

ABSTRACT

Title of Dissertation:

PROGRAMMED TRANSLATIONAL
RECODING SIGNALS AS A
THERAPEUTIC TARGET AGAINST
ALPHAVIRUSES

Joseph Aaron Kendra
Doctor of Philosophy
2018

Dissertation directed by:

Professor and Chair, Jonathan D. Dinman,
Department of Cell Biology and Molecular
Genetics

While infection from communicable diseases has posed a longstanding threat to human health throughout history, the modern realities of population expansion, global travel, and climate change have facilitated the rapid emergence and worldwide distribution of RNA viruses at an unprecedented scale. Of particular concern are the alphaviruses, mosquito borne viruses from the *Togaviridae* family. These viruses were previously relegated to rare outbreaks in isolated forested regions but have dramatically spread across the globe in the past decade. One of these viruses, Venezuelan equine encephalitis virus (VEEV), is a noted bioterror threat due to its ability for aerosol transmission and successful weaponization during the Cold War.

While no FDA approved drugs exist against alphaviruses, their reliance on programmed translational recoding mechanisms to regulate gene expression presents a potential vulnerability for therapeutic exploitation. Two instances of translational recoding have been identified but poorly characterized in the alphavirus genome. The first is a termination codon readthrough (TCR) event required for expression of the alphavirus replicase. The second is a programmed -1 ribosomal frameshift (-1 PRF) that produces a C-terminally extended variant of viroporin 6K. In this work, the *cis*-acting RNA elements that mitigate alphavirus recoding were functionally and structurally characterized. The predicted TCR and -1 PRF sequences were cloned into dual luciferase reporter vectors and their ability to promote efficient recoding was verified in several mammalian cell lines. Chemical probing assays elucidated the presence of highly structured stemloop elements downstream of the alphavirus recoding sites, which function as a kinetic trap for elongating ribosomes. Notably, mutations that abrogate efficient -1 PRF not only attenuated pathogenesis of VEEV in mice, but also provided protective immunity to subsequent wild-type challenge. These findings suggest a novel approach to the development of a safe and effective live attenuated vaccine strategy against VEEV, closely related alphaviruses, and potentially all viruses that rely on translational recoding mechanisms for optimal gene expression.

PROGRAMMED TRANSLATIONAL RECODING SIGNALS AS A
THERAPEUTIC TARGET AGAINST ALPHAVIRUSES

by

Joseph Aaron Kendra

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2018

Advisory Committee:

Professor Jonathan D. Dinman, Chair

Professor Antony M. Jose

Professor James N. Culver

Professor Jeffrey J. DeStefano

Professor George A. Belov, Dean's Representative

© Copyright by
Joseph Aaron Kendra
2018

Foreword

This thesis uses manuscripts that were previously published in the Journal of Virology and submitted to the Journal of Biological Chemistry, respectively. As first author of both manuscripts, I certify that I was a substantial contributor to the experimental design, data collection, analyses and writing of each work.

Dedication

I dedicate this work to my parents with all of my love. Thank you for everything along the way.

Acknowledgements

I would first and foremost like to extend my gratitude to my mentor, Dr. Jonathan Dinman. Under his tutelage these past years I have grown tremendously both as a scientist and as a person. Wherever life takes me from here, I will remain indebted to his knowledge, his guidance, and especially his patience.

I would also like to thank the members of my committee -Dr. Antony Jose, Dr. James Culver, Dr. Jeffrey DeStefano and Dr. George Belov- for taking the time to sit through many arduous progress meetings throughout my time here. The quality of my work would simply not be where it is without your advice and I thank you all for that.

To the past and present members of the Dinman lab, especially Dr. Joe Briggs, Dr. John Jones, Dr. Bin Chen, Dr. Denis Wafula, Dr. Vivek Advani, Dr. Carol Vieira, Dr. Alicia Cheek, Dr. Sharmishtha Musalgaonkar, Dr. Ryan Kobylarz, Dr. Ashton Trey Belew and Alex Olson: thank you for being wonderful colleagues. My time at the University of Maryland was richer for having known you, and I wish you all the very best in all future endeavors.

To my research collaborators, particularly those from the laboratories of Dr. Kylene Kehn-Hall at George Mason University, MRIGlobal, and the UMD FIRE program: thank you for lending your expertise to the manuscripts featured in this dissertation. This could not have happened without your support.

Special thanks to the UMD Virology Program, whose training grant funded my research for many years and presented numerous opportunities to present my work at the American Society for Virology conferences. It cannot be overstated how important that was for my time here.

To all the friends I met at the University of Maryland: thank you for all the nights of grading, brainstorming, cooking, commiserating and karaoke (activities not necessarily mutually exclusive). May we go on to have many others.

To the One Frame Link crew and the rest of my FGC family: thank you for giving me a competitive outlet and helping me hold on to a sense of purpose and accomplishment when the going got rough in grad school. There are far too many people to name in this regard, but I sincerely thank you all.

To my parents, John and Sylvia Kendra, and my brother Gabe: I thank you for being a constant source of love and support for my entire life. I am the person I am today because of you all, and it's the reason I even made it to grad school in the first place. Moreover, I am thankful that both of my parents are survivors of graduate school themselves, and thus had some advice to give on how to roll with the punches.

Lastly, I want to acknowledge my partner, Randi Turner, who has been an unending source of inspiration and support for me. You are my best friend and companion and I owe much more to you than I could ever hope to put into words.

The works in this dissertation were partially funded through grants from the Defense Threat Reduction Agency grant (HDTRA1-13-1-0005) and the National Institutes of Health/National Institutes of General Medical Sciences (R01 GM117177). Additional funding awarded through the University of Maryland Virology Program was supported by NIH institutional training grant 2T32AI051967-06A1.

Table of Contents

Foreword.....	ii
Dedication	iii
Acknowledgements	iv
Table of Contents.....	vii
List of Tables.....	ix
List of Figures.....	x
List of Abbreviations	xii
Chapter 1: Introduction	15
Historical and Modern Context of Emerging Viral Diseases	15
The Alphaviruses as an Emerging Global Threat	23
Programmed Translational Recoding Signals in Viruses.....	33
Targeting Translational Recoding in Alphaviruses	53
Chapter 2: Ablation of programmed -1 frameshifting in Venezuelan equine encephalitis virus results in attenuated neuropathogenicity ¹⁵⁰	56
Abstract.....	57
Importance	58
Introduction	58
Materials and Methods	60
<i>Computational Prediction of Viral PRF Signals</i>	60
<i>Dual reporter plasmid construction and bacterial transformation</i>	61
<i>Cell Culture</i>	62
<i>Plasmid and siRNA transfection</i>	62
<i>Assays of programmed -1 ribosomal frameshifting</i>	63
<i>Chemical modification analysis of -1 PRF promoting RNA structural elements</i>	63
<i>3-dimensional structural modeling of the tVEEV -1 PRF stimulatory element</i>	64
<i>Introducing -1 PRF mutations into infectious VEEV clones</i>	64
<i>VEEV viral stocks</i>	65
<i>Analysis of viral kinetics</i>	66
<i>Animal experiments.....</i>	68
<i>Statistical analysis</i>	68
Results.....	69
<i>Alphavirus-derived sequences promote efficient levels of -1 PRF</i>	69
<i>Efficient alphavirus -1 PRF is stimulated by stem-loop mRNA structural elements.....</i>	72
<i>Ablation of -1 PRF mildly decreases VEEV virus production in vitro</i>	74
<i>Ablation of -1 PRF strongly attenuates VEEV pathogenesis</i>	76
Discussion.....	79
Acknowledgements.....	84

