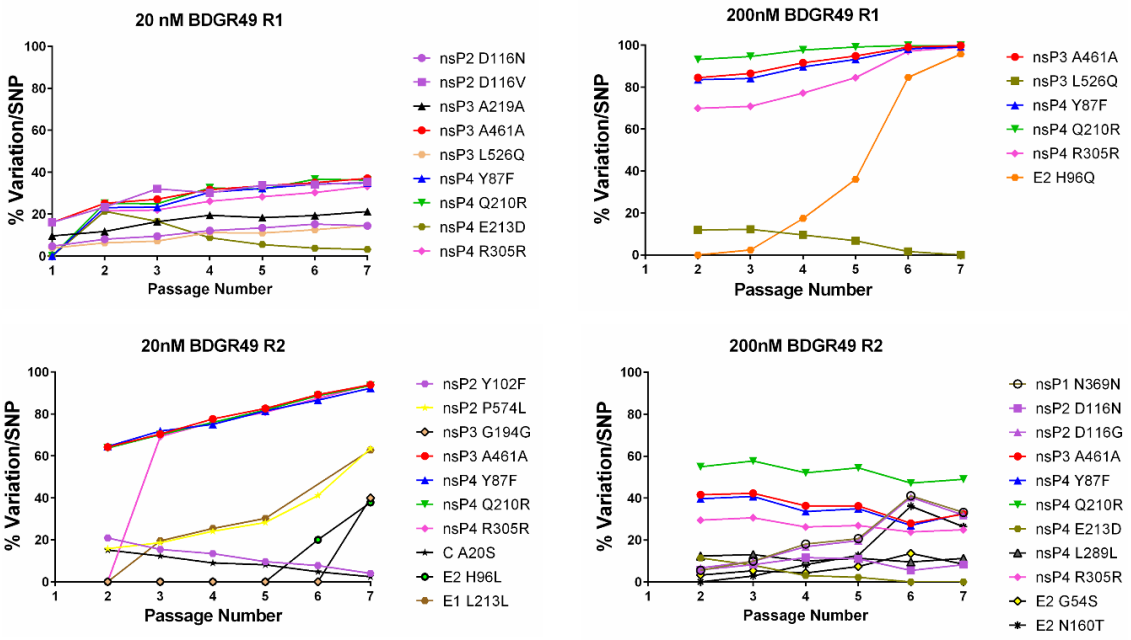


Figure 3-4. SNPs conferring resistance to BDGR-49 rapidly emerge even at low concentrations.

VEEV TC-83 was passaged in Vero 76 cells at either 20 nM or 200 nM of BDGR-49 seven times in succession, and each passage was sequenced for SNPs. Each SNP in the graphs appeared in at least 10% of all reads in at least one passage, and are represented by an individual color.

BDGR-163 and BDGR-164 Showed Divergent Evolutionary Trajectories Despite Very Similar Structures

BDGR-163 and BDGR-164 are very similar in structure, with the only difference being a trifluoromethyl group in BDGR-163 replacing the methyl group in BDGR-164. Despite this difference, it was hypothesized that this change in an R-group would lead to similar VEEV TC-83 was passaged seven times in doubling concentrations (50 to 3200 nM) of BDGR-163 and BDGR-164 and sequenced at each passage. The resulting SNPs are found in **Figure 3-5** and **Table 3-3**. VEEV TC-83 passaged in BDGR-163 showed starkly different trajectories compared to those of other compounds; namely, mutations in nsP2 dominated the early passages, while nsP4 Q210 mutations did not appear until 200-400 nM. In all three replicates, an A2000T mutation led to the amino acid change E117D in nsP2. An adjacent amino acid also changed in replicates R1 and R3 in the D116A mutation. In replicates R2 and R3, a Y102H mutation also emerged after a few passages. Of other particular note was the E213D amino acid change in nsP4, which either slowly increased in low percentage over time in R1, became the most dominant mutation at near 80% in R2, or weakly penetrated and disappeared entirely in R3. Taken together, BDGR-163 appears to cause TC-83 to evolve nsP2 mutations in response but allows for room for Q210 mutations to eventually emerge.

TC-83 passaged in BDGR-164 showed evolutionary trajectories in which mutations emerged rapidly (**Figure 3-6** and **Table 3-4**). In all three replicates, the G5414A, A5962T, A6331G, and G6617A mutations all appeared at 50 nM and, except for G6617A in replicates R2 and R3, all were above 50% penetrance. In addition, the Q210R mutation took as little as one passage to become near-consensus in the population, even starting at over 98% in R2. Overall, this data suggests that TC-83 rapidly evolved resistance-conferring mutations in order to overcome BDGR-164.

Network Analysis Shows Divergence in Evolutionary Trajectories

To examine the differences in evolutionary trajectory of VEEV TC-83 for the compounds tested, a network analysis was run using consensus sequences of each replicate at each passage, as well as the consensus sequence of the seed stock used to begin the passaging and a reference sequence from GenBank [11]. With the exception of replicate 2 at 800 nM of BDGR-4, which had low sequence quality, these consensus sequences are plotted in **Figure 3-7**. This graph shows that there is divergence of evolutionary path based on compound; BDGR-4-treated viruses took many passages to differentiate from the seed stock, while BDGR-49 and BDGR-164-treated VEEV TC-83 immediately diverged from the reference sequence, and BDGR-163-treated virus diverged in a unique direction, with multiple branches. Interestingly, the largest group of sequences was at a position four nucleotides away from the seed stock; this group was comprised of viruses treated with BDGR-4, 49, and 164, with the wide majority of BDGR-49-derived virus landing there. In summary, each compound drove diverging evolutionary trajectories in VEEV TC-83 in both nucleotide differences and the explored sequence space.

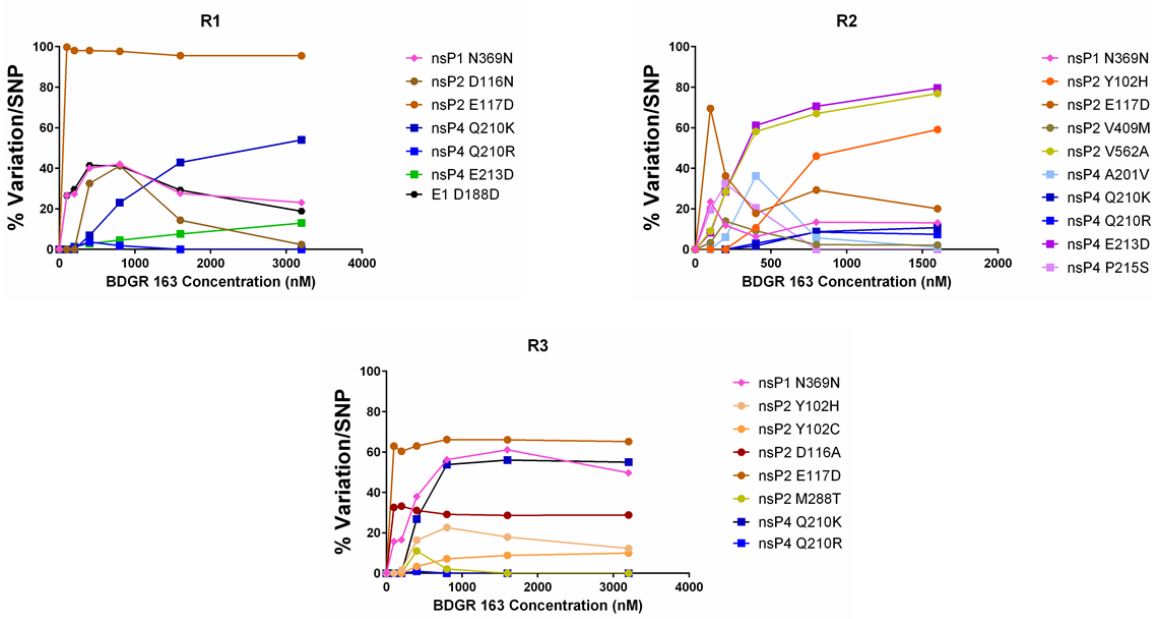


Figure 3-5. SNPs in nsP2 conferring resistance to BDGR-163 predominately appeared after repeated passaging.

Each well of a 6-well plate was seeded with approximately 600,000 cells and infected with an MOI of 0.1, then passaged in the presence of two-fold increasing concentrations of BDGR-163: 50-, 100-, 200-, 400-, 800-, 1600-, and 3200-nM. All seven passages are plotted. Each SNP in the graphs appeared in at least 10% of all reads in at least one passage and are represented by an individual color.

Table 3-3. Summary of point mutations and amino acid changes made at various concentrations of BDGR-163.

Nucleotide Change	Amino Acid Change	Gene	Replicate	50 nM	100 nM	200 nM	400 nM	800 nM	1600 nM	3200 nM
C1151T	**N369N**	nsP1	R1	-	26.52	27.33	40.09	42.09	27.64	23.06
			R2	-	23.51	11.82	6.32	13.41	13.1	
			R3	-	15.69	16.49	37.96	56.24	61.12	49.7
T1953C	**Y102H**	nsP2	R2	-	-	-	10.61	46.05	59.12	
			R3	-	-	1.51	16.39	22.62	17.98	12.29
A1954G	**Y102C**	nsP2	R3	-	-	-	3.33	7.16	8.87	10
G1995A	**D116N**	nsP2	R1	-	-	-	32.54	41.28	14.41	2.38
A1996C	**D116A**	nsP2	R3	-	32.61	33.21	31.11	29.17	28.65	28.87
A2000T	**E117D**	nsP2	R1	-	99.78	98.06	98.08	97.7	95.59	95.55
			R2	-	69.52	36.25	17.81	29.29	20.1	
			R3	-	32.61	33.21	31.11	29.17	28.65	28.87
T2512C	**M288T**	nsP2	R3	-	-	-	11.03	2.13	-	-
G2874A	**V409M**	nsP2	R2	-	3.26	13.92	9.16	2.37	2.1	
T3334C	**V562A**	nsP2	R2	-	9.03	28.48	58.09	67.04	76.88	
C6304T	**A201V**	nsP4	R2	-	-	5.99	36.24	5.76	1.12	
C6330A	**Q210K**	nsP4	R1	-	-	-	6.86	23.09	42.9	53.95
			R2	-	-	-	1.73	8.74	10.72	
			R3	-	-	-	26.83	53.77	56.04	55
A6331G	**Q210R**	nsP4	R1	-	-	1.11	3.67	1.82	-	-
			R2	-	-	-	2.97	8.54	7.43	
			R3	-	-	-	1.03	-	-	-

Table 3-3. Continued.

Nucleotide Change	Amino Acid Change	Gene	Replicate	50 nM	100 nM	200 nM	400 nM	800 nM	1600 nM	3200 nM
A6341C	**E213D**	nsP4	R1	-	-	1.34	2.92	4.55	7.62	12.97
			R2	-	8.32	28.5	61.14	70.57	79.61	
			R3	-	-	2.34	2.78	1.68	-	-
C6345T	**P215S**	nsP4	R2	-	19.55	32.57	20.4	-	-	
T10564C	**D188D**	E1	R1	-	26.53	29.68	41.42	41.2	29.2	18.83

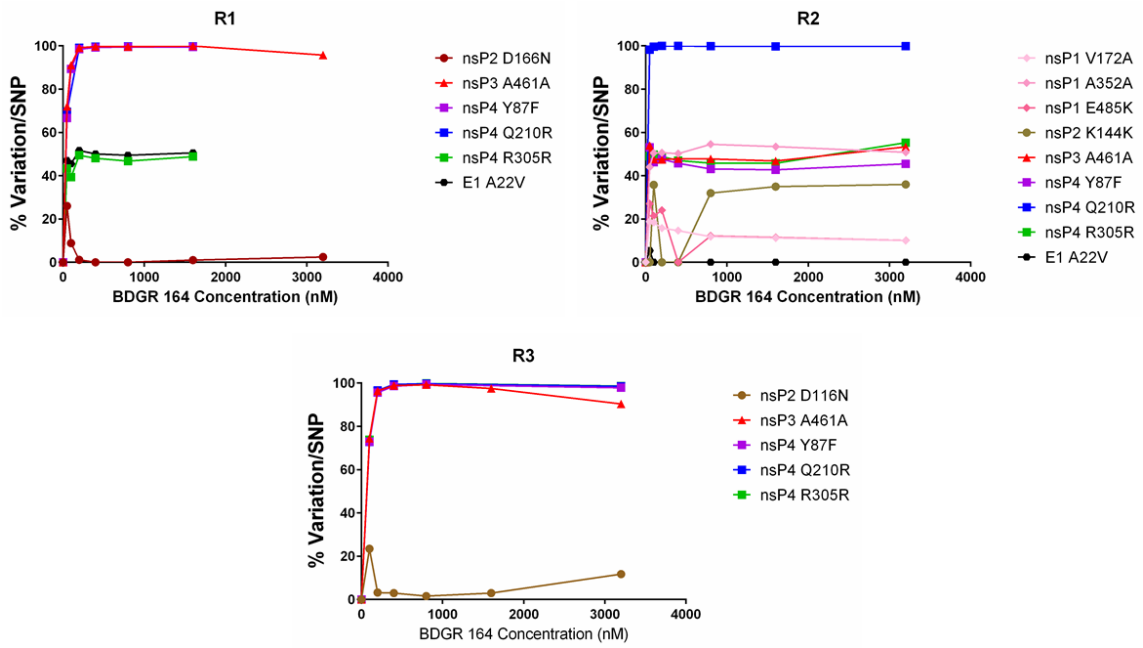


Figure 3-6. SNPs conferring resistance to BDGR-164 appear after one passage.

Each well of a 6-well plate was seeded with approximately 600,000 cells and infected with an MOI of 1, then passaged in the presence of two-fold concentrations of BDGR-164: 50-, 100-, 200-, 400-, 800-, 1600-, and 3200-nM concentrations. All seven passages are plotted. Each SNP in the graphs appeared in at least 10% of all reads in at least one passage and are represented by an individual color.

Table 3-4. Summary of point mutations and amino acid changes made at various concentrations of BDGR-164.