Chapter 3: Functional and structural characterization of the Chikungunya virus translational recoding signals	85
Abstract.....	86
Introduction	87
Materials and Methods	89
<i>Computational prediction of CHIKV recoding signals.....</i>	<i>89</i>
<i>Dual reporter plasmid construction.....</i>	<i>90</i>
<i>Cell Culture</i>	<i>90</i>
<i>Plasmid Transfections</i>	<i>91</i>
<i>Assays of translational recoding</i>	<i>91</i>
<i>Immunoblot Analyses.....</i>	<i>91</i>
<i>RNA structure analyses</i>	<i>92</i>
<i>Three-dimensional structural modeling of CHIKV -1 PRF and TCR stimulatory elements</i>	<i>94</i>
<i>Phylogenetic analyses</i>	<i>94</i>
<i>Statistical analyses</i>	<i>94</i>
Results.....	95
<i>The CHIKV translational recoding sequences are highly conserved</i>	<i>95</i>
<i>CHIKV-derived sequences promote efficient recoding in mammalian cell lines</i>	<i>97</i>
<i>Targeted mutations alter CHIKV recoding efficiencies</i>	<i>100</i>
<i>Chemical modification analyses identify complex stem-loop elements immediately 3' of CHIKV recoding sites.....</i>	<i>102</i>
Discussion.....	104
Acknowledgements	107
Chapter 4: Conclusions and future directions.....	109
Appendices.....	120
Appendix A. Experimental insert sequences.	120
(Table 1) 5' and 3' positions of the oligonucleotide sequence are indicated. Lowercase text denotes sequence overlap with plasmid backbone sequence. Uppercase text denotes the respective experimental sequence. Underlined regions indicate the position of the -1 PRF or TCR recoding site. Bold text denotes mutations from the wild type sequence. ..	120
Appendix B. Plasmids used in this work	126
(Table 2) List of plasmid names and descriptions that were used in this dissertation. ...	126
Appendix C. Supplementary Figures for Chapter 3	129
Bibliography	133

List of Tables

Table 1. Experimental insert sequences	120
Table 2. Plasmids used in this work.....	126

List of Figures

Figure 1. Communicable diseases remain the leading cause of death in low-income countries.	16
Figure 2. Cross-species transmission occurs more readily in RNA viruses than DNA viruses.	17
Figure 3. Worldwide presence of vector competent <i>Aedes</i> mosquitos.	21
Figure 4. Global proliferation of ZIKV is a recent phenomenon.	23
Figure 5. Alphavirus structural organization.	24
Figure 6. Global distribution of CHIKV.	26
Figure 7. Major outbreaks of VEEV in the Americas.	29
Figure 8. A single point mutation in VEEV TC-83 destabilizes a critical stemloop in the 5' UTR.	32
Figure 9. A multitude of proposed vaccine modalities against VEEV.	33
Figure 10. Initial discovery of -1 PRF in RSV.	35
Figure 11. Downstream stimulatory structure implicated in the efficient -1 PRF of RSV.	36
Figure 12. Biochemical evidence of a frameshifted gag-pol peptide in HIV-1.	38
Figure 13. Efficient HIV-1 -1 PRF occurs at a designated slippery site.	39
Figure 14. Canonical mechanism of -1 PRF.	40
Figure 15. Amber TCR is required to express gag-pol in some retroviruses.	41
Figure 16. Canonical mechanism of TCR.	43
Figure 17. Altering L-A -1 PRF efficiency diminishes propagation of M ₁ satellite virus.	45
Figure 18. Alterations to efficient -1 PRF diminish genomic and subgenomic RNA production in SARS-CoV.	47
Figure 19. Computational prediction of VEEV TCR signal.	50
Figure 20. Arginine substitution of the opal stop codon attenuates CHIKV infection in mice.	51
Figure 21. Discovery of alphavirus -1 PRF product TF.	53
Figure 22. Identification and monitoring of -1 PRF in the alphaviruses.	71
Figure 23. Structural analysis of alphaviral -1 PRF signals.	73
Figure 24. Ablation of -1 PRF decreases VEEV release <i>in vitro</i>	75
Figure 25. Ablation of -1 PRF strongly attenuates VEEV pathogenesis.	78
Figure 26. Identification and conservation analysis of CHIKV derived programmed translational recoding signals.	97
Figure 27. Characterization and validation of efficient CHIKV translational recoding in mammalian cell lines.	99
Figure 28. Mutagenesis-mediated destabilization of efficient CHIKV recoding.	101
Figure 29. Structural analyses of the CHIKV recoding signals.	103
Figure 30. VEEV TrD _{PRFm} induces high levels of neutralizing antibodies and protects mice from VEEV TrD infection.	110
Figure 31. pSGDluc corrects potential sources of confounding distortions from the pLuci reporter vector.	114

Figure 32. Efficient -1 PRF promoted by CHIKV and VEEV signals in bi-fluorescent (BiFL) reporter vectors.....	116
Figure 33. Structural analyses of the EEEV, VEEV and WEEV TCR signals.	117
Figure 34. Efficient translational recoding documented in other RNA viruses.	119
Figure 35. Multiple sequence alignments of CHIKV TCR and -1 PRF sequences.....	129
Figure 36. Supplemental information for Figure 27. Total protein of samples separated through 4-20% SDS-PAGE gel.	130
Figure 37. Supplemental information for Figure 28. Total protein of samples separated through 4-20% SDS-PAGE gels.	131

List of Abbreviations

-1 PRF	Programmed -1 ribosomal frameshifting
16S	Svedberg Unit
3'ter	3' premature termination codon
5'ter	5' premature termination codon
ABSL-3	Animal biosafety level 3
acGFP	<i>Aequorea coerulescens</i> green fluorescent protein
Af/As	Africa/Asia consensus sequence for CHIKV
Ago	Argonaute
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BaEV	Baboon endogenous retrovirus
BiFL	Bifluorescent vector
bp	Base pairs
BSA	Bovine serum albumin
BYDV	Barley yellow dwarf virus
°C	Celsius
Ca	Calcium
Carib	Caribbean consensus sequence for CHIKV
CCR5	C-C Chemokine receptor type 5
CDC	Centers for Disease Control
cDNA	Complementary DNA
CHIKV	Chikungunya virus
cm	Centimeter
CO ₂	Carbon dioxide
CoV	Coronavirus
Ct	Threshold cycle
Cys	Cysteine
DENV	Dengue virus
DMEM	Dulbecco's modified Eagle medium
DNA	Deoxyribonucleic acid
dpi	Days post-infection
DTRA	Defense Threat Reduction Agency
EEEV	Eastern equine encephalitis virus
Em	Emission
EMEM	Eagles minimum essential medium
eRF	Eukaryotic release factor
Ex	Excitation
Exp5	Exportin 5
FBS	Fetal bovine serum
FelV	Feline leukemia virus
gag	"Group-antigens" (Retrovirus capsid protein)
GALV	Gibbon ape leukemia virus

GMU	George Mason University
GNI	Gross national income
gRNA	Genomic RNA
GTP	Guanosine triphosphate
H	Hydrogen
HBSS	Hank's Balanced Salt Solution
HIV-1	Human Immunodeficiency Virus
hpe	Hours post electroporation
hpi	Hours post-infection
IACUC	Institutional animal care and use committee
IgG	Immunoglobulin G
kcal	Kilocalorie
kD	Kilo-Dalton
MaP	Mutational profiling
MAYV	Mayaro virus
MEM	Minimal essential medium
Met	Methionine
Mg	Magnesium
mg	Milligram
miRNA	MicroRNA
ml	Milliliter
mm	Millimeter
mM	Millimolar
MOI	Multiplicity of Infection
mol	Mole
mRNA	Messenger RNA
MuLV	Murine leukemia virus
N	Nitrogen
NIH	National Institutes of Health
nm	Nanometer
NMIA	N-methylisatoic anhydride
nPH85	Penny and Hendy metric
NS1	Nonstructural gene 1 (Flavivirus)
nsP	Nonstructural protein (Alphavirus disambiguation)
nt	Nucleotide
ONNV	O'Nyong'Nyong virus
ORF	Open reading frame
P	Phosphorus
PAGE	Poly-acrylamide gel electrophoresis
PBS	Phosphate buffered saline
PBST	1x PBS, 0.5% Tween20
PCR	Polymerase chain reaction
PEMV	Pea enation mosaic virus
PFU	Plaque-forming unit
PLB	Passive lysis buffer
pmol	Picomole