Nucleotide Change	Amino Acid Change	Gene	Replicate	50 nM	100 nM	200 nM	400 nM	800 nM	1600 nM	3200 nM
T559C	**V172A**	nsP1	R2	18.83	18.38	15.81	14.61	11.86	11.22	10.03
G1100A	**A352A**	nsP1	R2	43.97	50.5	50.65	50.19	54.53	53.5	50.85
G1497A	**E485K**	nsP1	R2	27.11	21.45	24.04	-	12.17	11.47	10.15
G1995A	**D116N**	nsP2	R1	26.03	8.89	1.17	-	-	1.01	2.49
			R3		23.45	3.24	2.96	1.57	2.96	11.74
A2081G	**K144K**	nsP2	R2	-	35.81	-	-	32.01	35.01	36.02
G5414A	**A461A**	nsP3	R1	72.11	91.09	99.18	99.74	99.68	99.94	95.76
			R2	54	47.65	47.5	47.82	47.75	46.86	53.38
			R3		74.43	96.42	98.93	99.22	97.52	90.36
A5962T	**Y87F**	nsP4	R1	66.67	89.36	98.47	99.27	99.51	99.46	
			R2	53.11	46.22	47.7	45.8	43.11	42.75	45.54
			R3		72.94	95.65	98.59	99.36		97.86
A6331G	**Q210R**	nsP4	R1	69.54		99.1	99.65	99.77	99.78	
			R2	98.19	99.61	100	99.95	99.84	99.79	99.82
			R3		72.76	96.57	99.43	99.71		98.51
G6617A	**R305R**	nsP4	R1	43.3	39.31	49.46	48.08	46.78	48.81	
			R2	48.95	49.75	48.9	47.04	45.81	45.74	55.31
			R3		73.82	96.44	99.22	99.86		98.63
C10065T	**A22V**	E1	R1	47.05	45.58	51.69	50.09	49.47	50.6	
			R2	5.44	-	-	-	-	-	-

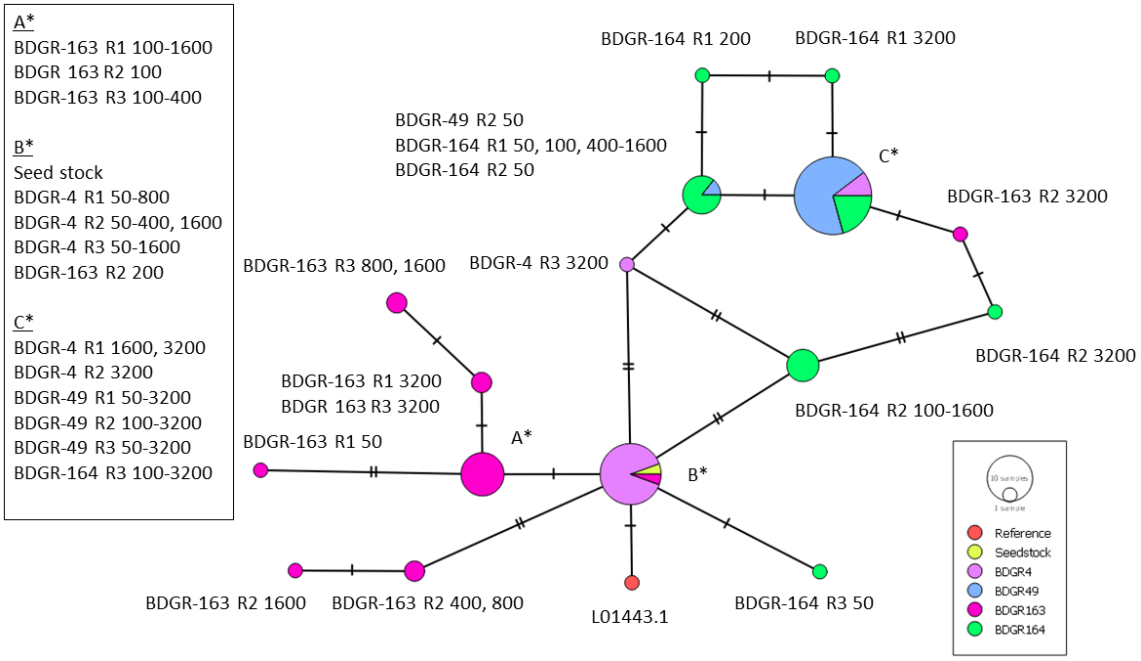


Figure 3-7. Network analysis of VEEV TC-83 consensus sequences shows that the virus evolves along divergent paths for each compound.

Consensus sequences were obtained from NGS sequencing data analysis, aligned using MUSCLE, and plotted based on nucleotide differences. Samples were grouped and labeled based on source and compound treatment. Each tick represents the number of nucleotide differences between sequences. The reference sequence was taken from GenBank (accession number L10433.1). Other than the reference and seed stock sequences, all sequences are labeled by compound treatment, replicate number, and compound concentration (nM).

Conclusions

In this study, VEEV TC-83 was repeatedly passaged in increasing concentrations of four different compounds from two closely related scaffolds. It was hypothesized that the difference in which scaffold a particular compound belonged to, particularly benzamidine or quinazolinone, would affect the evolutionary trajectory of TC-83 in terms of both the SNPs that emerged over time and the percentage they appeared in the population. In addition, it was thought that smaller changes, such as between different R groups, would also affect the evolutionary trajectory. The results shown here demonstrate that both hypotheses may be generally correct, but there may be more factors at work than simply whether a compound is classified as a benzamidine or quinazolinone, or simply because one functional group is different between compounds. Of the four compounds tested here, only BDGR-49 caused all experimental replicates to evolve similar SNP profiles and follow similar evolutionary trajectories.

It was surprising to see that the single difference of a methyl group versus a trifluoromethyl group caused a significant difference in both the SNPs that initially appeared and the trajectories of those SNPs. BDGR-163, which had the trifluoromethyl side group, was the only compound to induce the emergence of nsP2 mutations to a significant degree, particularly the D116N mutation. Mutations at Y102 emerged in virus treated with BDGR-163, which was particularly notable in that this mutation had only previously appeared in mutational studies with CID159977213, where a Y102C (as well as D116N) mutation was found [106]. In contrast, the quinazolinone compounds BDGR-4 and BDGR-164 followed similar trajectories in that the nsP3 A451A, and nsP4 Y87F, Q210R, and R305R mutations all emerged within a single passage, while these mutations were largely absent with BDGR-163. While D116N appeared in two out of three replicates in BDGR-164, it followed a trend of appearing at over 20% penetrance, then nearly disappearing. Combined with the non-immediate rise of Q210 mutations with BDGR-163, it may be possible that VEEV TC-83 first explores nsP2 mutations on the way to finding the nsP4 mutations. The Q210K mutation caused a higher fold loss in potency than the Y102C and D116N mutations with ML336, which makes it possible that it is the more optimal solution for VEEV to find [107].

These experiments show how virus populations can evolve even in response to a single selective pressure. However, observation of these population dynamics in more complex environments such as multiple-cell systems and *in vivo* could allow for more accurate predictions of how VEEV responds to drug treatment in terms of the sequence spaces the population is able to explore.

CHAPTER 4. EVOLUTION OF VENEZUELAN EQUINE ENCEPHALITIS VIRUS IN THE BRAINS OF MICE

Introduction

Venezuelan equine encephalitis virus (VEEV) is a New World alphavirus that has caused numerous outbreaks of VEE in humans and equines in the past century [5]. In humans, VEE is characterized by a febrile illness, symptoms of which include fever, malaise, myalgia, and headache. In some cases, particularly in children, VEE can progress to neurological disease, with symptoms including convulsions, seizures, stupor, and coma, and the possibility of death [9]. Outbreaks of VEE occur when infected mosquitos transmit virus from small mammals, which are believed to be the natural reservoir of VEEV, to equines, which act as an intermediate, amplifying host for the virus. Mosquitos can then transmit VEEV from equines to cause human outbreaks. VEEV can also be spread through aerosols, which was why it was developed as a bioweapon [2, 134]. Currently, there are currently no FDA-approved vaccines or antivirals against VEEV.

Knowledge of the pathogenesis of VEEV in humans is limited due to the lack of autopsies from past outbreaks. It is known, however, that lesions occur in the CNS, lymph nodes, spleen, liver, and lungs, lymphoid tissue of the gastrointestinal tract, and kidneys [23, 24]. Lesions in the brain include edema, meningitis, and congestion, while inflammatory infiltration is also observed. Follicular necrosis and lymphocyte degeneration also occurs in lymph nodes and the spleen.

Mice are commonly used as animal models for VEEV as they provide the ability to examine the neuroinvasion of the virus. During a natural or subcutaneous (s.c.) mouse infection, VEEV is taken up by dendritic cells and brought to lymphoid tissue, where it establishes infection [26, 27]. VEEV enters the bloodstream following infection of the draining lymph node and spreads to several organs in the body, including the spleen and brain. Neurological disease occurs when VEEV penetrates the central nervous system (CNS) by accessing olfactory neurons in the olfactory tract, which are easily accessible from either blood vessels or from nerve endings of olfactory receptor neurons that cross the nasal epithelium [27, 28]. From these neurons, VEEV quickly reaches the olfactory bulbs of the brain. From there, virus spreads through the rest of the brain, including the cerebellum and brain stem. In aerosol or intranasal (i.n.) infections, VEEV reaches the olfactory nerves and the brain much more rapidly [29]. Tissue damage in mice occurs in the form of apoptosis and necrosis of infected neurons, perivascular cuffing, and meningitis in areas of the brain the virus has reached [28]. In an outbred mouse line, ICR (CD-1®), with an i.n. infection of VEEV strain TC-83, virus can be detected in the brain at least three days before clinical symptoms develop [36]. A different study used mice lacking interferon-induced protein with tetratricopeptide repeats 1 (*Ifit1*), which inhibits the translation of viral RNA, and found that VEEV TC-83 reaches the cerebellum and spinal cord by 4 days post –infection (dpi) [34]. It is possible that VEEV enters the CNS through other means, such as by crossing the blood-brain barrier (BBB), but this is

currently a subject of debate since the BBB is broken down during infection and may not act as a barrier [34].

While the exact routes VEEV can use to enter the CNS are still being examined, our understanding of the viral population during the progression of infection within the brain is limited. VEEV, like other viruses with an RNA-based genome, has high mutation rates as its RNA-dependent RNA polymerase (RdRp) has a high error rate [94, 122]. Hence, replication to high titers generates a population of many individually distinct viruses that contribute to an overall phenotype. A high mutation rate may be required for its ability to infect both mammals and mosquitos. For example, in the mosquito, the VEEV population undergoes a bottleneck, resulting in a small number of infected midgut cells, even though this is an early step in establishing an infection [3]. This would suggest that the VEEV population needs to maintain a diverse population to ensure continued infection as it moves through the midgut, hemocoel, and salivary glands. Therefore, the mutation rate may be necessary for VEEV's ability to adapt to infect multiple hosts, as demonstrated in studies where altering the fidelity of the polymerase changes the infectivity of VEEV and increases survival of the host [103]. To our knowledge, deep sequencing to characterize the standing genetic variability or variants of VEEV in an insect or mammalian host has not been examined.

Based on the findings that VEEV undergoes a hard selection during infection of the mosquito midgut, we hypothesized that similar barriers may be present during the i.n. infection of VEEV TC83 in the C3H/HeN mouse as the virus moves from the olfactory bulb to the brain. The C3H/HeN strain has been developed as a lethal model for VEEV. TC-83 is an attenuated vaccine strain of VEEV that was derived from repeated passaging of TrD and has 12 nucleotide differences [11, 135]. C3H/HeN is the only known mouse strain in which TC-83 causes a lethal infection when infected i.n., intracranially, or through aerosol, whereas other strains such as BALB/c will experience symptoms but do not succumb [28, 136]. C3H/HeN has also been used as an animal model for antiviral efficacy studies [106, 109, 132].

To define the distribution and progression of VEEV infection from the olfactory bulb to the brain over time, we conducted a serial sacrifice study in which the olfactory bulbs and brains were taken each day for nine days post-TC-83 infection to investigate viral antigen and to investigate the population dynamics of VEEV, we examined the viral populations in the olfactory bulb, piriform cortex, caudate putamen (CPU), motor cortex, sensory cortex, hippocampus, thalamus, and cerebellum. Here, we show the extent of VEEV infection over time, starting from olfactory epithelium and olfactory bulb and ending with the cerebellum. We also show, using network analysis of TC-83 sequence reads and evaluation of single nucleotide polymorphisms (SNPs), that the VEEV genomes does not undergo any apparent selection resulting in mutations through Day 5, suggesting that during this time frame, there were no apparent barriers to viral replication using this route of infection.

Materials and Methods

Cells and Viral Culture

Vero 76 cells (ATCC CCL-131) were maintained at 37°C and 5% CO_2 in complete medium containing Dulbecco modified Eagle medium (DMEM) with Hi-glucose and L-glutamine supplemented with 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin. VEEV TC-83 (lyophilized vaccine) was obtained as a gift from Connie Schmaljohn (United States Army Medical Research Institute of Infectious Diseases, Maryland). Virus seed stocks were amplified in Vero 76 cells in infection medium comprising minimum essential medium with Earle's salts (MEM) with 2% FBS and 1% penicillin-streptomycin. All cell culture reagents were purchased from Thermo Fisher Scientific unless otherwise specified.

Plaque Assays

Twelve-well plates were seeded with 2.5×10^5 Vero 76 cells overnight. Ten series of 10-fold dilutions of virus samples were made in infection medium, and 200 l of each dilution was added to each well in duplicate and incubated at 37°C and 5% CO_2 for 1 h. Following incubation, overlay medium consisting of a 1:1 mix of 2% carboxymethylcellulose and 2X MEM supplemented with 5% FBS, 1% L-glutamine, and 2% penicillin-streptomycin was added to each well. Plates were incubated at 37°C and 5% CO_2 for 48 h. Wells were fixed with 10% formalin, stained with crystal violet, and washed with Dulbecco's phosphate-buffered saline (DPBS), and plaques were counted.

Mouse Studies for Pathology

All mouse studies were reviewed by and approved by the UTHSC IACUC in animal care and use protocol numbers 18-044 and 21-0258. Five- to six-week-old, female, C3H/HeN mice (Charles River Laboratories) were randomly assigned to one of ten groups based on day of sacrifice: 1, 2, 3, 4, 5, 6, 7, 8, or 9 days post infection (dpi). Mice were infected i.n. with 1×10^7 PFU of VEEV TC-83, 15 µl per naris in PBS. PBS was used in place of virus for the vehicle-only control group. Mice were weighed daily and checked twice daily for mortality and morbidity. Each group of VEEV TC-83 infected mice (four mice in each group) were humanely sacrificed at 1, 2, 3, 4, 5, 6, 7, 8, or 9 dpi or if they reached the criteria for euthanasia. Mice were humanely euthanized using isoflurane followed by cervical dislocation. Following euthanasia, the whole brain and the olfactory bulbs were removed from each mouse and fixed in 20 mL of 10% formalin for 24 h at 4°C. Brains were cut into two parts sagittally and cryoprotected with 30% sucrose in 0.1 M phosphate buffer pH 7.4. Brains were sectioned sagittally with a Leica CM3050 at 40 µM thickness.

Immunohistochemical Staining and Imaging of Mouse Brains

For immunohistochemical staining and imaging, brain sections were washed with PBS and endogenous peroxidases were quenched with 1% H2O2 in PBS for 5 min, rinsed with PBS three times and followed by blocking with 1% bovine serum albumin (BSA) in PBS and 0.3% Triton X-100. Floating sections were incubated with goat anti-VEE glycoprotein (GP) antibody (gift from Kurt Kamrud, AlphaVax, NC, USA) at a 1:12000 dilution for 2 days at 4°C on a shaker, rinsed with PBS x3 and incubated with biotinylated rabbit anti goat IgG at 1:200 (BA-5000 Jackson ImmunoResearch Laboratories PA USA) overnight. After rinsing with PBS three times, sections were incubated with avidin-biotin complex (ABC) using the Vectastain ABC Elite Kit PK-6100 (Vector Laboratories. USA) at a 1:100 dilution and developed with 3,3′-diaminobenzidine (DAB) (Sigma-Aldrich, USA). Sections were then counterstained with hematoxylin. Sections were mounted onto gelatin coated glass slides and air dried, dehydration through serial alcohol and xylene and coverslip with Permount. Digital images were taken with an Olympus VS 200 by scanning slides with a 40x objective. Images were captured and processed with the OlyVia version 3.2 software.

Histopathological and Immunohistochemical Staining and Imaging of the Nasal Cavity

To make paraffin sections for nasal epithelium, mouse skulls were fixed with 10% buffered formalin, dissected, decalcified with 10% ethylenediarninetetraacetic acid (EDTA), and embedded in paraffin. 5µm sections were obtained, mounted on Fisherbrand™ Superfrost™ Plus Microscope Slides (Fisher Scientific, USA), and stained with hematoxylin and eosin (H&E). For paraffin embedded nasal mucosa/epithelium sections, heat induced antigen retrieval was performed with Citrate-Based Antigen Unmasking Solution, (H3300 Vector Laboratories, USA). Sections were incubated with goat anti-VEEV GP antibody (gift from Kurt Kamrud) (AlphaVax, NC, USA) at a 1:1000 dilution. Sections were treated using ABC and DAB as described above. Digital images were taken with an Olympus VS 200 by scanning slides with a 40x objective. Images were captured and processed with the OlyVia v. 3.2 software.

Mouse Infection and Brain Harvesting for RNAseq

Five- to six-week-old, female, C3H/HeN mice were obtained from Charles River Laboratories and randomly assigned to one of four groups: group 1, uninfected control; group 2, TC-83 for day 1 sacrifice, group 3, TC-83 for day 3 sacrifice, group 4, TC-83 for day 5 sacrifice. Mice were infected i.n. with VEEV TC-83 diluted to a concentration of 1×10^7 PFU/30 µl, 15 µl per naris in PBS. PBS was used in place of virus for the mock-infection control group. Mice were weighed daily and checked twice daily for mortality and morbidity. Each group of VEEV TC-83 infected mice (four mice in each group) were humanely sacrificed at 1, 3, or 5 dpi. Following euthanasia, the whole brain, including the olfactory bulbs, was removed from each mouse. Each brain was separated into these eight

portions: olfactory bulb, piriform cortex, caudate-putamen, motor cortex, sensory cortex, thalamus, hippocampus, and cerebellum. Each section was homogenized in 1 mL of lysis buffer from the MagMAX™ mirVana™ Total RNA Isolation Kit (ThermoFisher Scientific). RNA was isolated from each brain section using the MagMAX™ mirVana™ Total RNA Isolation Kit and the KingFisher Flex system (ThermoFisher Scientific). RNA was sequenced using the Illumina HiSeq platform.

Sequencing Analysis

Sequence data was processed by first removing reads with < 50 nt length and Phred score < 20. Reads were mapped to the TC-83 seed stock sequence. Variants were called by identifying SNPs with a minimum frequency above 1% in all reads at positions throughout the entire genome with a minimum coverage of 400.

Network Analysis

Consensus sequences from each sample were obtained using CLC Genomics Workbench v.12 (Qiagen). Sequences were aligned using MUSCLE and mapped by the program PopArt v. 1.7 [133], utilizing a Minimum Spanning Tree to plot each sequence.

Results

Examining the Immunohistochemistry of VEEV Infection in the C3H/HeN Mouse

In this study, our goal was to examine the trafficking of VEEV strain TC-83 from the olfactory bulb to the brain following an i.n. infected C3H/HeN mice overnine days. Mock and infected mice were sacrificed from 1 to 9 dpi, inclusive. Brains and olfactory bulbs were harvested and sagittally sectioned, then examined through IHC for the presence of VEEV GP. A map of the brain sections examined are presented in **Figure 4-1**. Nasal turbinates at 1 dpi were also harvested, sectioned, and stained through H&E and IHC (**Figure 4-2**). These histological images were prepared by Dr. Yi Xue of UTHSC.

In nasal turbinate sections, viral antigen was visualized by IHC in the olfactory sensory neurons, covering their length from the cilia at the top of the epithelial layer down through a thin axon to the circular shape of the cell body (**Figure 4-2**). Deeper in the epithelium, the darker staining indicated infection of basal cells, supporting cells, and areas around Bowman's glands. In both cases, infected cells showed evidence of cell death through pyknosis and karyorrhexis, in which cell nuclei are condensed and fragmented respectively. The top of the epithelium layer was also disrupted at sites of infection, as shown with the gap at the largest site of infection.

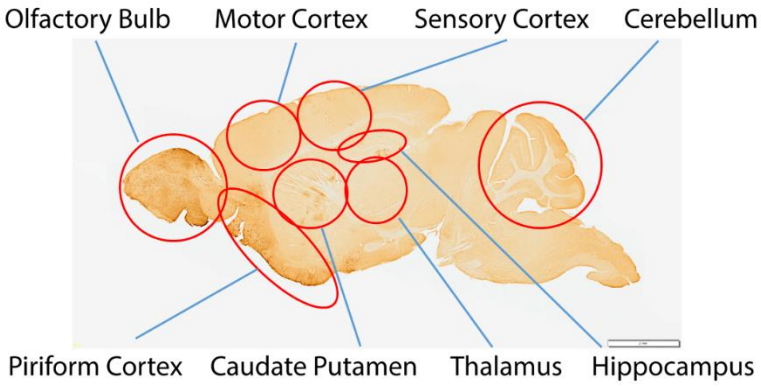


Figure 4-1. Mouse brain labeled with sections selected for RNA sequencing.

Mice were infected with a lethal dose of TC-83 and sacrificed at 5 dpi. Sagittal sections of brains were treated with anti-VEEV GP through the ABC method.

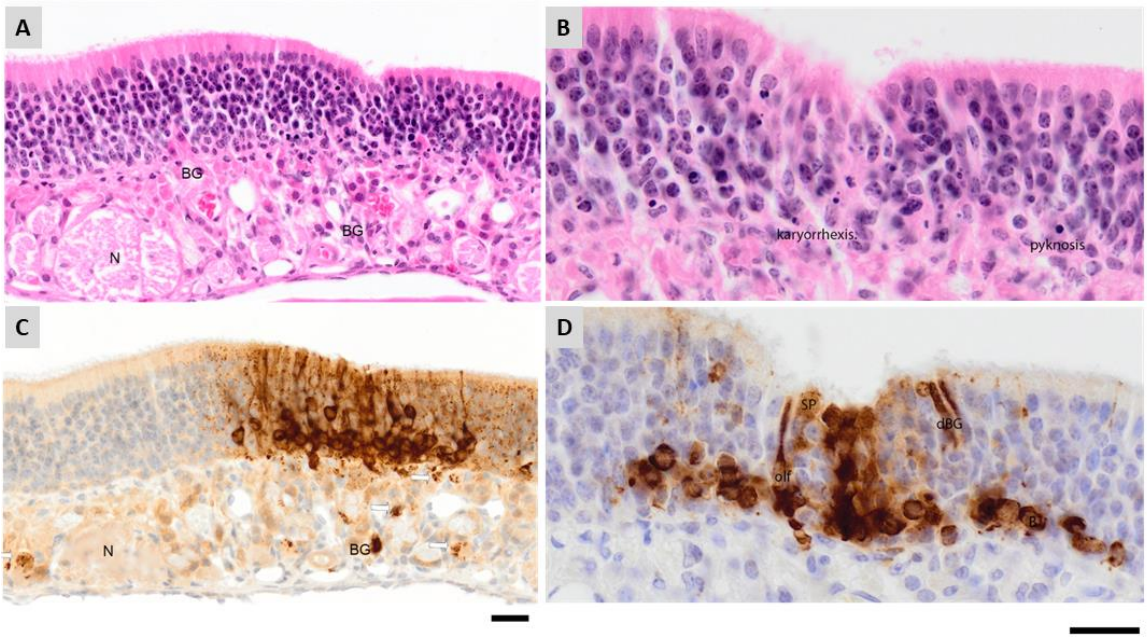


Figure 4-2. Nasal turbinate sections of C3H/HeN mice following i.n. infection with VEEV strain TC-83.

Images were taken at 1 dpi. Images A and C are images of the same location, as are B and D. (A) H&E staining shows the location of olfactory nerve bundles and Bowman's glands. (B) H&E staining shows karyorrhexis and pyknosis occurring in infected cells. (C) IHC staining for VEEV GP shows a variety of locations being infected, including olfactory neurons, basal cells, Bowman's glands, and supporting cells. (D) A closer look at the olfactory epithelium shows virus within olfactory neurons, supporting cells, basal cells, and the secretory duct of Bowman's glands. BG, Bowman's glands; N, olfactory nerve bundles; SP, supporting cells, dBG; secretory duct of Bowman's glands, olf, olfactory neuron. Scale bars = 20 µm.

IHC stains of VEEV GP each brain section over time are depicted in **Figure 4-3** and **Figure 4-4**. At 1 dpi, viral antigen was noted in the olfactory bulbs and the nasal cavity, although infection remained predominately in areas closest to the olfactory neurons. Tissue damage began at 2 dpi, after which virus expanded further in the olfactory bulb. The first entry of TC-83 into the brain occurs at 2 dpi, where it reaches the piriform cortex. Tissue damage in the piriform cortex began the following day and escalated in patches in successive time points (see subsequent days in **Figure 4-3**). At 3 dpi, viral GP was observed in the CPU, cerebral cortex, and thalamus. CPU tissue damage was strongest on the striated areas, with microscopic lesions appearing at 4 dpi. Individually infected neurons were seen in the cerebral cortex at 3 and 4 dpi, after which more widespread tissue damage occurred. While TC-83 reaches the thalamus at 3 dpi, infection was not established until 5 dpi, when larger groups of cells are infected. Viral GP antigen was detected in the hippocampus at 4 dpi, and finally, in the cerebellum at 5 dpi. In the hippocampus, lesions began to appear at 5 dpi, mainly within the dentate gyrus, where neurons connecting the outer and inner layers can be seen. Individual mossy fiber neurons connecting the dentate gyrus to the pyramidal cell layer showed GP antigen staining at 5 dpi (**Figure 4-4**). Lesions did not develop in the cerebellum during the infection. Individual Purkinje cells were infected as shown from 7 dpi onward.

Sequencing of VEEV TC-83 in the Brain Revealed a Wide Variety of SNPs but No Evidence of Bottlenecks or Adaptive Mutations Through Day 5

Building on the IHC study, we asked if the viral standing genetic variation from different portion of the brain were undergoing any bottlenecks or selective pressures during neuroinvasion of the brain. During this time period, the average weights of each mouse group did not significantly differ from each other, aside from between the PBS and day 3 sacrifice groups (**Figure 4-5**). Mice were sacrificed on days 1, 3, and 5, and eight different sections were individually removed for RNAseq, namely the olfactory bulb, piriform cortex, CPU, motor cortex, sensory cortex, thalamus, hippocampus, and cerebellum.

Because total RNA was used as the input for sequencing, a wide majority of sequencing reads, which ranged from sixty million to over eighty million reads per sample, did not align to the TC-83 genome, but for most samples there were enough reads to analyze at a depth of 400 or 1000 (**Figure 4-6**). However, all day 1 and cerebellum (all days) samples did not have enough viral genomic reads to provide sufficient depth for low variant. This was expected as very little virus had appeared to penetrate into the cerebellum based on the IHC. The highest number of TC-83 reads were measured in the olfactory bulb and piriform cortex on day 3, where the number exceeded 10^6 viral reads. TC-83 reads reached over 10^5 in the caudate putamen and the hippocampus at 3 dpi, while virus in the cerebral cortex and thalamus only exceeded this number at 5 dpi.

Next, we analyzed the mapped reads for SNPs that appeared with at least 1% frequency at locations with at least 400x coverage. We used a cutoff of 10% to identify

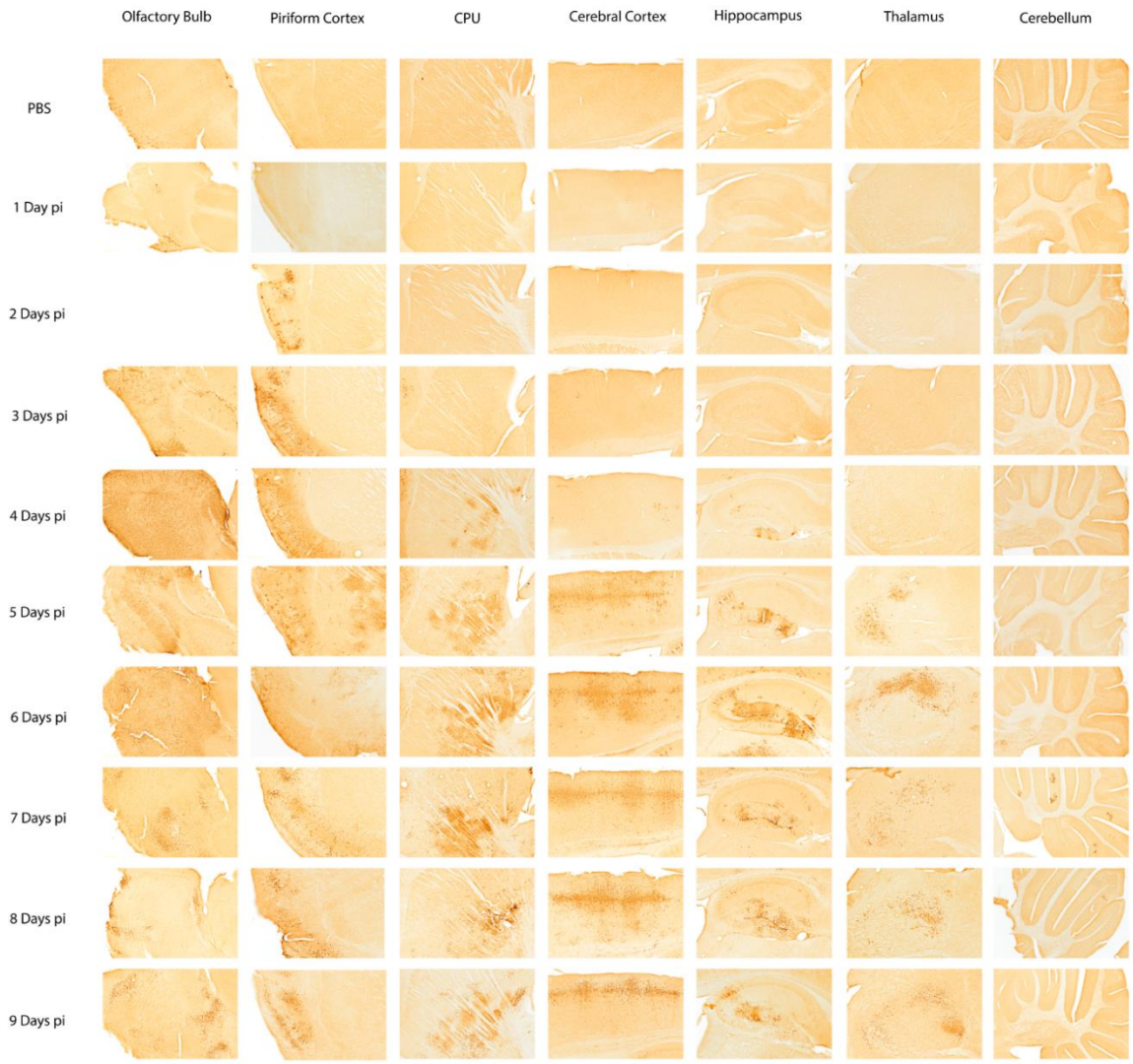


Figure 4-3. IHC of the VEEV GP in selected sections of C3H/HeN mouse brains following TC-83 infection.

Sagittal sections of brains are shown at seven representative regions of the brain. GP antigen was detected with anti-VEEV GP as described in the Materials and Methods. Antigen is indicated by a dark brown color. Scale bar = 200 µm.

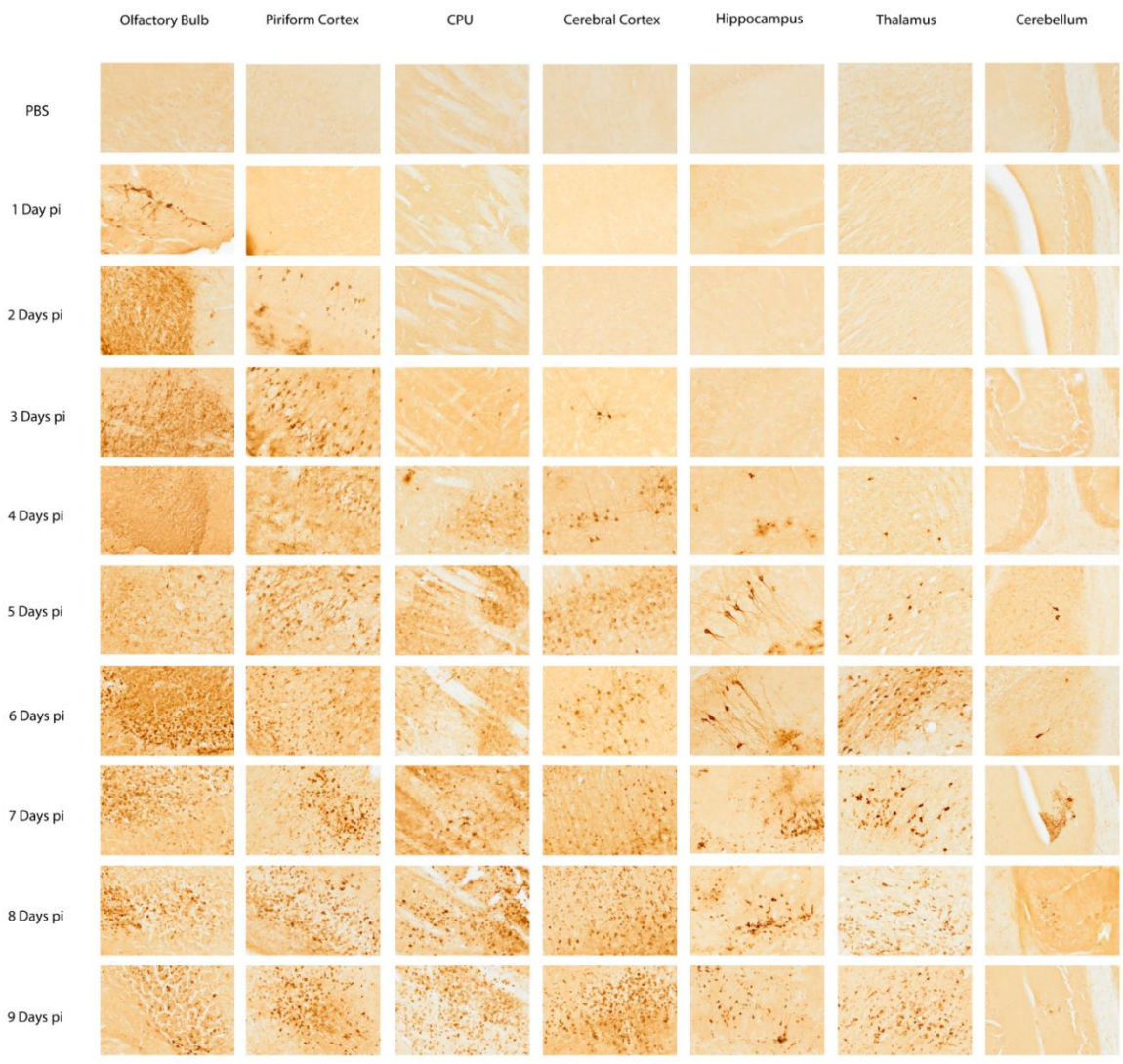


Figure 4-4. Close-up images of selected brain sections over time in mice following a TC-83 infection.