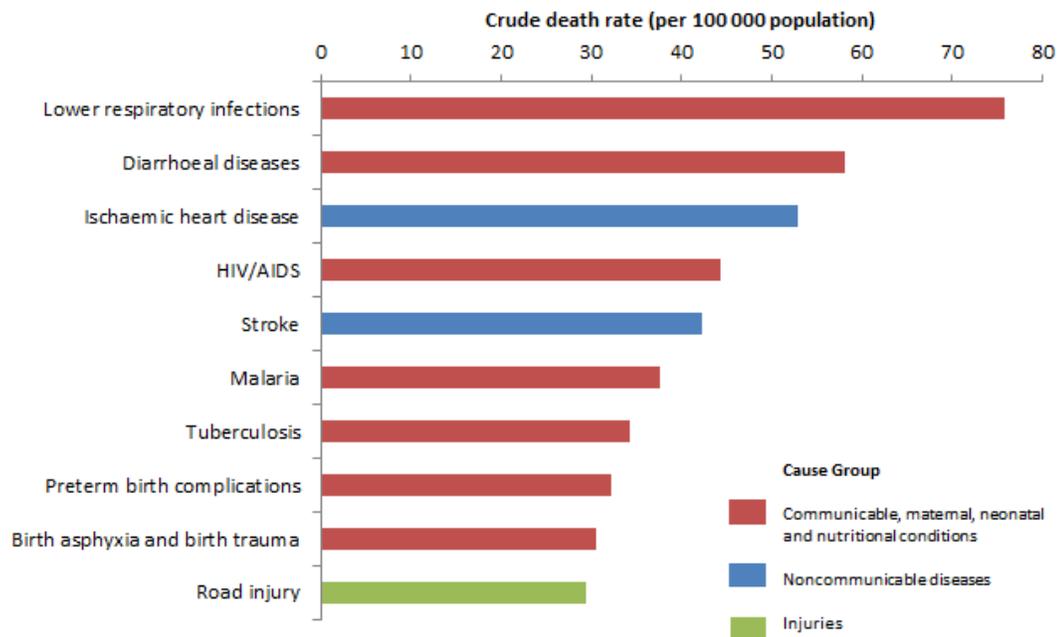
pol	“Polymerase” (Retrovirus reverse transcriptase)
PRFm	PRF mutant
qRT-PCR	Quantitative reverse transcription PCR
RCNMV	Red clover necrotic mosaic virus
rmp	Rotations per minute
RNA	Ribonucleic acid
RSV	Rous-sarcoma virus
S	Sulfur
SARS	Sudden Acute Respiratory Syndrome
Scr	Scrambled
SDS	Sodium dodecyl sulfate
SFV	Semliki Forest virus
sgRNA	Sub-genomic RNA
SHAPE	Selective 2' hydroxyl acylation and primer extension
SINV	Sinbis virus
siRNA	Small interfering RNA
ssM	Slip-site mutant
ssRNA	Single-strandedRNA
STLV	Simian T-cell leukemia virus
t2EEEV	Secondary truncation of EEEV -1 PRF
TCR	Termination codon readthrough
TCV	Turnip crinkle virus
tEEEV	Truncated EEEV -1 PRF signal
TF	<i>Trans</i> -Frame (Alphavirus)
TrD	Trinidad Donkey (VEEV)
tRNA	Transfer RNA
tVEEV	Truncated VEEV -1 PRF signal
tWEEV	Truncated WEEV -1 PRF signal
UMD	University of Maryland
USUV	Usutu virus
UTR	Untranslated region
VEEV	Venezuelan equine encephalitis virus
vRNA	Viral RNA
VRP	Virus replicon particles
WEEV	Western equine encephalitis virus
WHO	World Health Organization
WNV	West Nile Virus
WT	Wild-type
ZIKV	Zika virus
μF	Microfarad
μg	Microgram
μl	Microliter
μM	Micromolar
Ω	Ohm

Chapter 1: Introduction

Historical and Modern Context of Emerging Viral Diseases

Human history is no stranger to scourges wrought by microbial infections. Despite advances in public health over the last century¹, communicable diseases remain the predominant cause of death in low income countries (Figure 1)^{2,3}. Of particular global concern is the threat of emerging diseases from RNA viruses. These viruses have higher mutation rates compared to their DNA counterparts and are capable of opportunistic infections that can even transgress species boundaries (Figure 2)⁴.

Top 10 causes of deaths in low-income countries in 2016



Source: Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016. Geneva, World Health Organization; 2018. World Bank list of economies (June 2017). Washington, DC: The World Bank Group; 2017 (<https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>).

Figure 1. Communicable diseases remain the leading cause of death in low-income countries. Top 10 causes of death in low income countries between the years of 2000 and 2016, as collected by the World Health Organization (WHO) and the World Bank². “Low-income” is defined as any country with a gross national income (GNI) per capita of 1005 or lower⁵.

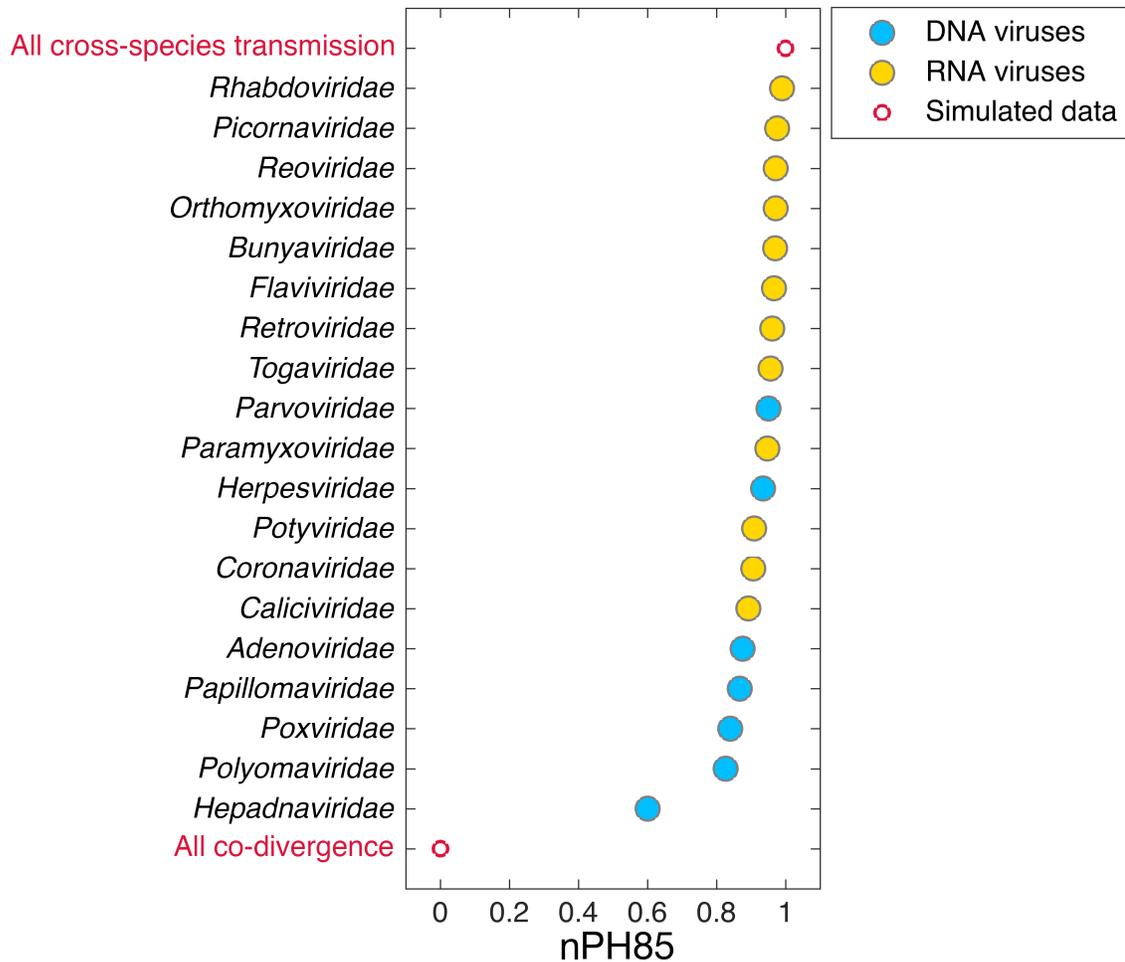


Figure 2. Cross-species transmission occurs more readily in RNA viruses than DNA viruses. Range of topological distance between DNA (blue) and RNA (yellow) virus families as normalized by the Penny and Hendy metric (nPH85). Simulated data (red) denotes the extremes of the nPH85 scale, where 0 is indicative of 1:1 virus/host co-divergence and 1 denotes a propensity for frequent host switching. Virus families are ranked in descending order for visual clarity. Image adapted from Geoghegan *et al.* 2017⁴.