VEEV GP was detected in sagittal sections of brains by IHC as described in the Materials and Methods. Antigen is indicated by a dark brown color. Images are at 20X magnification. Scale bar = 50 µm.

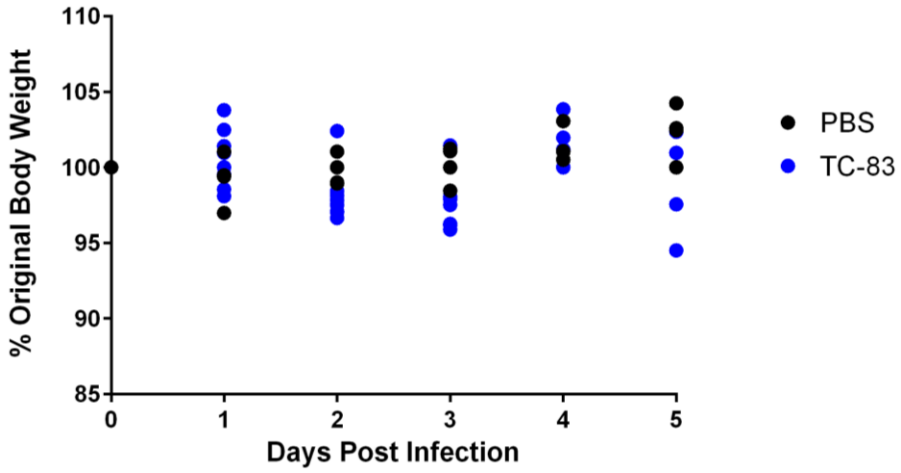


Figure 4-5. Weight change of mice.

Mice were i.n. infected with VEEV TC-83 and monitored over 5 days.

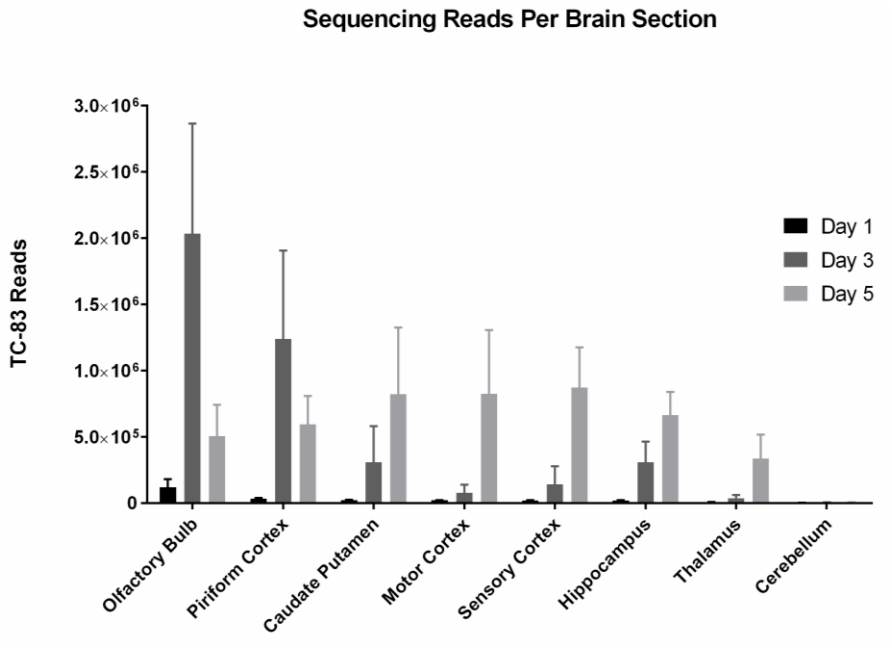


Figure 4-6. Average number of TC-83 sequencing reads in different C3H/HeN brain sections following infection.

Total RNA was sequenced and aligned to the TC-83 seed stock reference sequence.

SNPs, a summary of which are shown in **Table 4-1**. As previously stated, no SNPs were observed in day 1 samples as well as cerebellum samples due to the low number of reads. More SNPs were found in day 5 compared to day 3 samples in all brain sections. The most common SNP was an insertion of an adenine at genome position 2628, which never appeared at more than 20% frequency but was present in most samples. Other commonly found SNPs include C1151T in nsP1, which mostly appeared in Day 3 samples, C3958T in nsP2 which led to a threonine to methionine amino acid (AA) change, and A6257G in nsP4. Along with C1151T and C3958T, and A6257G, the SNPs that reached a majority in at least one sample included C5029T, which did so in the CPU at 3 dpi. However, aside from the 2628 adenine insertion, the C1151T SNP, and the G11149T SNP in the E1 gene, none of the SNPs listed appeared in more than one mouse; for example, both the A3154T and C3958T nucleotide changes were only found in a singular mouse brain at 5 dpi. In conclusion, the SNPs obtained from TC-83 populations in the brain were varied in number and frequency and covered most genes, but aside from 5 dpi samples generally containing more SNPs, there were no particular patterns in their appearance.

Network Analysis of the TC-83 Population Consensuses from Each Brain Section Suggested No Significant Changes

To further evaluate the consensus nucleotide genomes of all samples, we used a network analysis using the consensus sequence of each sample. All sequences, except for cerebellum sequences, were aligned along with the seed stock TC-83 genome consensus and a GenBank reference sequence (accession number L01443.1, GI: 323714) using MUSCLE, then plotted using PopART v.1.3 (**Figure 4-7**). Most of the sequences were grouped into the same circle as the seed stock sequence, showing no change from the reference sequence. The remaining sequences branched off in three different directions. TC-83 from the piriform cortex, motor cortex, and hippocampus of single mouse sacrificed at 3 dpi had the C1151T nucleotide substitution. One mouse from day 3 had the nsP3 T333I mutation, and a day 5 mouse had the nsP2 T770M and the nsP4 A6257G nucleotide substitution.

Discussion

In this study, we examined the prevalence of viral antigen, viral genomes, and genetic variation of viral genomes over time in the mouse brain following an infection with VEEV TC-83. The IHC analysis of TC-83-infected C3H/HeN brains show that infection reaches the olfactory bulbs within 24 hours after intranasal infection. By day 3, the piriform cortex shows signs of cellular damage, while virus is detected in the caudate putamen, cerebral cortex, and thalamus. By day 5, these areas as well as the hippocampus show lesions, while infection progresses to the cerebellum. These results match up well with previous data on the presence of viral antigen in TC-83-infected

Table 4-1. Summary of SNPs that appeared in VEEV TC-83 when infecting mouse brains.

Nucleotide Change	Amino Acid Change	Gene	Olfactory		Piriform		CPU		Motor		Sensory		Thalamus		Hippo-campus	
			Day 3	Day 5	Day 3	Day 5	Day 3	Day 5	Day 3	Day 5	Day 3	Day 5	Day 3	Day 5	Day 3	Day 5
C421T	P126L	nsP1														
C1151T		nsP1														
A1362G	I440V	nsP1														
T1547C		nsP1														
C1969T	A107V	nsP2														
C2018T		nsP2														
C2161T	T171I	nsP2														
C2627A	D326E	nsP2														
-2628A		nsP2														
C2850T		nsP2														
A2985C	I446L	nsP2														
C3029T		nsP2														
C3045A	P466T	nsP2														
A3154T	D502V	nsP2														
T3396C		nsP2														
C3626T		nsP2														
C3932T		nsP2														
C3958T	T770M	nsP2														
G4386A	D119N	nsP3														
C5029T	T333I	nsP3														
G5468T		nsP3														

Table 4-1. Continued.

Nucleotide Change	Amino Acid Change	Gene	Olfactory		Piriform		CPU		Motor		Sensory		Thalamus		Hippo-campus	
			Day 3	Day 5	Day 3	Day 5	Day 3	Day 5	Day 3	Day 5	Day 3	Day 5	Day 3	Day 5	Day 3	Day 5
A6257G		nsP4														
T6323C		nsP4														
G6533A		nsP4														
A6939T	I413L	nsP4														
C6977T		nsP4														
C7951T		C														
C8924T	H121Y	E2														
G10139A	V47I	E1														
C11036T		E1														
G11149T		E1														

Samples are grouped based on brain portion and dpi. Cells are shaded based on the highest percentage they appeared in at least one sample in the given brain portion and day; red, > 50%; orange, 20-50%; yellow, 10-20%.

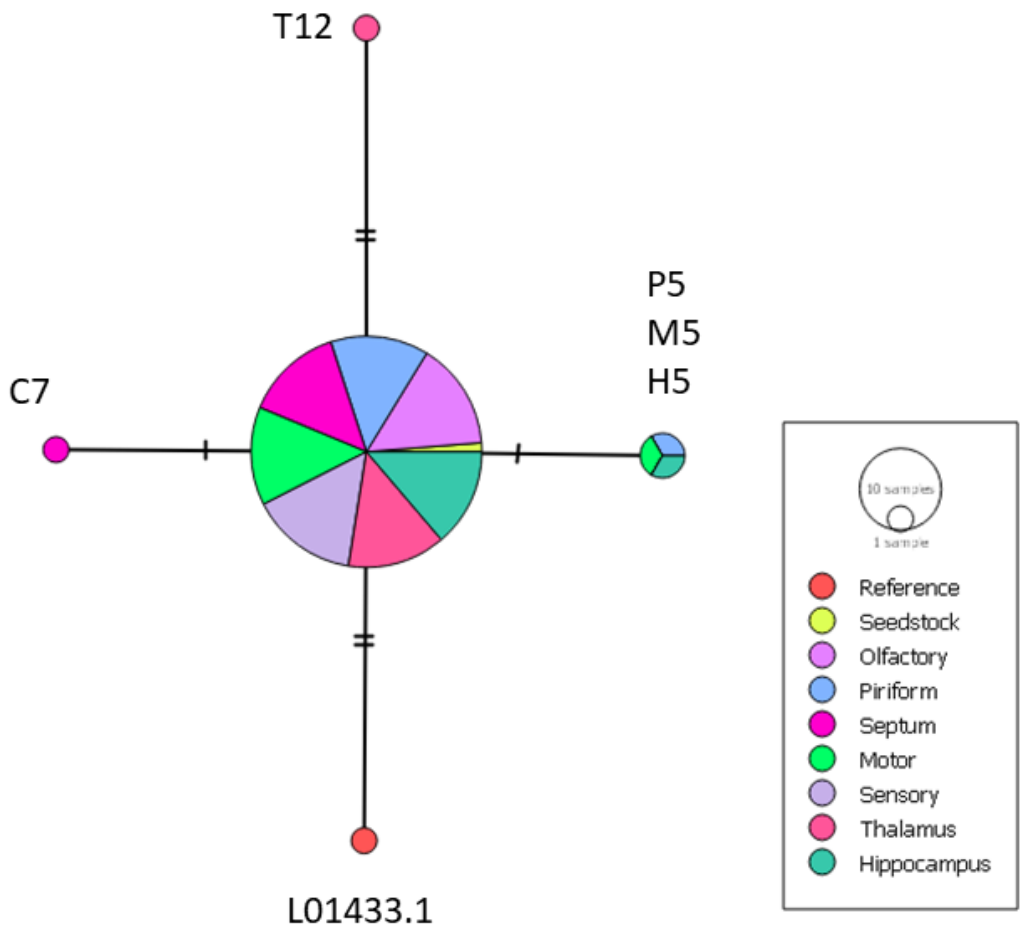


Figure 4-7. Network analysis of VEEV TC-83 consensus sequences shows that the consensus generally does not change when inside the brain.

Consensus sequences were obtained from NGS sequencing data analysis, aligned using MUSCLE, and plotted based on nucleotide differences. Samples were grouped and labeled based on portion of brain and day of extraction. Each tick represents the number of nucleotide differences between sequences. The reference sequence was taken from GenBank (accession number L10433.1). Other than the reference and seed stock sequences, all sequences are labeled by brain section and individual mouse number (1-4, Day 1 sacrifice; 5-8, Day 3 sacrifice; 9-12, Day 5 sacrifice). Septum = Caudate Putamen.

C3H/HeN mice, in which the time points for the appearance of virus in the piriform cortex, caudate putamen, thalamus, and hippocampus were the same [28]. For the first time, we also show individual neurons that are infected by VEEV, including neurons crossing the inner and outer layers of the dentate gyrus in hippocampus and Purkinje cells in the cerebellum. The most common symptoms of mouse infection with VEEV include weight loss, hunching, and lethargy, causing a lack of movement that is observed in infected mice that may be explained by the areas of the brain with extensive infection, as the caudate putamen, motor cortex, and cerebellum are all involved in motor control and movement, while the hippocampus helps with spatial orientation. In addition, the thalamus assists in consciousness and alertness while acting as a relay center for most sensory pathways. While these symptoms have been observed in previous studies of TC-83-infected C3H/HeN mice, none of those studies attempted to make direct links between disease symptoms and the presence of viral antigen in specific portions of the brain over time [28, 132].

In line with the histological studies, at 1 dpi, the number of viral reads in the olfactory bulb were high enough to run low variant analysis, while they were too low in all other sections of the brain. However, no SNPs were detected at 1 dpi, which suggests that VEEV TC-83 was already well-adapted to infecting olfactory bulb neurons. At 3 and 5 dpi, the piriform cortex, caudate putamen, motor cortex, sensory cortex, thalamus, and hippocampus contained virus populations where mutations started to appear. Through 5 dpi, there were not enough viral reads to conduct low variant analysis on cerebellum samples, corroborating with the low number of infected neurons detected at that time point through histology. Since C3H/HeN mice usually succumb to VEEV TC-83 infection at 6-7 dpi onwards, as shown here and in previous research, it may be worth investigating along further time points to examine the dynamics of these SNPs and whether they are retained over a longer period of time [28, 132, 136].