The introduction of viral diseases to naïve human populations has historically been catastrophic. One of the earliest notable instances in the American continents was the unintentional import of viruses such as smallpox and influenza from European invaders⁶. Lacking immunity to these foreign pathogens, indigenous civilizations of the New World quickly succumbed to

disease, particularly in densely populated cities. Mere months after the arrival of the Spanish Conquistadors in April 1520, nearly half the population of the Aztec capital of Tenochtitlán had died of smallpox, rendering the remainder helpless against the assault of Hernán Cortés⁶. Similarly, the Taino people –who had an estimated population of 60,000 to 8 million during the initial contact with Christopher Columbus in 1492- were reduced to less than 500 people by 1548⁶. Smallpox and measles were also introduced to indigenous populations to similar consequence as Europeans colonized what would eventually become the United States of America^{7,8}. Here, transmission was not always unintentional. In an early example of bioterrorism, settlers would intentionally distribute smallpox-infected blankets to native groups⁷. To this day, there are still isolated indigenous tribes in central and south America that are deemed at-risk to transmission of foreign-borne pathogens, whom local governments have taken steps to protect from epidemics⁶.

The threat of emerging viral disease persists despite -and in many cases *because of*- advances in modern technology. The increasing global connectivity of human civilization allows for the rapid transportation of people and goods throughout the world. This also facilitates the unpredictable dissemination of infectious diseases at a previously unprecedented scope⁹. It is not by coincidence that the deadliest recorded pandemic in human history –the H1N1 Spanish influenza that infected a third of the world's population between 1918 and 1919 and killed an estimated 50-100 million people- occurred during the 20th century¹⁰. While no influenza outbreak in the past century has rivaled the

casualties of Spanish flu, the World Health Organization (WHO) considers the advent of the next pandemic flu imminent and urges preparations to control the spread of influenza at both a national and global level¹¹. Another consideration regarding emerging diseases is the geographic expansion of the human population into forested regions of the world, which increases the probability of contact with reservoir host animals harboring viruses that can easily jump across species¹². The progenitor to Human Immunodeficiency Virus (HIV-1) was originally hosted in chimpanzees and the local bushmeat trade has been implicated in its introduction to human hosts¹³. While there is evidence that this jump had occurred several times previously, international trade and global sex tourism in the 1970s exposed HIV-1 to a larger human population than ever before^{14,15}. The virus quickly reached pandemic status due to predominately infecting high-risk groups (male homosexuals, intravenous drug users, and blood transfusion recipients) and delayed government response^{12,14,16}. As of 2017, an estimated 40 million people have died of complications from HIV-1, while 36.9 million are currently living with the virus, 1.8 million of those newly infected that year^{17,18}. The coronavirus responsible for Severe Acute Respiratory Syndrome (SARS-CoV) is another example of a zoonotic virus that was propelled to pandemic status through a globalized economy. The 2003 outbreak started in Guandong, China, where the virus jumped from the initial bat reservoir to civets, and then to a doctor spending the night at a hotel in Hong Kong¹⁹. Thirteen additional hotel occupants were unknowingly infected before returning to their home countries⁹. Over the next 114 days, the SARS-CoV pandemic would

spread to 29 other countries with an estimated 8098 reported cases, 774 subsequent fatalities and a damaging blow to the Asian economy^{20,21}. Worryingly, the capabilities of rapid global travel can also facilitate the transmission of viruses that are typically self-limiting due the severity of pathology. In 2014, an outbreak of Ebola was reported in Guinea that quickly spread to the neighboring countries of Liberia and Sierra Leone²². The situation was declared an emergency of international concern by the WHO that summer after the virus spread to the densely populated cities of all three countries, including the capitols²². The West Africa Ebola epidemic ended in June 2016 with a reported 28,600 cases and 11,325 deaths^{22,23}. An additional 36 cases of Ebola and 15 deaths were also reported amongst seven other countries (Italy, Mali, Nigeria, Senegal, Spain, the United Kingdom, and the United States) over the course of the outbreak, primarily in medical workers²². These reports serve as a cautionary tale of how quickly an emerging virus can be inadvertently transported to other parts of the globe.

Global climate change also exerts an influence on the dynamics of emerging viral diseases^{24,25}. Many RNA viruses utilize hematophagous insect vectors such as ticks and mosquitoes as a central component of their transmission cycle^{26,27}. Increasing global temperatures and extreme climate phenomena such as *El Niño* have expanded the effective breeding and habitat range of insect vectors throughout the world^{28,29}. Of particular concern are mosquitos from the *Aedes* genus that are competent vectors for many medically important viruses^{27,30,31}. Their global distribution has increased dramatically over the years, including

previously inhospitable regions such as Europe (Figure 3)^{32,33,28,34}. Consequently, viruses of an African or tropical origin now have a global presence, particularly from the *Flaviviridae* family³⁵. Dengue virus (DENV), for example, is endemic in over 100 countries, with an estimated 2.5 billion people at risk of infection, and 50 million infections and 500,000 annual hospitalizations from hemorrhagic complications^{36,37,26}.

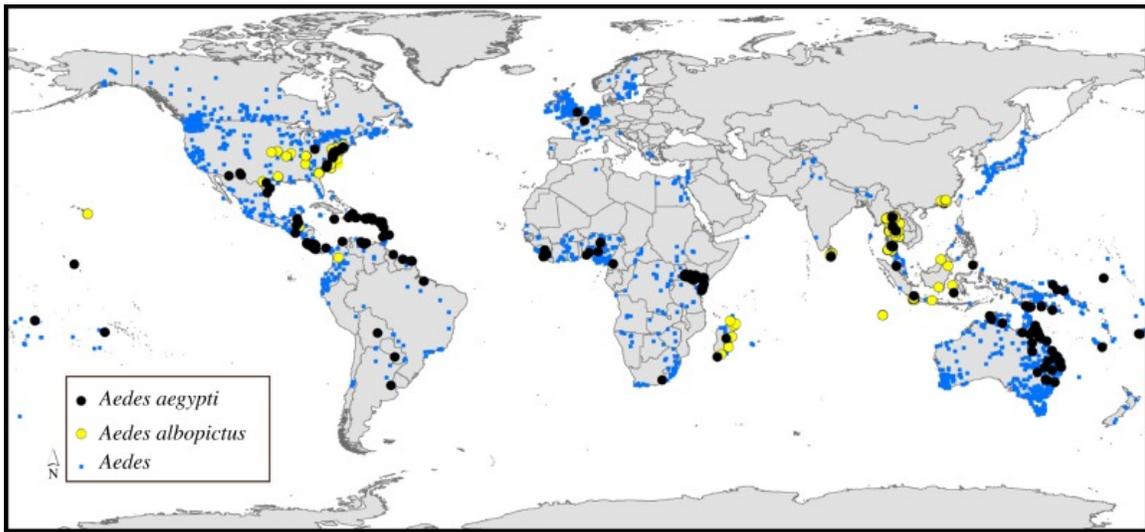


Figure 3. Worldwide presence of vector competent *Aedes* mosquitos. Summary data of global distribution of *Aedes* mosquitos (blue) as of 2015. Specific incidences of *Aedes aegypti* (black) and *Aedes albopictus* (yellow) noted where applicable. Image derived from Campbell *et al.* 2015³⁸.

The combined effects of globalization and climate change were recently responsible for the epidemic emergence of the closely related Zika virus (ZIKV) (Figure 4)^{39,40,41}. First described in 1947 and isolated from a human host in 1952, cases of ZIKV infections were rarely reported and limited to rural forested regions of central Africa^{42,43}. This changed in the past 15 years, in which outbreaks of ZIKV were reported in previously unaffected regions of the world. The first

notable outbreak occurred in 2007 on Yap Island of Micronesia, infecting 5000 people^{44,45}. A second outbreak of ZIKV occurred in French Polynesia in 2013⁴⁵. An estimated 30,000 individuals were infected, and “DENV like symptoms” and neurological complications were reported for the first time in ZIKV cases^{45,40}. The largest recorded epidemic of ZIKV started in Brazil, 2015, shortly after the 2014 World Cup⁴⁶. ZIKV infections once again manifested as DENV-like symptoms of fever, rash headache and arthralgia, but worryingly were also implicated in a sudden increase of Guillain-Barré syndrome and severe congenital malformations including but not limited to neonatal microcephaly^{47,39,48}. The ZIKV virus epidemic quickly spread throughout South and Central America, spurred along by a general lack of immunity to the disease, densely populated urban centers, and an abundance of permissible mosquito vectors^{45,41,49}. Comprehensive statistics are not readily available, but the Brazilian Ministry of Health estimates between 497,593 and 1,482,701 ZIKV virus infections occurred in that country alone⁴¹. Though the ZIKV epidemic didn’t reappear the following summer, it remains a cautionary example of risk posed by an obscure virus re-emerging into the modern world.

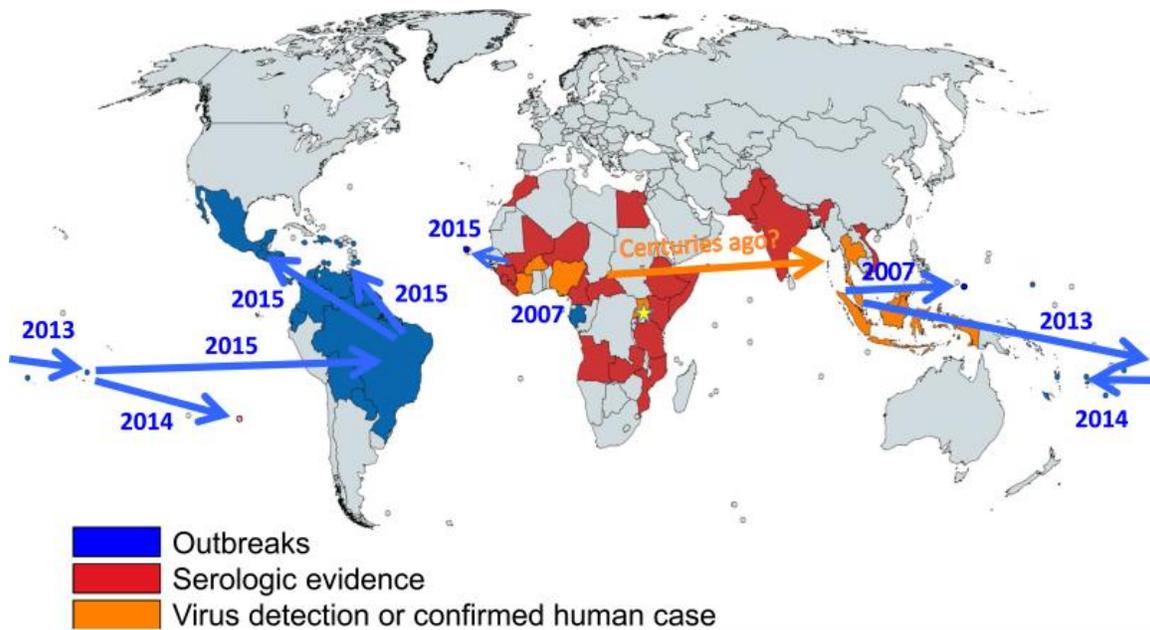


Figure 4. Global proliferation of ZIKV is a recent phenomenon. Chronological spread of ZIKV based on outbreak reports (blue) serologic surveys (red) and laboratory diagnosed individual cases (orange). Initially discovered in 1947, ZIKV originated in central Africa and remained in relative obscurity until a sudden large-scale outbreak on Yap Island of Micronesia in 2007. Blue arrows denote the subsequent spread of the virus to the American continents, precipitating the 2015 Brazilian outbreak. Figure derived from Weaver *et al.* 2016⁴¹.

The Alphaviruses as an Emerging Global Threat

The alphaviruses represent another emerging threat to public health and security. Alphaviruses are small, enveloped, single stranded, positive-sense viruses that comprise a genus within the *Togaviridae* family (Figure 5). The alphavirus genome is roughly 11 kilobases in length, with a 5' cap and a polyadenylated tail. It is organized into two open reading frames (ORFs). The first ORF encodes a polypeptide that is post-translationally cleaved into four non-structural proteins responsible for host evasion, transcription and replication: nsP1, nsP2, nsP3 and nsP4. The second ORF –accessible only as a

subgenomic transcript- encodes a polyprotein comprising the structural proteins C, E3, E2, 6K and E1. Mature alphavirus particles are organized as spherical capsids ~70 nm in diameter with T=4 icosahedral symmetry. The nucleocapsid is enveloped by a host cell-derived lipid membrane obtained through budding⁵⁰. The envelope is further modified with spike protein trimers of E1, E2 and E3 that play a role in both the entry and egress from host cells^{51,52,53}. These viruses are typically maintained in a zoonotic cycle between small mammal or avian reservoirs and blood-feeding arthropod vectors, predominately mosquitos from the *Aedes* and *Culex* genera⁵⁴. As a consequence of increased globalization and urbanized environments, humans can serve as amplifying reservoirs as well⁵⁵.

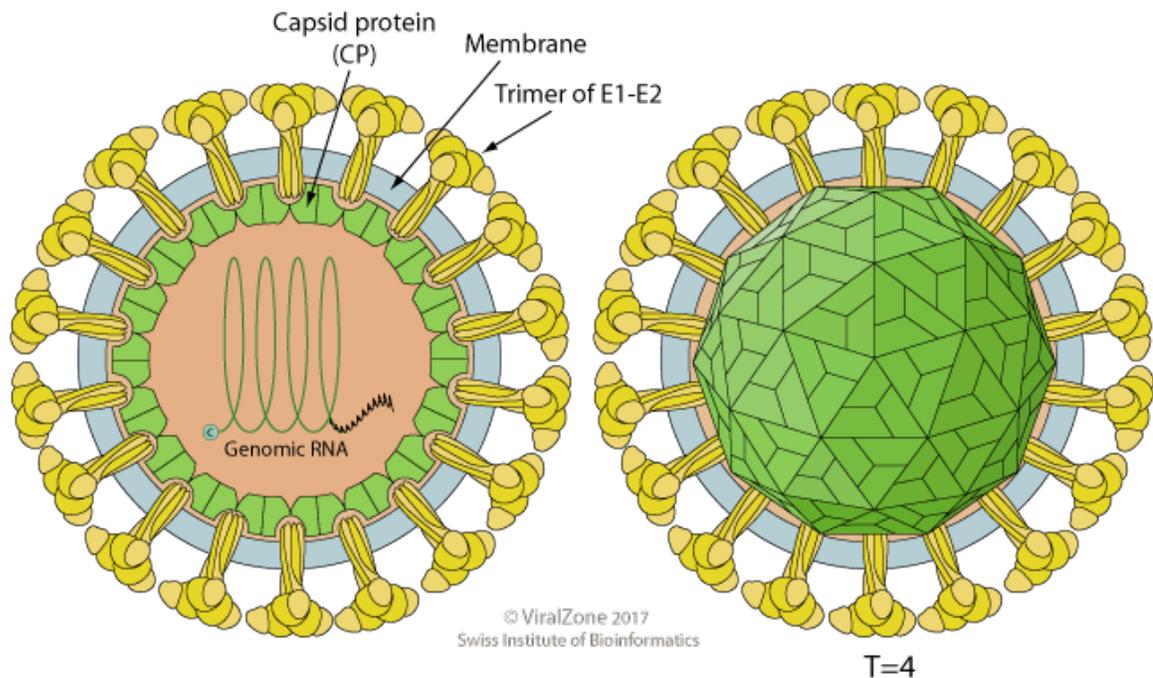


Figure 5. Alphavirus structural organization. Generic schematic of a mature alphavirus virion. A capped, polyadenylated, positive sense RNA genome is encapsidated a 70 nm diameter icosahedral capsid with T=4 symmetry. A host-derived membrane envelops the capsid. Trans-membrane spike proteins are used for host cell entry and membrane fusion. Image derived from ViralZone⁵⁶.

Over 26 species of alphavirus have been described, several of which are of medical concern to humans⁵⁷. They can be roughly organized into two major groups based on geographic distribution and pathology: old world and new world. Old world alphaviruses are endemic to Africa and Asia and include medically relevant viruses such as Chikungunya (CHIKV), O’Nyong’Nyong (ONNV), Sindbis (SINV) and Semliki Forest (SFV). Infection by these viruses are predominately characterized by febrile illness and severe arthralgia⁵⁸. These symptoms typically result in high morbidity in human hosts for weeks or even months^{59,60,61}. Additionally, complications during the infection can lead to painful sequelae that can persist for years after viral clearance⁶². Climate change and human globalization over the past half-century have exerted a tremendous influence over the global distribution patterns of these viruses²⁵. CHIKV, for example, was an uncommonly reported disease since its initial discovery in 1952⁶³, primarily associated with small, localized outbreaks in forested regions of Africa. In 2004, the first large-scale outbreaks of CHIKV were reported in Africa and Asia⁶⁴. Over the next few years CHIKV would become established to 44 other countries, including the Floridian peninsula of the United States in 2014 (Figure 6)^{26,65,66}.