The low variant analysis found a variety of SNPs, including in all of the nonstructural genes, but most of these were only present in a single mouse, and there were few, if any, changes in the consensus sequences of the VEEV populations, suggesting that these SNPs arose randomly instead of to a particular selective pressure, at least through Day 5. To our knowledge, none of the mutations that we have shown here have been previously found in mutational studies with VEEV, other than the D326E mutation in nsP2, which was denoted as a mutational hotspot that remained even in a low-fidelity polymerase mutant [103]. This is also the spot in the VEEV genome where an insertion was frequently observed in most of the sequenced brain samples (**Figure 4-7**), suggesting that it may be a feature of the virus that may aid replication in some manner. D326 is located in the helicase/NTPase domain of nsP2, so this mutation could be aiding that activity. The number of recurring synonymous mutations such as C1151T could suggest that these are also features of VEEV, and that they may play a role in RNA secondary structure influencing replication. For example, C1151 itself is adjacent to a packaging signal, which takes the form of a secondary structure that is found in all alphaviruses [39, 137]. However, contributions of secondary structure in the VEEV genome to pathogenesis and the life cycle are currently largely unexplored [138]. On the other hand, the recurring mutations could also simply be artifacts from the sequencing

process, where errors can be introduced at various steps, including during library preparation and sequencing [139]. Regardless, the fact that all of the other mutations were only found in a single mouse suggests that they are random and were not selected for in VEEV TC-83 to confer a fitness advantage in mouse neurons.

The findings in this study come from the C3H/HeN mouse strain infected i.n. with the TC-83 strain of VEEV. While TC-83 does not appear to adapt and evolve to selective pressures within the brain, at least during the time points given, there may be many factors that could influence the evolution of the viral population if it is not given a fast track to the brain through intranasal or aerosol infection. It takes VEEV more than a day to reach the olfactory epithelium in a natural or subcutaneous infection, and during that time the virus replicates in lymph nodes and spreads to many parts of the body, including organs such as the liver, kidneys, and spleen. It may be possible that viruses that are able to infect and inflict lesions on these organs are different in sequence to those of the initial infection, or to the viruses that eventually make it to the nasal epithelium and the CNS. During this time, VEEV looks to evade the innate immune response, which adds another selective pressure. The population dynamics may also slightly differ when using alternative mouse models of VEEV infection, such as with the infection of BALB/c mice with the Trinidad Donkey strain of VEEV. Application of these results to what may happen in a human infection is also limited, as currently there is not enough information on the pathogenesis of VEEV in human infections to make meaningful comparisons.

In summary, we show both the progression of VEEV TC-83 infection in the C3H/HeN mouse model and the sequences of virus within eight different sections of infected brains. We showed that there is a lack of selective pressure on VEEV as it spreads in the brain during an infection as there are no mutations that point to a specific adaption. With an understanding how VEEV may evolve during an infection, testing its potential by observing its evolution through the duration of the infection, and determining how that evolutionary trajectory is altered by the addition of selective pressures such as antiviral compounds, will add to the knowledge of how viruses adapt to their environments and to selective pressures.

CHAPTER 5. CONCLUSIONS

Conclusions

Venezuelan equine encephalitis virus (VEEV) is an emerging zoonotic pathogen that has the potential to cause outbreaks in the Americas. This potential, along with its potential use as a bioweapon, gives rise to a need for vaccines and antivirals against the virus, as none so far have been FDA-approved. Our lab has been part of an effort to develop anti-VEEV antivirals, which resulted in the discovery of compounds such as ML336. However, the emergence of antiviral resistance is a concern with any compound, so our efforts herein were to identify the mutations that confer antiviral resistance, as well as explore the evolutionary trajectory and population dynamics of VEEV as resistance develops. First, we designed a blind passaging experiment with increasing concentrations of compounds to select for mutants that were resistant to those compounds over time. Secondly, we also examined the effect of natural selective pressures on the evolution of VEEV populations in different cell types and in the mouse brain. In this chapter, the main findings of the experiments are summarized and put into context of current knowledge of VEEV, RNA virus population dynamics, and the alphavirus nonstructural proteins nsP2 and nsP4. Future directions to follow up on the results herein are also proposed.

In Chapter 2, we developed a pipeline for next-generation sequencing (NGS) sequencing VEEV populations to examine the overall sequence composition and detect any SNPs that were present. We then applied that pipeline to studies in which we examined how the VEEV population sequences changed over time in response to increasing concentrations of the antiviral compound ML336. We determined that the resistance profiles in biological replicates of VEEV show different paths to evolving resistance to ML336, but in all cases reached a consensus at amino acid Q210 within the amino terminus of the nsP4 gene to a basic amino acid (i.e. K, R, or H). Upon removal of ML336, the mutations were for the most part stable in the population through seven passages. Using a similar approach with increasing concentrations of ML336, we showed that the emergence of the Q210 mutations when passaging in the astrocyte cell line SVGA was in general lower over most time points than Vero 76 cells.

In Chapter 3, we applied the same approach to repeated passaging and sequencing to VEEV, using a variety of hit compounds from the quinazolinone and benzamidine scaffolds. We found that there was a large difference in evolutionary trajectory in the emergence of resistance-conferring mutations, in both the timing of emergence and the mutations that appeared. Mutations did not appear until the third passage with BDGR-4, at 200 nM. In contrast, mutations emerged in the very first passage, at 50 nM, of both BDGR-49 and BDGR-164, with the Q210R rapidly reaching nearly 100% penetration in the population. For BDGR-163, mutations also emerged quickly, but a completely different SNP profile emerged as the primary mutations were those found in nsP2, with the Q210 mutations appearing but at much lower frequencies.

In Chapter 4, we examined the natural evolution of VEEV in the central nervous system (CNS) of mice to determine if there were any selective pressures acting on the virus as it moves through the brain during an infection. Based on our studies on the spatial migration and pathogenesis of VEEV in the mouse brain after an intranasal (i.n.) infection, we selected multiple sections of the brain to harvest and sequenced the virus within those sections at multiple time points. A variety of SNPs were found, but most were unique to each individual mouse, and there was no indication of any particular selective pressures upon VEEV in the CNS through day 5. A network analysis showed that the virus consensus barely changed over 5 days.

The Ability of VEEV to Evolve and Rapidly Mutate in Response to Antivirals

As an RNA virus with an error-prone RNA-dependent RNA polymerase (RdRp), VEEV has a high mutation rate that, when combined with the high titers generated in cell culture and in animal models, leads to the creation of a large population of related, yet genetically distinct viruses. Herein, we used next-generation sequencing methods to capture that genetic diversity over time, and address how it changes when selective pressures are applied. We showed that small molecules with an antiviral effect on VEEV, including ML336 and its derivatives, impose a strong selective pressure in vitro. We also showed that neurons and neuron-adjacent cells such as astrocytes do not impart a strong positive selective pressure on the VEEV genome. In every experiment, VEEV populations are not homogenous, and that even in systems where there is a lack of selective pressure, the virus will gain and lose mutations through genetic drift. This is likely the best explanation for the SNPs seen in structural genes, particularly in E1 and E2, and for the majority of synonymous mutations.

Mutations at five particular locations in the genome were recurrent throughout the studies and their approximate locations on the genome are shown in **Figure 5-1**. Amino acids A461 in the nsP3 gene, and Y87, Q210, E213, and R305 in nsP4 appeared in in vitro studies with multiple compounds. In addition, our experiments revealed that multiple mutations of Q210, sometimes within the same virus population, were possible. Schroeder et al. [107] found the Q210K mutation, but we discovered that Q210R and Q210H could also be present. While Q210R was the most common of the three, it was found alongside Q210H in some samples when TC-83 was passaged with ML336 and BDGR-4. In contrast, passaging TC-83 with BDGR-163 resulted in virus populations carrying Q210K and Q210R, with the former being more dominant, but in this case neither were the dominant mutations found in those populations, with nsP2 mutations at Y102, D116, and E117 the most prominent.

However, none of these mutations were detected *in vivo* when ML336 was used to treat TC-83-infected mice. In this case, the addition of ML336 mostly prevented infections from becoming lethal, but while virus titers were fairly high in treated mice through 14 days post infection (dpi), few virus sequences were able to be isolated and sequenced, leading to incomplete coverage across the TC-83 genome after NGS for many samples, and thus a paucity of SNP data. Indeed, most of the major SNPs found from

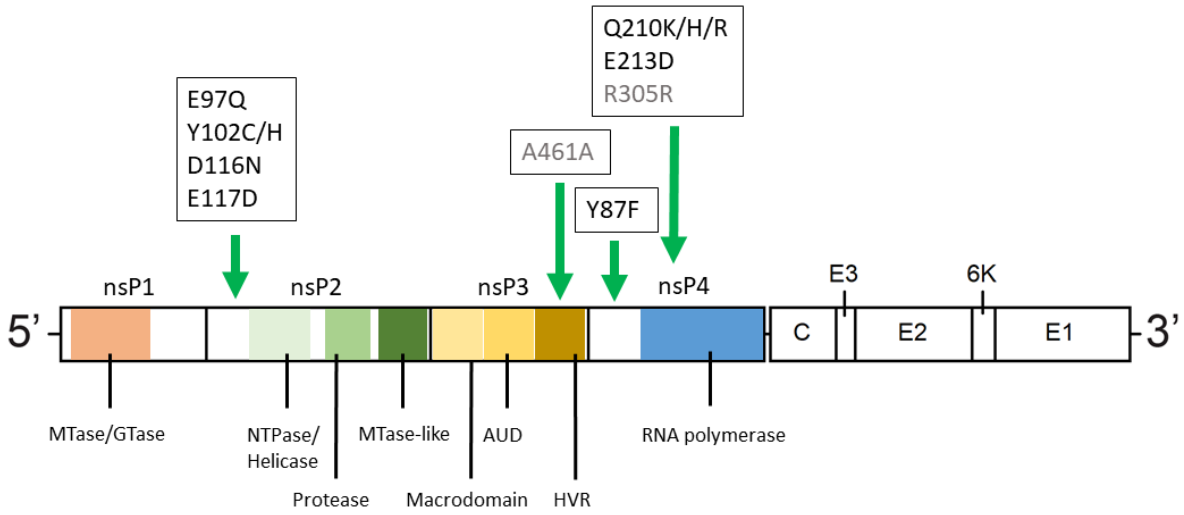


Figure 5-1. VEEV genome with major domains and SNPs found in this dissertation's studies marked.

Nonsynonymous mutations are written in black, while synonymous mutations are written in gray. The green arrows point towards the domain and approximate nucleotide position each SNP is located in within the VEEV genome.

this experiment were from untreated mice, appearing to be random mutations that emerged by 7 and 8 dpi. A more direct method of sequencing TC-83 from brain sections at earlier time points of 1, 3, and 5 dpi showed a larger variety of SNPs that, for the most part, were unique to each individual mouse. Remarkably, in terms of SNPs that reached a penetrance of at least 10% in a given sample, zero SNPs were found in common between the two mouse experiments. The SNPs that were found in more than one mouse, namely the C1151T and G11149T substitutions and the 2628A insertion, were already discovered to be present in the TC-83 genome at low levels when analyzing NGS data from seed stock sequences, showing that they were likely already present in the population before infection.

The timing of the emergence of antiviral resistance-conferring mutations was the other key difference between the compounds tested in these studies. For ML336, resistance-conferring mutations started emerging at the second passage, at 100 nM, while for BDGR-4 it took until 200 nM for mutations to appear. On the opposite end, TC-83 evolved in the very first passage, at 50 nM, for BDGR-49, BDGR-163, and BDGR-164. This result also appeared to hold true for BDGR-49 when passaged at 20 nM, although to a much lesser extent than at 50 nM.

Roles and Functions of nsP2 and nsP4 Mutations in the Replicase Complex

Previous studies with CID15997213 and ML336 suggest that compounds in these benzamidine and quinazolinone scaffolds work by direct interaction with the replication complex, interfering with the function of nsP2 and nsP4 in particular. The data shown in this study corroborates this theory in that the major mutations that emerged through repeated passaging appeared in nsP2 and nsP4. Some of these mutations either had been already previously-discovered, such as nsP4 Q210K, which was first discovered with ML336 [107]. Another category of mutations were those adjacent to these previously-discovered mutations, such as nsP2 E117D, which is adjacent to the D116N mutation and appeared with BDGR-163. Others were simply different mutations of the same amino acid, such as with Q210R and Q210H appearing at the same time when passaging TC-83 with BDGR-4. Lastly, there was the Y87F mutation, which appears in the N-terminal domain of nsP4, an area without a known function despite being essential to function of the protein. As far as we are aware, these mutations have not been previously discovered in other circumstances, which strengthens the idea that they are particular to the compounds we have tested. Another interesting observation with these mutations was that, despite nsP2 mutations dominating in TC-83 when challenged with BDGR-163, Q210 mutations slowly emerged in those populations over time, suggesting that the nsP2 mutations could be more transitory and give way to the Q210 mutants later on. In that case, Q210 mutations could be a stronger adaptation against the antivirals than the nsP2 mutations.

We currently theorize that, because of these mutations and other studies done to determine the mechanism of action, the compounds act by inhibiting viral replication through direct interaction with nsP2 and nsP4 (**Figure 5-2**). A study with ML336

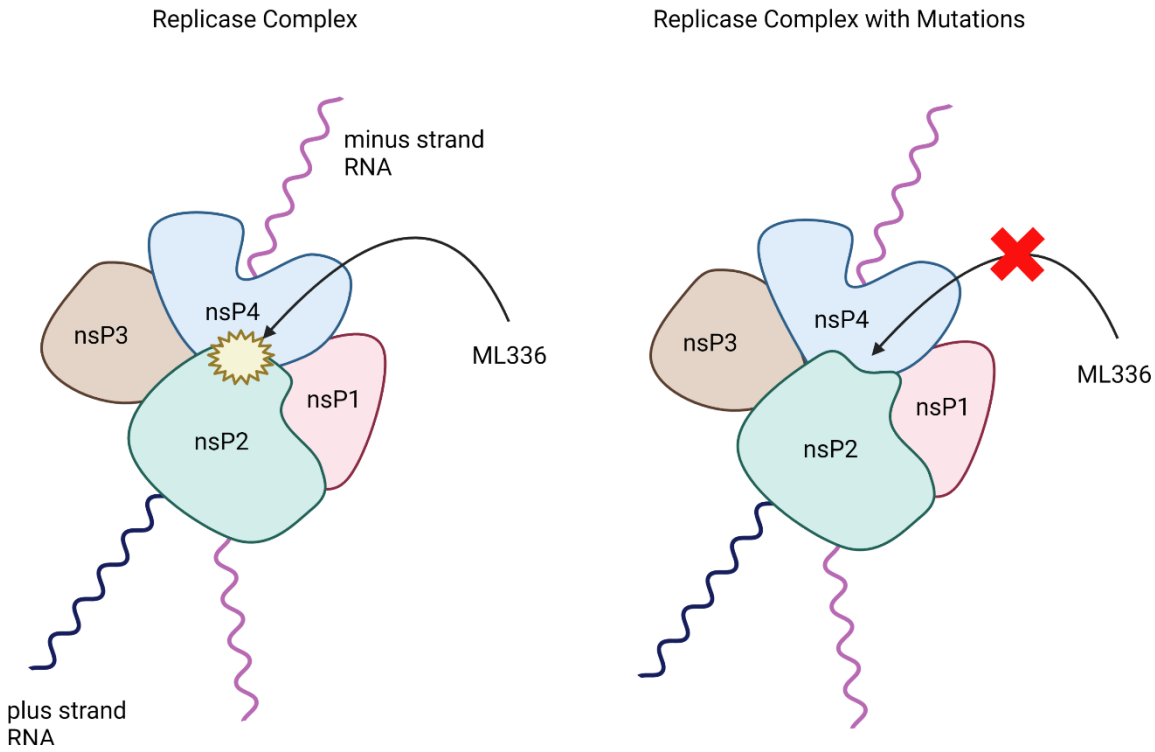


Figure 5-2. Schematic of the mature replicase complex showing all of the nonstructural proteins and where ML336 and other compounds may be inhibiting viral replication.

showed that both positive and negative strand synthesis are inhibited by ML336, and that ML336 inhibits the mature replicase complex containing all four individual non-structural proteins, including the synthesis of both genomic and subgenomic RNA [118]. Since both the RdRp of nsP4 and the helicase and NTPase of nsP2 are active during plus-strand synthesis, there may be a direct interaction between nsP2 and nsP4 which ML336 and the other compounds are inhibiting. Because minus-strand synthesis is also inhibited, however, there may also be interactions between nsP4 and the nsP123 polyprotein, although the instability of the nsP123/4 complex would make it difficult to observe compared to the nsP1/2/3/4 complex. Skidmore et al. [118] also showed that the nsP2 Y102C and nsP4 Q210K mutations conferred resistance to the inhibition of RNA synthesis, and since both mutations are located in regions of currently unknown function in their respective proteins, we hypothesize that these regions provide structural support to the replication complex and may be in direct contact near these particular amino acids. There is some evidence using Chikungunya virus that nsP2 and nsP4 are directly interacting within the replicase complex, so while the formation of the spherule and the arrangement of viral and host proteins within the replicase complex have not been explored with VEEV in particular, research with other alphaviruses appears to point towards it being the case [140].