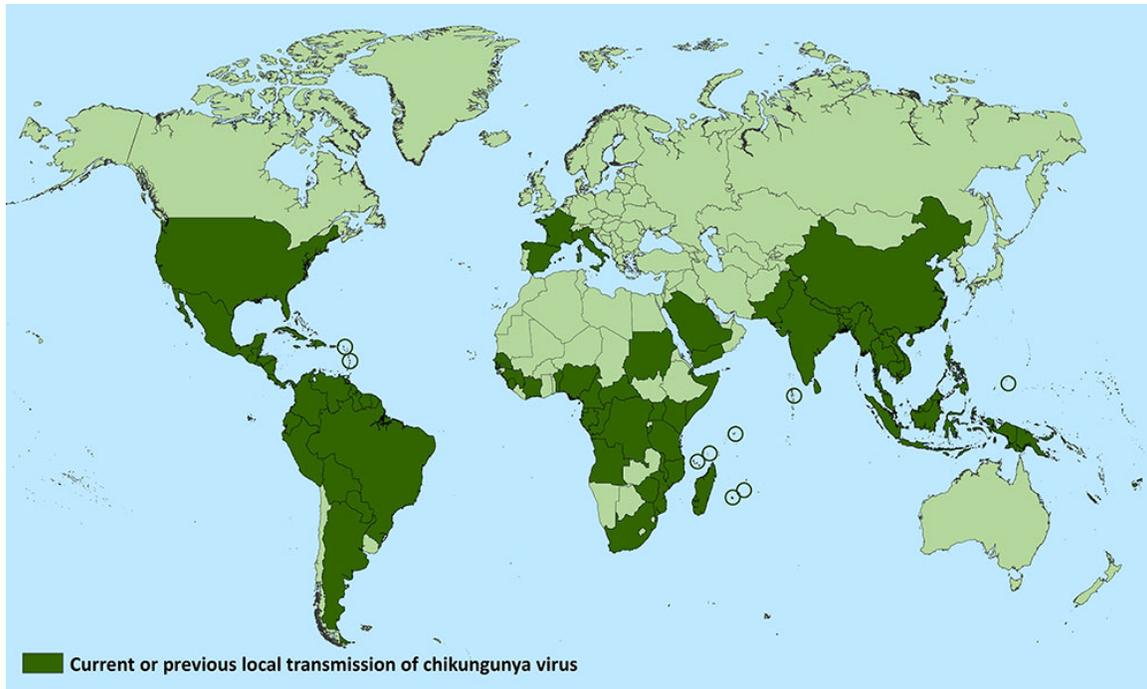


Figure 6. Global distribution of CHIKV. Map of the worldwide presence of CHIKV as of May 2018. Dark green countries highlight countries where local mosquito transmission has been reported. Circles are used to highlight islands for added visual clarity. Imported cases of CHIKV are not shown on this map. Image obtained from the CDC⁶⁷.

The new world alphaviruses are endemic to the American continents and are comprised of the Eastern, Venezuelan and Western equine encephalitis viruses (EEEV, VEEV and WEEV, respectively). Similar to their old world counterparts, these viruses are transmitted through mosquito vectors and can infect large mammal hosts such as humans and equines⁵⁴. New world alphavirus infection exhibits a biphasic pattern of disease. The initial phase is characterized by high-titer viremia and replication in lymphoid tissue or osteoblasts in the case of EEEV^{68,69}. Symptoms during this phase consist of fever, chills, vomiting and myalgia that manifest two to five days post infection⁵⁸. The second phase occurs in the event of the virus bypassing the blood-brain barrier and infecting the

neurons of the brain and spinal cord, leading to cases of acute encephalomyelitis⁷⁰. In equines, progression to encephalitis occurs five to ten days post-infection and has a high-case fatality rate up to 90% for EEEV, 70% for VEEV and 50% for WEEV, respectively^{71,72,73}. This fact alone makes these viruses severe social and economic threat to livestock producing countries.

The incidence and severity of encephalitis in human infections is heavily influenced by the specific alphavirus. EEEV has been implicated in the most severe cases of encephalitis in humans. Infected individuals experience an initial phase of escalating fever and headaches that ultimately culminates in neurological disease with the possibility of paralysis, coma and death in 30 to 70% of cases^{71,74}. Complications from encephalitis will persist as debilitating sequelae in 35% of surviving patients⁷⁴, the long term care of which can total up to several million dollars⁷⁵. Currently, the neurovirulence of EEEV is mitigated by a low incidence of infection, with only 270 reported cases between 1964 and 2010⁷⁶. However, the recent geographic expansion of demonstrated alphavirus vector *Aedes albopictus* in the United States has elevated the risk of larger scale EEEV outbreaks in the future⁷⁷.

In contrast, the case fatality rate of WEEV in humans is considerably lower at an estimated 3-4%⁷⁸. Fewer than 700 cases of WEEV have been documented since 1964⁷⁹. However, WEEV has demonstrated epidemic potential since its discovery in 1930, with fatal outbreaks peaking between 1940-1950^{80,81}. The decrease of reported WEEV infections is not attributed to any notable decline in the virulence of the virus itself, and is more likely associated with changing

environmental factors unfavorable for the competence of their associated insect vectors^{79,82,83}. As with EEEV, the risk remains that shifting environmental circumstances can lead to the reemergence of WEEV as an epidemic threat.

Presently, VEEV is the most active of the encephalitic alphaviruses (Figure 7). First isolated in Yaracuay state, Venezuela in 1938, the virus is endemic to Central and South America. VEEV is comprised of 14 subtypes and 7 different virus species (figure 7)⁵⁵. Six of these subtypes have been implicated as antigenic to human hosts, though the IA/B and IC epizootic strains are responsible for the largest outbreaks of encephalitis^{72,84}. VEEV has been implicated in many large-scale outbreaks in the American continents with occasional reported fatalities⁸⁵. One of the largest of these outbreaks occurred in Venezuela and Columbia in 1995 and infected over 100,000 people⁸⁶. Outbreaks of VEEV have also been reported in North America as well, particularly in Mexico and the southern border of the United States⁸⁷. A comprehensive look at the scope of VEEV outbreaks is difficult to achieve, as the virus inhabits the same geographic range as other mosquito-borne arboviruses such as DENV and share many similarities in clinical presentation⁸⁸. VEEV infection in humans is characterized by acute, incapacitating morbidity consisting of fever, headaches and myalgia, followed by a longer two-to-three week period of lethargy and anorexia^{70,71,89}. The progression of encephalitis from VEEV infection is uncommon, with an incidence of >1% that predominantly afflicts children and the elderly⁸⁹. Mortality from VEEV-instigated encephalitis varies by age, ranging from

10% in adults and up to 35% in children, with the threat of lifelong neurological complications persisting after viral clearance^{72,90}.



Nature Reviews | Microbiology

Figure 7. Major outbreaks of VEEV in the Americas. Geographic distribution and specific incidences of VEEV outbreaks in North, Central and South America. Shaded purple areas denote regions of major VEEV outbreaks primarily from the IA/B and IC epizootic strains. Dates of major outbreaks and the responsible VEEV subtype are identified in the shaded area. Colored symbols are used to represent additional enzootic subtypes of VEEV. Figure originally appears in Weaver and Barrett, 2004⁹¹.

Separate from natural outbreak concerns, the encephalitic alphaviruses also pose a threat to public health and national security as a potential bioterror weapon⁹². These viruses have been classified as a category B select agent by the National Institutes of Health (NIH) and Centers for Disease Control and Prevention (CDC) due to its high morbidity in human hosts and characteristics conducive for weaponization⁹³. Firstly, they can be quickly grown to high titers, require a low infectious dose and can be easily lyophilized^{94,95,96}. Secondly, all three of the encephalitic alphavirus have demonstrated a potential to achieve lethal infection through the aerosol route, likely infecting the central nervous system through the olfactory bulb^{97,98,99,100}. This bypasses the requirement for insect vectors or other mechanical means of transmission. Thirdly, VEEV was successfully weaponized by both the United States and the former Soviet Union during the 1950s and 1960s^{101,95}. This means that the procurement, production and dissemination of biothreat encephalitic alphaviruses at a large scale is achievable to malicious actors with even modest modern laboratory facilities¹⁰².