The location of Q210 on the protein gives some further insight as to how the protein complex is arranged, and how ML336, BDGR-4, etc. may interact with the viral proteins. A very recent study used Ross River virus to produce what likely the first to produce a crystal structure of alphavirus nsP4, showing that Q210 falls into the fingers structural portion of the RdRp, specifically the pinky [141]. If VEEV nsP4 has a similar structure, it would support the hypothesis that Q210 is located on the exterior of the protein facing nsP2. Whether this is a site of contact with nsP2 remains to be addressed, however.

In regards to the nsP2 mutations, their simultaneous appearance with Q210 mutations in some passaging experiments, as well as their position within the gene, may suggest that they assist in conferring resistance but are not as fit, instead causing fitness loss to the virus. Many RNA viruses, including alphaviruses, have a helicase and polymerase that act in tandem during replication [142, 143]. Mutations in nsP2 and nsP4 conferring resistance to nucleoside analogues that appeared in tandem in Chikungunya virus (CHIKV) have been observed [144]. However, this idea may not be able to explain why there was a lack of nsP2 mutations in some of the passaging experiments, particularly with BDGR-49.

Interestingly, some recent studies with other compounds with an inhibitory effect on the VEEV replicase complex have also found mutations in nsP2 and nsP4. A study using the nucleoside analogue β-D-N4-hydroxycytidine found that mutations conferring resistance to it include P187S, A189V, and I190T in nsP4 [113]. Later passages also revealed presence of A201V and V308I mutations, also in nsP4. A study altering the fidelity of the RdRp found that there were certain mutational hotspots, mostly in the structural genes, but a couple, namely C2627A and A2634Del, were found in nsP2 [103]. Interestingly, the former is a SNP we have also found, but from TC-83-infected mouse

brains, which suggests that mutations at this point in the genome are regions of inherent variability for VEEV. This is compounded with the repeated appearance of the 2628A insertion, which appears near 10% in many samples throughout all of the studies described in previous chapters. However, the 2628A insertion, if added to the genome, causes a frameshift that results in a truncated nsP2 protein. This seemingly deleterious mutation could simply be explained by the fact that it appears in a string of adenine repeats, however. Another alternate explanation is that defective virus genomes are being captured when sequencing total RNA from a sample, as studies with other alphaviruses show they are produced, but defective virus genomes are a topic for VEEV that has yet to be explored [145].

Future Directions

Despite finding that there is little selective pressure on VEEV on the first five days of infection as it moves through the CNS, our results show that it is very much possible to detect and analyze viral sequences from smaller portions of the brain. Obviously, this means that the same technique can be used to sequence viruses from infected mice treated with antivirals in order to see if the mutations found with repeated passaging *in vitro* will be found in the mouse, but this may be challenging as shown in **Table 2-6**. Another adaptation of this experiment would be to characterize the VEEV population as it moves through entire body as opposed to just brain, using a subcutaneous or hock infection to better mimic a natural infection. Spleen, liver, lungs, kidneys, and lymphoid tissue would be harvested at multiple time points for RNA isolation and sequencing to determine the SNPs found within the viral population and whether the populations differ and change over time during the course of the infection.

The C3H/HeN model of the mouse is used mainly because it is the only model in which TC-83, a strain of VEEV allowed at biosafety Level 2, can cause lethal infection. However, it may be more informative to use a nonlethal model using wild-type, pathogenic strains of VEEV, such as Trinidad Donkey, to more accurately model natural infections and their resolution. Use of an alternative mouse model, such as BALB/c, that are symptomatic but survive the infection over time, and sequencing the virus within the brain at different time points could present a different picture than the one shown here, particularly if the virus is found to persist within the CNS [28]. A further application would be to apply this method of examining virus population variation within hosts to other alphaviruses, particularly the lesser-explored Eastern equine encephalitis virus (EEEV) and Western equine encephalitis virus; indeed, there are already studies showing that there is intrahost variation in EEEV in human infection [146].

Because VEEV is only spread between humans through mosquitos that carry the virus from one infected host to another, it may be worth asking whether mosquitos are able to carry virus populations with antiviral-resistant mutations. Studies have already demonstrated that the mosquito midgut and salivary glands, while major sites of replication, present a bottleneck for a VEEV population looking to establish an infection [147]. Would mutations conferring antiviral resistance make it through these

bottlenecks? Some research has already shown that polymerase mutations become attenuated in both mosquito and mammalian hosts, but the mutant strain was designed to increase fidelity, thus decreasing the possible sequence space for the viral population [102]. Regardless, it would be worth examining how VEEV populations may change during the course of infection in a mosquito, as a major hypothesis of how epizootic outbreaks occur is that VEEV mutates from an enzootic (ID, IE) to an epizootic (IAB, IC) strain that allows it to replicate in equids. Directly observing this result in the mosquito would not only confirm this hypothesis but also provide a more direct window into how VEEV populations evolve over time in nature. An experiment to test this would be to infect mosquitos with VEEV containing antiviral-resistant mutations, then isolating and sequencing the viral RNA found in the midgut and salivary glands to determine the population composition at those locations and to detect the presence or absence of those mutations [99].

Q210 and Y87 are conserved residues throughout all alphaviruses, while the nsP2 residues Y102 and D116 are seen in some Old World Alphaviruses to be a lysine and glutamic acid respectively [106]. The compounds used in our experiments have not been extensively tested with Old World Alphaviruses, with Chikungunya virus (CHIKV) being an exception; CID15997213 was found to have little antiviral activity against CHIKV [106]. The more-recently synthesized compounds have not been tested against CHIKV or other Old World Alphaviruses, which leaves open the possibility that these derivatives may be more effective at inhibiting them compared to the original compounds. However, the comparison of amino acid sequences still raises the question of why these compounds work on some alphaviruses and not on others. Determining differences between the protein structures of nsP2 and nsP4 would help in that regard, but another method of testing these differences would be to, for example, changing nsP2 Y102 to lysine, then running titer reduction assays with the antiviral compounds to determine the extent of viral replication inhibition.

The repeated presence of mutations such as C1151T and the 2628A insertion raises some interesting questions. Are these artifacts from sequencing, or are they SNPs that may serve a purpose by being present in the viral population? C1151T is a synonymous mutation, so if it serves a function it may be related to RNA secondary structure. Little work has been done to examine RNA secondary structure of alphaviruses thus far, but a computational study predicted that most secondary structures were not conserved between alphaviruses, or particularly between Sindbis virus (SINV) and VEEV [137]. An experiment in which positive and negative-sense RNA are selectively isolated and sequenced may also be helpful in determining the possible roles of these mutations, especially when combined with secondary structure predictions to delve further into why VEEV may need these seemingly silent mutations.

There has been no major outbreak of VEEV since the 1995 Colombian-Venezuelan outbreak of VEEV that resulted in the infection of over 100,000 individuals. However, the virus has not disappeared but has stayed endemic in rural areas throughout South and Central America, where it is still spilling over into humans, but only causing the milder, flu-like symptoms that keeps VEEV indistinguishable from other viral

diseases, allowing it to stay under the radar without molecular tools to identify it [1]. However, as the world has recently learned with the COVID-19 pandemic, and with many more pandemics and epidemics that have preceded it, emerging RNA viruses can cause outbreaks with widespread, expensive, and long-lasting effects on human civilization and society, and we need to be sufficiently prepared to deal with outbreaks and the propensity of viruses to constantly evolve over time.

LIST OF REFERENCES

1. Aguilar, P.V., et al., *Endemic Venezuelan equine encephalitis in the Americas: hidden under the dengue umbrella.* Future Virology, 2011. **6**(6): p. 721-740. DOI: https://doi.org/10.2217/fvl.11.50.
2. Zacks, M.A. and S. Paessler, *Encephalitic alphaviruses.* Vet Microbiol, 2010. **140**(3-4): p. 281-6. DOI: https://doi.org/10.1016/j.vetmic.2009.08.023.
3. Smith, D.R., et al., *Venezuelan equine encephalitis virus in the mosquito vector Aedes taeniorhynchus: infection initiated by a small number of susceptible epithelial cells and a population bottleneck.* Virology, 2008. **372**(1): p. 176-86. DOI: https://doi.org/10.1016/j.virol.2007.10.011.
4. Paessler, S. and S.C. Weaver, *Vaccines for Venezuelan equine encephalitis.* Vaccine, 2009. **27 Suppl 4**: p. D80-D85. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2764542/.
5. Weaver, S.C., et al., *Venezuelan equine encephalitis.* Annu Rev Entomol, 2004. **49**: p. 141-74. DOI: https://doi.org/10.1146/annurev.ento.49.061802.123422.
6. Rossi, A.L., *Rural epidemic encephalitis in Venezuela caused by a group A arbovirus (VEE).* Progress in Medical Virology, 1967. **9**(0079-645X (Print)): p. 27. https://pubmed.ncbi.nlm.nih.gov/4383416/.
7. Zehmer, R.B., et al., *Venezuelan equine encephalitis epidemic in Texas, 1971.* Health services reports, 1974. **89**(3): p. 278-282. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1616214/.
8. Weaver, S.C., et al., *Re-emergence of epidemic Venezuelan equine encephalomyelitis in South America.* The Lancet, 1996. **348**(9025): p. 436-440. DOI: https://doi.org/10.1016/s0140-6736(96)02275-1.
9. Rivas, F.D., L A; Cardenas, V M; Daza, E; Bruzon, L; Alcala, A; De la Hoz, O; Caceres, F M; Aristizabal, G; Martinez, J W; Revelo, D; De la Hoz, F; Boshell, J; Camacho, T; Calderon, L; Olano, V A; Villarreal, L I; Roselli, D; Alvarez, G; Ludwig, G; Tsai, T., *Epidemic Venezuelan Equine Encephalitis in La Guajira, Colombia, 1995.* J. Infect. Dis., 1995. **1997**(175): p. 828-832. DOI: https://doi.org/10.1086/513978.
10. Medina, G., et al., *Genetic diversity of Venezuelan alphaviruses and circulation of a Venezuelan equine encephalitis virus subtype IAB strain during an interepizootic period.* Am J Trop Med Hyg, 2015. **93**(1): p. 7-10. DOI: https://doi.org/10.4269/ajtmh.14-0543.
11. Kinney, R.M., et al., *The full-length nucleotide sequences of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and its attenuated vaccine derivative, strain TC-83.* Virology, 1989. **170**(1): p. 19-30. DOI: https://doi.org/10.1016/0042-6822(89)90347-4.
12. Kinney, R.M., et al., *Nucleotide sequences of the 26S mRNAs of the viruses defining the Venezuelan equine encephalitis antigenic complex.* The American journal of tropical medicine and hygiene Am J Trop Med Hyg Am. J. Trop. Med. Hyg., 1998. **59**(6): p. 952-964. DOI: https://doi.org/10.4269/ajtmh.1998.59.952.
13. Kinney, R.M., et al., *Genetic evidence that epizootic Venezuelan equine encephalitis (VEE) viruses may have evolved from enzootic VEE subtype I-D*

virus. Virology, 1992. **191**(2): p. 569-580. DOI: https://doi.org/10.1016/0042-6822(92)90232-E.
14. Brault, A.C., et al., *Potential sources of the 1995 Venezuelan equine encephalitis subtype IC epidemic.* Journal of virology, 2001. **75**(13): p. 5823-5832. DOI: https://doi.org/10.1128/JVI.75.13.5823-5832.2001.
15. Oberste, M.S., M.D. Parker, and J.F. Smith, *Complete Sequence of Venezuelan Equine Encephalitis Virus Subtype IE Reveals Conserved and Hypervariable Domains within the C Terminus of nsP3.* Virology, 1996. **219**(1): p. 314-320. DOI: https://doi.org/10.1006/viro.1996.0254.
16. Wang, E., et al., *Reverse transcription-PCR-enzyme-linked immunosorbent assay for rapid detection and differentiation of alphavirus infections.* Journal of clinical microbiology, 2006. **44**(11): p. 4000-4008. DOI: https://doi.org/10.1128/JCM.00175-06.
17. Sudia, W.D., et al., *EPIDEMIC VENEZUELAN EQUINE ENCEPHALITIS IN NORTH AMERICA IN 1971: VECTOR STUDIES1.* American Journal of Epidemiology, 1975. **101**(1): p. 17-35. DOI: https://doi.org/10.1093/oxfordjournals.aje.a112068.
18. Turell, M.J., G.V. Ludwig, and J.R. Beaman, *Transmission of Venezuelan Equine Encephalomyelitis Virus by Aedes sollicitans and Aedes taeniorhynchus (Diptera: Culicidae).* Journal of Medical Entomology, 1992. **29**(1): p. 62-65. DOI: https://doi.org/10.1093/jmedent/29.1.62.
19. Ferro, C., et al., *Natural Enzootic Vectors of Venezuelan equine encephalitis virus in the Magdalena Valley, Colombia.* Emerging Infectious Disease journal, 2003. **9**(1): p. 49. DOI: https://doi.org/10.3201/eid0901.020136.
20. Carrara, A.-S., et al., *Venezuelan equine encephalitis virus infection of spiny rats.* Emerging infectious diseases, 2005. **11**(5): p. 663-669. DOI: https://doi.org/10.3201/eid1105.041251.
21. Carrara, A.-S., et al., *Venezuelan equine encephalitis virus infection of cotton rats.* Emerging infectious diseases, 2007. **13**(8): p. 1158-1165. DOI: https://doi.org/10.3201/eid1308.061157.
22. Guzmán, C., et al., *Eco-epidemiology of the Venezuelan equine encephalitis virus in bats of Córdoba and Sucre, Colombia.* Acta Tropica, 2019. **191**: p. 178-184. DOI: https://doi.org/10.1016/j.actatropica.2018.12.016.
23. Johnson, K.M., et al., *Recovery of Venezuelan Equine Encephalomyelitis Virus in Panam?: A Fatal Case in Man.* The American Journal of Tropical Medicine and Hygiene, 1968. **17**(3): p. 432-440. DOI: https://doi.org/10.4269/ajtmh.1968.17.432.
24. de la Monte, S.M., et al., *The Systemic Pathology of Venezuelan Equine Encephalitis Virus Infection in Humans.* The American Journal of Tropical Medicine and Hygiene, 1985. **34**(1): p. 194-202. DOI: https://doi.org/10.4269/ajtmh.1985.34.194.
25. Steele, K.E. and N.A. Twenhafel, *REVIEW PAPER: pathology of animal models of alphavirus encephalitis.* Vet Pathol, 2010. **47**(5): p. 790-805. DOI: https://doi.org/10.1177/0300985810372508.