It therefore constitutes an urgent problem that there are currently no therapeutics against alphaviruses that the Food and Drug Administration (FDA) has approved for civilian use. Veterinary and investigational vaccines have been approved for at-risk military personal and laboratory researchers, but generally produce unacceptably severe side effects. One prominent example is TC-83, a live attenuated strain of VEEV. TC-83 is derived from serial passage of the neurovirulent VEEV Trinidad Donkey (TrD) strain through guinea pig heart cells¹⁰³. TC-83 attenuation is attributed to two point mutations. The first of which

is located in the 5' UTR and destabilizes a critical stem-loop element (figure 8)¹⁰⁴, while the second mutation is found in the E2 envelope glycoprotein¹⁰⁵. While TC-83 vaccination elicits a neutralizing antibody response in 80% of human recipients, moderate flu-like symptoms have been observed in 40% of cases⁹⁴. TC-83 is a similarly sub-optimal vaccine candidate for animals. Significant levels of viremia are observed in vaccinated horses and acute illness and death has been observed in mouse strains following subcutaneous or intra-cranial inoculation^{106,107}. While other vaccination modalities against VEEV are in development against encephalitic alphaviruses (summarized in figure 9)⁹⁴, live attenuation models consistently remain the most immunogenic and efficacious⁹⁴. Therefore, emphasis should be placed on the discovery of attenuating mutations that minimize pathology and are resistant to wild-type reversion.

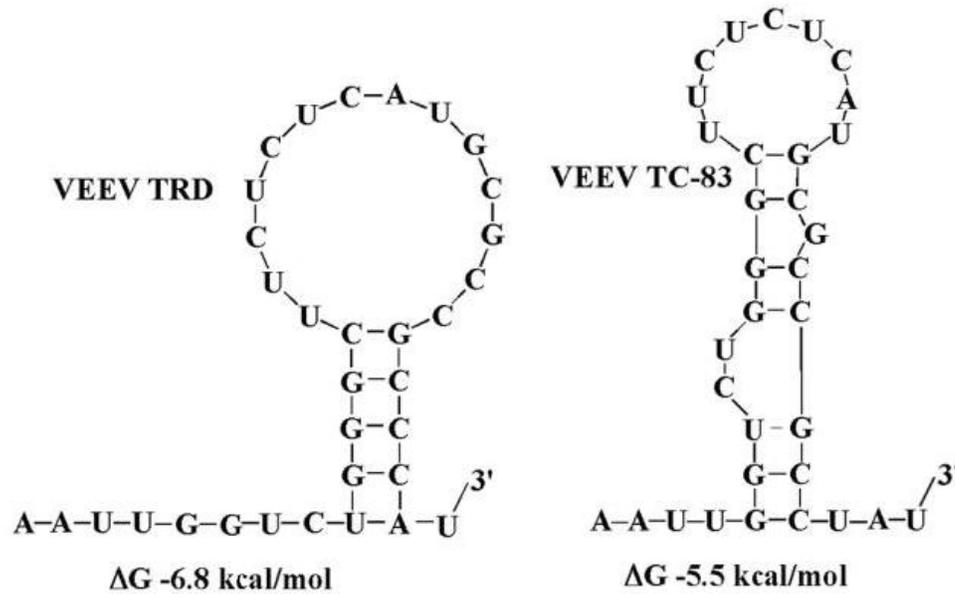


Figure 8. A single point mutation in VEEV TC-83 destabilizes a critical stemloop in the 5' UTR. Secondary structure and free energy comparison of a 5' UTR stemloop between VEEV TRD and VEEV TC-83. A notable characteristic of VEEV TC83 is a C to U mutation near the base of the stem on the 3' side. Consequences of this mutation include a markedly less stable stem and smaller loop. Figure derived from Kulasegaran-Shylini *et al.* 2009¹⁰⁴.

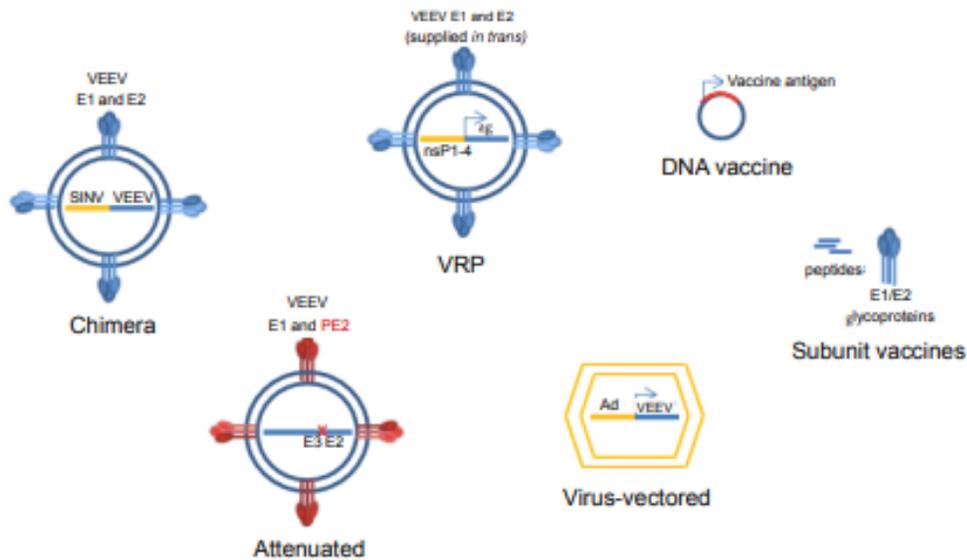


Figure 9. A multitude of proposed vaccine modalities against VEEV. Examples of other vaccine strategies in development against VEEV beyond the live-attenuated model. Chimera models are a synthesis between the non-structural proteins of old world SINV and the structural proteins of VEEV. Viral replicon particles (VRPs) are infectious but non-replicating variants of VEEV. Virus-vectorized models use viruses such as adenovirus to transmit key antigens of VEEV to targeted cells. DNA vaccines are comprised of plasmids that express VEEV antigens into target cells. Subunit vaccines contain whole or fragment peptides from VEEV designed to provoke a host immune response. Image derived from Spurgers and Glass, 2011⁹⁴.

Programmed Translational Recoding Signals in Viruses

Programmed translational recoding mechanisms are an underexplored potential target for vaccine development. Translational recoding is broadly defined as processes that direct ribosomes to transgress canonical translational mechanisms. Translational recoding is heavily utilized by many RNA viruses as a means to expand the coding capacity of what are typically very limited genomes. Many modalities of translation recoding have been documented, but of particular

interest to this document are programmed -1 ribosomal frameshifting (-1 PRF) and termination codon readthrough (TCR).

Evidence of -1 PRF was initially documented in the Rous-Sarcoma retrovirus (RSV) in 1985 by Tyler Jacks and Harold Varmus. Analysis of the RSV genome revealed that sequence for coat protein *gag* overlapped with the 5' open reading frame for replicase gene *pol* by 205 to 241 nucleotides. Expression of *pol* was contingent on the synthesis of a 160 kDa *gag-pol* fusion polypeptide at a 1:20 ratio to standalone *gag* expression¹⁰⁸. Closer sequencing analysis revealed presence of an amber (UAG) termination codon 58 nucleotides into the overlapping reading frames and more crucially that the overlapping *pol* sequence was encoded in a -1 reading frame respective to *gag* (figure 10A)¹⁰⁸. The synthesis of the *gag-pol* fusion protein, then, was reasoned to only occur consequent to an RNA splicing event (despite a marked absence of traditional donor or acceptor sites) or a -1 ribosomal frameshifting event. Experiments utilizing a cell free rabbit reticulocyte lysate system demonstrated the ability of both *gag* and *gag-pol* products to be synthesized from the same RNA transcript (figure 10B)¹⁰⁸, lending strong support to the latter hypothesis. Subsequent work sought to elucidate the precise location of the frameshift event^{108,109}. RSV *gag-pol* synthesis was shown to be unaffected by large deletions to the *gag* sequence, including the termination codon. The minimum sequence required to promote efficient RSV -1 PRF was ultimately localized to a 147 nucleotide sequence along the overlapping reading frames that notably harbored an A AAU

UUA sequence amenable to tRNA repositioning and a downstream stemloop structure (figure 11)¹⁰⁹.