26. MacDonald, G., H. and R. Johnston, E., *Role of Dendritic Cell Targeting in Venezuelan Equine Encephalitis Virus Pathogenesis.* Journal of Virology, 2000. **74**(2): p. 914-922. DOI: https://doi.org/10.1128/JVI.74.2.914-922.2000.
27. Charles, P.C.W., Eric; Margolis, Frank; Johnston, Robert E., *Mechanism of Neuroinvasion of Venezuelan Equine Encephalitis Virus in the Mouse.* Virology, 1995. **208**: p. 662-671. DOI: https://doi.org/10.1006/viro.1995.1197.
28. Steele, K.E., et al., *Comparative Neurovirulence and Tissue Tropism of Wild-type and Attenuated Strains of Venezuelan Equine Encephalitis Virus Administered by Aerosol in C3H/HeN and BALB/c Mice.* Veterinary Pathology, 1998. **35**(5): p. 386-397. DOI: https://doi.org/10.1177/030098589803500508.
29. Ryzhikov, A.B., et al., *Venezuelan equine encephalitis virus propagation in the olfactory tract of normal and immunized mice.* Biomed Sci, 1991. **2**(0955-9701 (Print)): p. 8. https://pubmed.ncbi.nlm.nih.gov/1841630/.
30. Jackson, A.C. and J.P. Rossiter, *Apoptotic cell death is an important cause of neuronal injury in experimental Venezuelan equine encephalitis virus infection of mice.* Acta Neuropathol, 1997. **93**: p. 349–353. DOI: https://doi.org/10.1007/s004010050626.
31. Schoneboom, B.A.C., Kristen M.K.; Marty, Aileen M.; Grieder, Franziska B. , *Inflammation is a component of neurodegeneration in response to Venezuelan equine encephalitis virus infection in mice.* J Neuroimmunol, 2000. **109**: p. 132–146. DOI: https://doi.org/10.1016/s0165-5728(00)00290-3.
32. Baer, A., et al., *Venezuelan Equine Encephalitis Virus Induces Apoptosis through the Unfolded Protein Response Activation of EGR1.* J Virol, 2016. **90**(7): p. 3558-72. DOI: https://doi.org/10.1128/JVI.02827-15.
33. Schafer, A., et al., *The role of the blood-brain barrier during Venezuelan equine encephalitis virus infection.* J Virol, 2011. **85**(20): p. 10682-90. DOI: https://doi.org/10.1128/JVI.05032-11.
34. Cain, M.D., et al., *Virus entry and replication in the brain precedes blood-brain barrier disruption during intranasal alphavirus infection.* J Neuroimmunol, 2017. **308**: p. 118-130. DOI: https://doi.org/10.1016/j.jneuroim.2017.04.008.
35. Phillips, A.T., et al., *Entry Sites of Venezuelan and Western Equine Encephalitis Viruses in the Mouse Central Nervous System following Peripheral Infection.* J Virol, 2016. **90**(12): p. 5785-96. DOI: https://doi.org/10.1128/JVI.03219-15.
36. Patterson, M., et al., *Rapid, non-invasive imaging of alphaviral brain infection: reducing animal numbers and morbidity to identify efficacy of potential vaccines and antivirals.* Vaccine, 2011. **29**(50): p. 9345-51. DOI: https://doi.org/10.1016/j.vaccine.2011.09.130.
37. Rupp, J.C., et al., *Alphavirus RNA synthesis and non-structural protein functions.* J Gen Virol, 2015. **96**(9): p. 2483-2500. DOI: https://doi.org/10.1099/jgv.0.000249.
38. Vasiljeva, L., et al., *Identification of a Novel Function of the AlphavirusCapping Apparatus: RNA 5'-TRIPHOSPHATASE ACTIVITY OF Nsp2*.* Journal of Biological Chemistry, 2000. **275**(23): p. 17281-17287. DOI: https://doi.org/10.1074/jbc.M910340199.

39. Kim, D.Y., et al., *Conservation of a Packaging Signal and the Viral Genome RNA Packaging Mechanism in Alphavirus Evolution.* Journal of Virology, 2011. **85**(16): p. 8022. DOI: https://doi.org/10.1128/JVI.00644-11.
40. Rupp, J.C., N. Jundt, and R.W. Hardy, *Requirement for the amino-terminal domain of sindbis virus nsP4 during virus infection.* J Virol, 2011. **85**(7): p. 3449-60. DOI: https://doi.org/10.1128/JVI.02058-10.
41. Tomar, S., et al., *Heterologous production, purification and characterization of enzymatically active Sindbis virus nonstructural protein nsP1.* Protein expression and purification, 2011. **79**(2): p. 277-284. DOI: https://doi.org/10.1016/j.pep.2011.05.022.
42. Gomez de Cedrón, M., et al., *RNA helicase activity of Semliki Forest virus replicase protein NSP2.* FEBS Letters, 1999. **448**(1): p. 19-22. DOI: https://doi.org/10.1016/S0014-5793(99)00321-X.
43. Hu, X., et al., *Kinetic, Mutational, and Structural Studies of the Venezuelan Equine Encephalitis Virus Nonstructural Protein 2 Cysteine Protease.* Biochemistry, 2016. **55**(21): p. 3007-19. DOI: https://doi.org/10.1021/acs.biochem.5b00992.
44. Russo, A.T., M.A. White, and S.J. Watowich, *The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease.* Structure, 2006. **14**(9): p. 1449-58. DOI: https://doi.org/10.1016/j.str.2006.07.010.
45. Atasheva, S., et al., *Development of Sindbis viruses encoding nsP2/GFP chimeric proteins and their application for studying nsP2 functioning.* J Virol, 2007. **81**(10): p. 5046-57. DOI: https://doi.org/10.1128/JVI.02746-06.
46. Kim, D.Y., et al., *Venezuelan equine encephalitis virus nsP2 protein regulates packaging of the viral genome into infectious virions.* J Virol, 2013. **87**(8): p. 4202-13. DOI: https://doi.org/10.1128/JVI.03142-12.
47. Garmashova, N., et al., *The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff.* Journal of virology, 2007. **81**(5): p. 2472-2484. DOI: https://doi.org/10.1128/JVI.02073-06.
48. Malet, H., et al., *The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket.* J Virol, 2009. **83**(13): p. 6534-45. DOI: https://doi.org/10.1128/JVI.00189-09.
49. Shin, G., et al., *Structural and functional insights into alphavirus polyprotein processing and pathogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(41): p. 16534-16539. DOI: https://doi.org/10.1073/pnas.1210418109.
50. Galbraith, S.E., B.J. Sheahan, and G.J. Atkins, *Deletions in the hypervariable domain of the nsP3 gene attenuate Semliki Forest virus virulence.* Journal of General Virology, 2006. **87**(4): p. 937-947. DOI: https://doi.org/10.1099/vir.0.81406-0.
51. Vihinen, H., et al., *Elimination of Phosphorylation Sites of Semliki Forest Virus Replicase Protein nsP3 *.* Journal of Biological Chemistry, 2001. **276**(8): p. 5745-5752. DOI: https://doi.org/10.1074/jbc.M006077200.
52. Foy, N.J., et al., *Hypervariable domain of nonstructural protein nsP3 of Venezuelan equine encephalitis virus determines cell-specific mode of virus*

replication. J Virol, 2013. **87**(13): p. 7569-84. DOI: https://doi.org/10.1128/JVI.00720-13.
53. Barton, D.J., S.G. Sawicki, and D.L. Sawicki, *Solubilization and immunoprecipitation of alphavirus replication complexes.* Journal of Virology, 1991. **65**(3): p. 1496. DOI: https://doi.org/10.1128/jvi.65.3.1496-1506.1991.
54. O'Reilly, E.K. and C.C. Kao, *Analysis of RNA-Dependent RNA Polymerase Structure and Function as Guided by Known Polymerase Structures and Computer Predictions of Secondary Structure.* Virology, 1998. **252**(2): p. 287-303. DOI: https://doi.org/10.1006/viro.1998.9463.
55. Tomar, S., et al., *Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity.* J Virol, 2006. **80**(20): p. 9962-9. DOI: https://doi.org/10.1128/JVI.01067-06.
56. Rubach, J.K., et al., *Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro.* Virology, 2009. **384**(1): p. 201-8. DOI: https://doi.org/10.1016/j.virol.2008.10.030.
57. Garmashova, N., et al., *Analysis of Venezuelan Equine Encephalitis Virus Capsid Protein Function in the Inhibition of Cellular Transcription.* Journal of Virology, 2007. **81**(24): p. 13552. DOI: https://doi.org/10.1128/JVI.01576-07.
58. Atasheva, S., et al., *Venezuelan Equine Encephalitis Virus Capsid Protein Forms a Tetrameric Complex with CRM1 and Importin α/β That Obstructs Nuclear Pore Complex Function.* Journal of Virology, 2010. **84**(9): p. 4158. DOI: https://doi.org/10.1128/JVI.02554-09.
59. Holmes, A.C., et al., *A molecular understanding of alphavirus entry.* PLoS Pathog, 2020. **16**(10): p. e1008876. DOI: https://doi.org/10.1371/journal.ppat.1008876.
60. Jose, J., J.E. Snyder, and R.J. Kuhn, *A structural and functional perspective of alphavirus replication and assembly.* Future Microbiology, 2009. **4**(7): p. 837-856. DOI: https://doi.org/10.2217/fmb.09.59.
61. de Curtis, I. and K. Simons, *Dissection of Semliki Forest virus glycoprotein delivery from the trans-Golgi network to the cell surface in permeabilized BHK cells.* Proceedings of the National Academy of Sciences, 1988. **85**(21): p. 8052. DOI: https://doi.org/10.1073/pnas.85.21.8052.
62. Uchime, O., W. Fields, and M. Kielian, *The Role of E3 in pH Protection during Alphavirus Assembly and Exit.* Journal of Virology, 2013. **87**(18): p. 10255-10262. DOI: https://doi.org/doi:10.1128/JVI.01507-13.
63. Zhang, R., et al., *4.4 A cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus.* EMBO J, 2011. **30**(18): p. 3854-63. DOI: https://doi.org/10.1038/emboj.2011.261.
64. Yao, J.S., E.G. Strauss, and J.H. Strauss, *Interactions between PE2, E1, and 6K required for assembly of alphaviruses studied with chimeric viruses.* Journal of virology, 1996. **70**(11): p. 7910-7920. DOI: https://doi.org/10.1128/JVI.70.11.7910-7920.1996.
65. Snyder, J.E., et al., *Functional Characterization of the Alphavirus TF Protein.* Journal of Virology, 2013. **87**(15): p. 8511-8523. DOI: https://doi.org/doi:10.1128/JVI.00449-13.

66. Firth, A.E., et al., *Discovery of frameshifting in Alphavirus 6K resolves a 20-year enigma.* Virology Journal, 2008. **5**(1): p. 108. DOI: https://doi.org/10.1186/1743-422X-5-108.
67. Kendra, J., A., et al., *Ablation of Programmed −1 Ribosomal Frameshifting in Venezuelan Equine Encephalitis Virus Results in Attenuated Neuropathogenicity.* Journal of Virology, 2017. **91**(3): p. e01766-16. DOI: https://doi.org/10.1128/JVI.01766-16.
68. Lusa, S., H. Garoff, and P. Liueström, *Fate of the 6K membrane protein of semliki forest virus during virus assembly.* Virology, 1991. **185**(2): p. 843-846. DOI: https://doi.org/10.1016/0042-6822(91)90556-Q.
69. Nickens, D.G. and R.W. Hardy, *Structural and functional analyses of stem–loop 1 of the Sindbis virus genome.* Virology, 2008. **370**(1): p. 158-172. DOI: https://doi.org/10.1016/j.virol.2007.08.006.
70. Pfeffer, M., R.M. Kinney, and O.-R. Kaaden, *The Alphavirus 3'-Nontranslated Region: Size Heterogeneity and Arrangement of Repeated Sequence Elements.* Virology, 1998. **240**(1): p. 100-108. DOI: https://doi.org/10.1006/viro.1997.8907.
71. Hardy, R.W. and C.M. Rice, *Requirements at the 3' end of the sindbis virus genome for efficient synthesis of minus-strand RNA.* Journal of virology, 2005. **79**(8): p. 4630-4639. DOI: https://doi.org/10.1128/JVI.79.8.4630-4639.2005.
72. Levis, R., S. Schlesinger, and H.V. Huang, *Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription.* Journal of virology, 1990. **64**(4): p. 1726-1733. DOI: https://doi.org/10.1128/JVI.64.4.1726-1733.1990.
73. Kolokoltsov, A.A., E.H. Fleming, and R.A. Davey, *Venezuelan equine encephalitis virus entry mechanism requires late endosome formation and resists cell membrane cholesterol depletion.* Virology, 2006. **347**(2): p. 333-42. DOI: https://doi.org/10.1016/j.virol.2005.11.051.
74. Ma, H., et al., *LDLRAD3 is a receptor for Venezuelan equine encephalitis virus.* Nature, 2020. **588**(7837): p. 308-314. DOI: https://doi.org/10.1038/s41586-020-2915-3.
75. Malygin, A.A., et al., *C-terminal fragment of human laminin-binding protein contains a receptor domain for Venezuelan equine encephalitis and tick-borne encephalitis viruses.* Biochemistry (Moscow), 2009. **74**(12): p. 1328-1336. DOI: https://doi.org/10.1134/S0006297909120050.
76. Bernard, K.A., W.B. Klimstra, and R.E. Johnston, *Mutations in the E2 Glycoprotein of Venezuelan Equine Encephalitis Virus Confer Heparan Sulfate Interaction, Low Morbidity, and Rapid Clearance from Blood of Mice.* Virology, 2000. **276**(1): p. 93-103. DOI: https://doi.org/10.1006/viro.2000.0546.
77. Klimstra, W.B., K.D. Ryman, and R.E. Johnston, *Adaptation of Sindbis Virus to BHK Cells Selects for Use of Heparan Sulfate as an Attachment Receptor.* Journal of Virology, 1998. **72**(9): p. 7357. DOI: https://doi.org/10.1128/JVI.72.9.7357-7366.1998.
78. Helenius , A., et al., *On the entry of semliki forest virus into BHK-21 cells.* Journal of Cell Biology, 1980. **84**(2): p. 404-420. DOI: https://doi.org/10.1083/jcb.84.2.404.
79. Li, G. and C.M. Rice, *The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream*

of the termination codon. Journal of virology, 1993. **67**(8): p. 5062-5067. DOI: https://doi.org/10.1128/JVI.67.8.5062-5067.1993.
80. de Groot, R.J., et al., *Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo.* The EMBO journal, 1990. **9**(8): p. 2631-2638. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC552296/.
81. Shirako, Y. and J.H. Strauss, *Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis, whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis.* Journal of virology, 1994. **68**(3): p. 1874-1885. DOI: https://doi.org/10.1128/JVI.68.3.1874-1885.1994.
82. Frolova, E.I., et al., *Functional Sindbis virus replicative complexes are formed at the plasma membrane.* Journal of virology, 2010. **84**(22): p. 11679-11695. DOI: https://doi.org/10.1128/JVI.01441-10.
83. Spuul, P., et al., *Phosphatidylinositol 3-kinase-, actin-, and microtubule-dependent transport of Semliki Forest Virus replication complexes from the plasma membrane to modified lysosomes.* Journal of virology, 2010. **84**(15): p. 7543-7557. DOI: https://doi.org/10.1128/JVI.00477-10.
84. Melancon, P. and H. Garoff, *Processing of the Semliki Forest virus structural polyprotein: role of the capsid protease.* Journal of virology, 1987. **61**(5): p. 1301-1309. DOI: https://doi.org/10.1128/JVI.61.5.1301-1309.1987.
85. von Bonsdorff, C.H. and S.C. Harrison, *Sindbis virus glycoproteins form a regular icosahedral surface lattice.* Journal of virology, 1975. **16**(1): p. 141-145. DOI: https://doi.org/10.1128/JVI.16.1.141-145.1975.
86. Acheson, N.H. and I. Tamm, *Replication of semliki forest virus: An electron microscopic study.* Virology, 1967. **32**(1): p. 128-143. DOI: https://doi.org/10.1016/0042-6822(67)90261-9.
87. Tang, J., et al., *Molecular links between the E2 envelope glycoprotein and nucleocapsid core in Sindbis virus.* Journal of molecular biology, 2011. **414**(3): p. 442-459. DOI: https://doi.org/10.1016/j.jmb.2011.09.045.
88. Pavan, A., et al., *Regional distribution of Sindbis virus glycoproteins on the plasma membrane of infected baby hamster kidney cells.* Experimental Cell Research, 1987. **168**(1): p. 53-62. DOI: https://doi.org/10.1016/0014-4827(87)90415-0.
89. Lu, Y.E., et al., *In Vivo Generation and Characterization of a Soluble Form of the Semliki Forest Virus Fusion Protein.* Journal of Virology, 2001. **75**(17): p. 8329-8339. DOI: https://doi.org/10.1128/JVI.75.17.8329-8339.2001.
90. Taylor, G.M., P.I. Hanson, and M. Kielian, *Ubiquitin depletion and dominant-negative VPS4 inhibit rhabdovirus budding without affecting alphavirus budding.* Journal of virology, 2007. **81**(24): p. 13631-13639. DOI: https://doi.org/10.1128/JVI.01688-07.
91. Wilkinson, T.A., et al., *Association of Sindbis Virus Capsid Protein with Phospholipid Membranes and the E2 Glycoprotein: Implications for Alphavirus Assembly.* Biochemistry, 2005. **44**(8): p. 2800-2810. DOI: https://doi.org/10.1021/bi0479961.

92. Martinez, M.G. and M. Kielian, *Intercellular Extensions Are Induced by the Alphavirus Structural Proteins and Mediate Virus Transmission.* PLOS Pathogens, 2016. **12**(12): p. e1006061. DOI: https://doi.org/10.1371/journal.ppat.1006061.
93. Sanjuán, R., et al., *Viral mutation rates.* Journal of virology, 2010. **84**(19): p. 9733-9748. DOI: https://doi.org/10.1128/JVI.00694-10.
94. Peck, K.M., A.S. Lauring, and C.S. Sullivan, *Complexities of Viral Mutation Rates.* Journal of Virology, 2018. **92**(14): p. e01031-17. DOI: https://doi.org/10.1128/JVI.01031-17.
95. Patterson, E.I., et al., *Measuring Alphavirus Fidelity Using Non-Infectious Virus Particles.* Viruses, 2020. **12**(5): p. 546. DOI: https://doi.org/10.3390/v12050546.
96. Scherer, W.F., C.A. Ellsworth, and A.K. Ventura, *Studies of viral virulence. II. Growth and adsorption curves of virulent and attenuated strains of Venezuelan encephalitis virus in cultured cells.* The American journal of pathology, 1971. **62**(2): p. 211-219. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2047534/.
97. Lauring, A.S. and R. Andino, *Quasispecies theory and the behavior of RNA viruses.* PLoS Pathog, 2010. **6**(7): p. e1001005. DOI: https://doi.org/10.1371/journal.ppat.1001005.
98. Lauring, A.S., J. Frydman, and R. Andino, *The role of mutational robustness in RNA virus evolution.* Nat Rev Microbiol, 2013. **11**(5): p. 327-36. DOI: https://doi.org/10.1038/nrmicro3003.
99. Grubaugh, N.D., et al., *Genetic Drift during Systemic Arbovirus Infection of Mosquito Vectors Leads to Decreased Relative Fitness during Host Switching.* Cell Host Microbe, 2016. **19**(4): p. 481-92. DOI: https://doi.org/10.1016/j.chom.2016.03.002.
100. Sim, S., et al., *Tracking Dengue Virus Intra-host Genetic Diversity during Human-to-Mosquito Transmission.* PLoS Negl Trop Dis, 2015. **9**(9): p. e0004052. DOI: https://doi.org/10.1371/journal.pntd.0004052.
101. Vignuzzi, M., et al., *Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population.* Nature, 2006. **439**(7074): p. 344-8. DOI: https://doi.org/10.1038/nature04388.
102. Warmbrod, K.L., et al., *Viral RNA-dependent RNA polymerase mutants display an altered mutation spectrum resulting in attenuation in both mosquito and vertebrate hosts.* PLoS Pathog, 2019. **15**(4): p. e1007610. DOI: https://doi.org/10.1371/journal.ppat.1007610.
103. Kautz, T.F., et al., *Low-fidelity Venezuelan equine encephalitis virus polymerase mutants to improve live-attenuated vaccine safety and efficacy.* Virus Evol, 2018. **4**(1): p. vey004. DOI: https://doi.org/10.1093/ve/vey004.
104. Wyles, D.L. and A.F. Luetkemeyer, *Understanding Hepatitis C Virus Drug Resistance: Clinical Implications for Current and Future Regimens.* Topics in antiviral medicine, 2017. **25**(3): p. 103-109. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5935211/.
105. Hussain, M., et al., *Drug resistance in influenza A virus: the epidemiology and management.* Infect Drug Resist, 2017. **10**: p. 121-134. DOI: https://doi.org/10.2147/IDR.S105473.

106. Chung, D.H., et al., *Discovery of a novel compound with anti-venezuelan equine encephalitis virus activity that targets the nonstructural protein 2.* PLoS Pathog, 2014. **10**(6): p. e1004213. DOI: https://doi.org/10.1371/journal.ppat.1004213.
107. Schroeder, C.E., et al., *Development of (E)-2-((1,4-dimethylpiperazin-2-ylidene)amino)-5-nitro-N-phenylbenzamide, ML336: Novel 2-amidinophenylbenzamides as potent inhibitors of venezuelan equine encephalitis virus.* J Med Chem, 2014. **57**(20): p. 8608-21. DOI: https://doi.org/10.1021/jm501203v.
108. Markland, W., et al., *Broad-Spectrum Antiviral Activity of the IMP Dehydrogenase Inhibitor VX-497: a Comparison with Ribavirin and Demonstration of Antiviral Additivity with Alpha Interferon.* Antimicrobial Agents and Chemotherapy, 2000. **44**(4): p. 8. DOI: https://doi.org/10.1128/aac.44.4.859-866.2000.
109. Julander, J.G., et al., *Treatment of Venezuelan equine encephalitis virus infection with (-)-carbodine.* Antiviral Res, 2008. **80**(3): p. 309-15. DOI: https://doi.org/10.1016/j.antiviral.2008.07.002.
110. Kehn-Hall, K., et al., *Modulation of GSK-3beta activity in Venezuelan equine encephalitis virus infection.* PLoS One, 2012. **7**(4): p. e34761. DOI: https://doi.org/10.1371/journal.pone.0034761.
111. Madsen, C., et al., *Small molecule inhibitors of Ago2 decrease Venezuelan equine encephalitis virus replication.* Antiviral Res, 2014. **112**: p. 26-37. DOI: https://doi.org/10.1016/j.antiviral.2014.10.002.
112. Delang, L., et al., *Mutations in the chikungunya virus non-structural proteins cause resistance to favipiravir (T-705), a broad-spectrum antiviral.* J Antimicrob Chemother, 2014. **69**(10): p. 2770-84. DOI: https://doi.org/10.1093/jac/dku209.
113. Urakova, N., et al., *β-d-N (4)-Hydroxycytidine Is a Potent Anti-alphavirus Compound That Induces a High Level of Mutations in the Viral Genome.* Journal of virology, 2018. **92**(3): p. e01965-17. DOI: https://doi.org/10.1128/JVI.01965-17.
114. Lundberg, L., et al., *Repurposed FDA-Approved drug sorafenib reduces replication of Venezuelan equine encephalitis virus and other alphaviruses.* Antiviral Res, 2018. **157**: p. 57-67. DOI: https://doi.org/10.1016/j.antiviral.2018.07.005.
115. Lundberg, L., et al., *Nuclear import and export inhibitors alter capsid protein distribution in mammalian cells and reduce Venezuelan Equine Encephalitis Virus replication.* Antiviral Res, 2013. **100**(3): p. 662-72. DOI: https://doi.org/10.1016/j.antiviral.2013.10.004.
116. DeBono, A., et al., *Novel RU486 (mifepristone) analogues with increased activity against Venezuelan Equine Encephalitis Virus but reduced progesterone receptor antagonistic activity.* Sci Rep, 2019. **9**(1): p. 2634. DOI: https://doi.org/10.1038/s41598-019-38671-y.
117. Jonsson, C.B., et al., *Efficacy of a ML336 derivative against Venezuelan and eastern equine encephalitis viruses.* Antiviral Res, 2019. **167**: p. 25-34. DOI: https://doi.org/10.1016/j.antiviral.2019.04.004.
118. Skidmore, A.M., et al., *Benzamidine ML336 inhibits plus and minus strand RNA synthesis of Venezuelan equine encephalitis virus without affecting host RNA*

production. Antiviral Res, 2020. **174**: p. 104674. DOI: https://doi.org/10.1016/j.antiviral.2019.104674.
119. Emeny, J.M.M., Michael J., *Regulation of the Interferon System: Evidence that Vero Cells have a Genetic Defect in Interferon Production.* J Gen Virol., 1979. **43**(1): p. 247-52. DOI: https://doi.org/10.1099/0022-1317-43-1-247.
120. Lee, J., et al., *Emergence and Magnitude of ML336 Resistance in Venezuelan Equine Encephalitis Virus Depend on the Microenvironment.* Journal of Virology, 2020. **94**(22): p. e00317-20. DOI: https://doi.org/10.1128/JVI.00317-20.
121. De Clercq, E. and G. Li, *Approved Antiviral Drugs over the Past 50 Years.* Clin Microbiol Rev, 2016. **29**(3): p. 695-747. DOI: https://doi.org/10.1128/CMR.00102-15.
122. Drake, J.W. and J.J. Holland, *Mutation rates among RNA viruses.* Proceedings of the National Academy of Sciences, 1999. **96**(24): p. 13910. DOI: https://doi.org/10.1073/pnas.96.24.13910.
123. Irwin, K.K., et al., *Antiviral drug resistance as an adaptive process.* Virus Evol, 2016. **2**(1): p. vew014. DOI: https://doi.org/10.1093/ve/vew014.
124. Farrukee, R., et al., *Characterization of substitutions in the neuraminidase of A(H7N9) influenza viruses selected following serial passage in the presence of different neuraminidase inhibitors.* Antiviral Res, 2019. **168**: p. 68-75. DOI: https://doi.org/10.1016/j.antiviral.2019.05.009.
125. Xu, S., et al., *In vitro selection of resistance to sofosbuvir in HCV replicons of genotype-1 to -6.* Antiviral therapy, 2017. **22**(7): p. 587-597. DOI: https://doi.org/10.3851/imp3149.
126. Donaldson, E.F., et al., *Clinical evidence and bioinformatics characterization of potential hepatitis C virus resistance pathways for sofosbuvir.* Hepatology, 2015. **61**(1): p. 56-65. DOI: https://doi.org/10.1002/hep.27375.
127. Renzette, N., et al., *Evolution of the influenza A virus genome during development of oseltamivir resistance in vitro.* J Virol, 2014. **88**(1): p. 272-81. DOI: https://doi.org/10.1128/JVI.01067-13.
128. Foll, M., et al., *Influenza virus drug resistance: a time-sampled population genetics perspective.* PLoS Genet, 2014. **10**(2): p. e1004185. DOI: https://doi.org/10.1371/journal.pgen.1004185.
129. Reed, L.J.M., H., *A Simple Method of Estimating Fifty Per Cent Endpoints.* Am J Hyg, 1938. **27**: p. 493–497. DOI: https://doi.org/10.1016/j.jviromet.2005.05.005.
130. Wang, C., et al., *Characterization of mutation spectra with ultra-deep pyrosequencing: application to HIV-1 drug resistance.* Genome Res, 2007. **17**(8): p. 1195-201. DOI: https://doi.org/10.1101/gr.6468307.
131. Ladner, J.T., et al., *Standards for sequencing viral genomes in the era of high-throughput sequencing.* MBio, 2014. **5**(3): p. e01360-14. DOI: https://doi.org/10.1128/mBio.01360-14.
132. Julander, J.G., et al., *C3H/HeN mouse model for the evaluation of antiviral agents for the treatment of Venezuelan equine encephalitis virus infection.* Antiviral Research, 2008. **78**(3): p. 230-241. DOI: https://doi.org/10.1016/j.antiviral.2008.01.007.

133. Leigh, J.W., D. Bryant, and S. Nakagawa, *PopART: Full-feature software for haplotype network construction.* Methods in Ecology and Evolution, 2015. **6**(9): p. 1110-1116. DOI: https://doi.org/10.1111/2041-210x.12410.
134. Paessler, S., et al., *Alpha-beta T cells provide protection against lethal encephalitis in the murine model of VEEV infection.* Virology, 2007. **367**(2): p. 307-23. DOI: https://doi.org/10.1016/j.virol.2007.05.041.
135. Kinney, R.M., et al., *Attenuation of Venezuelan equine encephalitis virus strain TC-83 is encoded by the 5'-noncoding region and the E2 envelope glycoprotein.* Journal of Virology, 1993. **67**(3): p. 1269-1277. DOI: https://doi.org/10.1128/jvi.67.3.1269-1277.1993.
136. Ludwig, G.V., et al., *Comparative neurovirulence of attenuated and non-attenuated strains of Venezuelan equine encephalitis virus in mice.* The American journal of tropical medicine and hygiene Am J Trop Med Hyg Am. J. Trop. Med. Hyg., 2001. **64**(1): p. 49-55. DOI: https://doi.org/10.4269/ajtmh.2001.64.49.
137. Kutchko, K.M., et al., *Structural divergence creates new functional features in alphavirus genomes.* Nucleic Acids Res, 2018. **46**(7): p. 3657-3670. DOI: https://doi.org/10.1093/nar/gky012.
138. LaPointe, A.T. and K.J. Sokoloski, *De-Coding the Contributions of the Viral RNAs to Alphaviral Pathogenesis.* Pathogens, 2021. **10**(6). DOI: https://doi.org/10.3390/pathogens10060771.
139. Orton, R.J., et al., *Distinguishing low frequency mutations from RT-PCR and sequence errors in viral deep sequencing data.* BMC Genomics, 2015. **16**: p. 229. DOI: https://doi.org/10.1186/s12864-015-1456-x.
140. Sreejith, R., et al., *Mapping interactions of Chikungunya virus nonstructural proteins.* Virus Res, 2012. **169**(1): p. 231-6. DOI: https://doi.org/10.1016/j.virusres.2012.08.006.
141. Tan, Y.B., et al., *A crystal structure of alphavirus nonstructural protein 4 (nsP4) reveals an intrinsically dynamic RNA-dependent RNA polymerase.* bioRxiv, 2021: p. 2021.05.27.445971. DOI: https://doi.org/10.1101/2021.05.27.445971.
142. Piccininni, S., et al., *Modulation of the Hepatitis C Virus RNA-dependent RNA Polymerase Activity by the Non-Structural (NS) 3 Helicase and the NS4B Membrane Protein* *. Journal of Biological Chemistry, 2002. **277**(47): p. 45670-45679. DOI: https://doi.org/10.1074/jbc.M204124200.
143. Chen, J., et al., *Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex.* Cell, 2020. **182**(6): p. 1560-1573.e13. DOI: https://doi.org/10.1016/j.cell.2020.07.033.
144. Stapleford, K., A., et al., *Viral Polymerase-Helicase Complexes Regulate Replication Fidelity To Overcome Intracellular Nucleotide Depletion.* Journal of Virology, 2015. **89**(22): p. 11233-11244. DOI: https://doi.org/10.1128/JVI.01553-15.
145. Levi, L.I., et al., *Defective viral genomes from chikungunya virus are broad-spectrum antivirals and prevent virus dissemination in mosquitoes.* PLOS Pathogens, 2021. **17**(2): p. e1009110. DOI: https://doi.org/10.1371/journal.ppat.1009110.
146. Hughes, H.R., et al., *Fatal Human Infection with Evidence of Intrahost Variation of Eastern Equine Encephalitis Virus, Alabama, USA, 2019.* Emerging infectious

diseases, 2021. **27**(7): p. 1886-1892. DOI: https://doi.org/10.3201/eid2707.210315.
147. Forrester, N.L., et al., *Vector-borne transmission imposes a severe bottleneck on an RNA virus population.* PLoS Pathog, 2012. **8**(9): p. e1002897. DOI: https://doi.org/10.1371/journal.ppat.1002897.

VITA

Jasper Lee was born to Mouh-Wahng and Wenyin Lee in Knoxville, Tennessee in 1992. He graduated from Solon High School of Solon, Ohio, in 2010, then obtained a Bachelor of Science degree in Microbiology at The Ohio State University in 2014. In 2015, he entered the Graduate Program in Microbiology at the University of Tennessee Knoxville, where he joined Dr. Colleen Jonsson's laboratory. In 2017, he moved with Dr. Jonsson to the University of Tennessee Health Science Center in Memphis, Tennessee, where he transferred into the Integrated Biomedical Sciences PhD program in the College of Graduate Health Sciences. He expects to complete his doctoral degree by December of 2021.