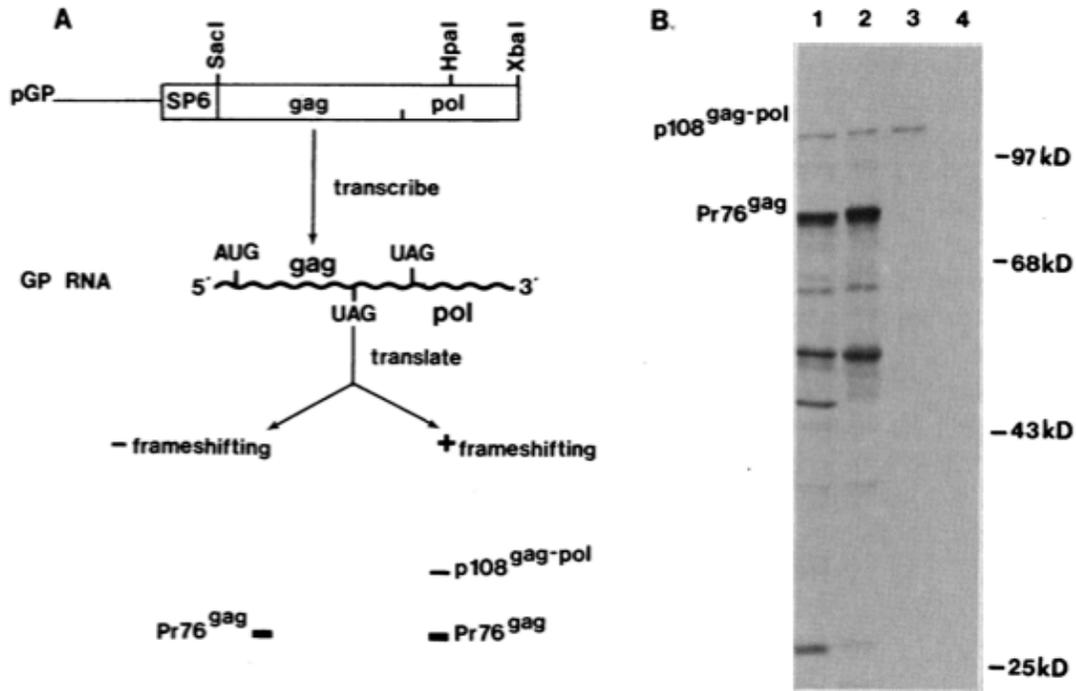
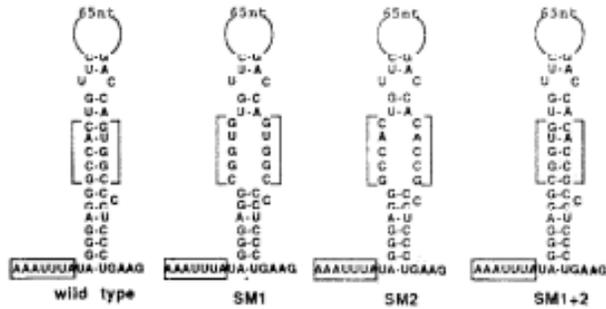


Figure 10. Initial discovery of -1 PRF in RSV. (A) *In vitro* experimental schematic for assaying -1 PRF. Gag-pol region of RSV was cloned into SP6 vector SP65. Vectors were linearized with XbaI prior to *in vitro* translation in a rabbit reticulocyte system. Anticipated kDa weights of RSV gag and gag-pol shown below. (B) Autoradiogram of a 10% SDS-PAGE of ³⁵S-labeled translation products. Lane 1 is unprecipitated product, lane 2 is immunoprecipitated with rabbit serum to p19^{gag}, lane 3 is immunoprecipitated with rabbit anti-reverse transcriptase serum, and lane 4 is immunoprecipitated with non-immuned rabbit serum. Weight markers are shown on the right of the gel, while the specific identities of PR76^{gag} and P108^{gag-pol} are shown on the left. Image taken from Jacks and Varmus, 1985¹⁰⁸.

A.



B.

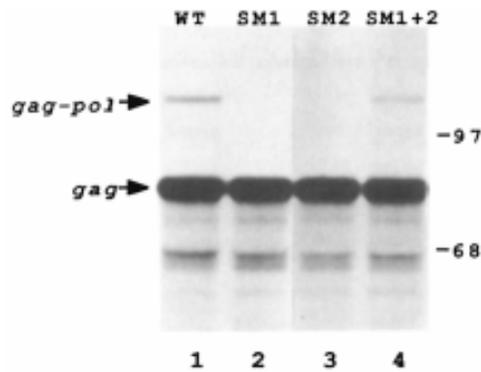


Figure 11. Downstream stimulatory structure implicated in the efficient -1 PRF of RSV. (A) Predicted secondary structure of stimulatory hairpin element downstream of predicted RSV slippery sequence. A collection of mutants –SM1, SM2, and intended rescue mutant SM1+2- and the predicted consequences on the wild-type secondary structure are also depicted. (B) Subsequent fluorogram of a 10% SDS-PAGE of ^{35}S -labeled, rabbit reticulocyte-derived translation products of (lane 1) wildtype RSV (wt), (lane 2) SM1, (lane 3) SM2, or (lane 4) SM1+2. Band identities of RSV gag and gag-pol are indicated on the left and molecular weights on the right. Image taken from Jacks and Varmus, 1988¹⁰⁹.

The *gag-pol* -1 PRF junction was further explored in the human immunodeficiency virus (HIV-1)¹¹⁰. Similar to the arrangement described in RSV, the reading frames for *gag* and *pol* overlapped, with the latter gene only expressed as a fusion protein consequent to a -1 PRF event. Recapitulating the

results found with RSV, *gag* and *pol* were arranged as overlapping reading frames on the HIV-1 genome, the presence of the resulting fusion protein once again detectable following expression in a rabbit reticulocyte lysate system¹¹⁰. As a shift into the -1 frame by definition alters the codon identity of the downstream nucleotide sequence, the radioactivity profiles of the fusion polypeptide were analyzed via Edman degradation to verify the presence of the trans-frame polypeptide (figure 12)¹¹⁰. The position of the trans-frame peptide once again implicated the conserved UUA leucine codon near the 5' end of the overlap as the locus of the frameshift, and subsequent experimental mutations ablated synthesis of *gag-pol* (figure 13)¹¹⁰. Later studies investigated the role of the structure downstream of this slippery region. Silent mutations that destabilized the secondary structure of this element were correlated with a decrease in -1 PRF in HIV-1 and related retroviruses¹¹¹.

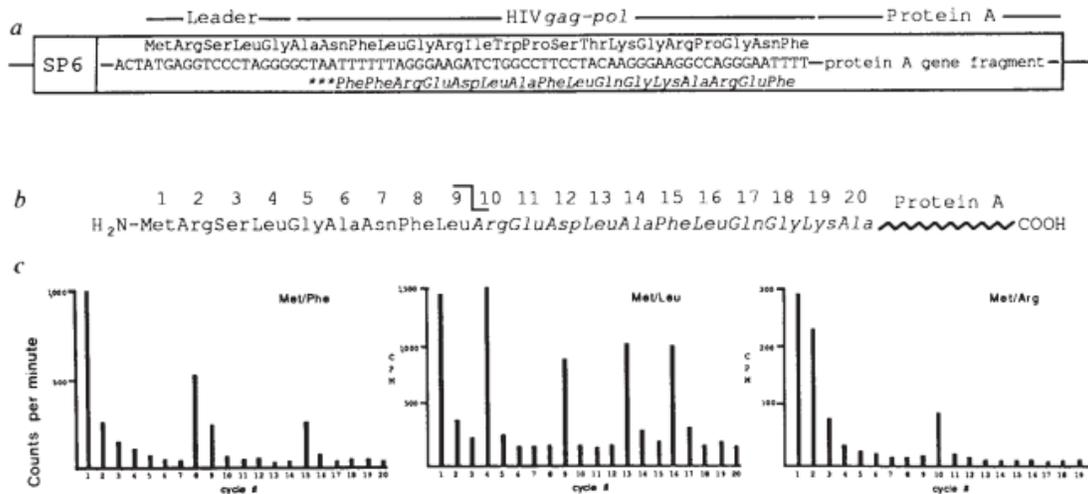


Figure 12. Biochemical evidence of a frameshifted gag-pol peptide in HIV-1. (A) Nucleotide and corresponding amino-acid sequences of the HIV-1 -1 PRF signals cloned into and SP6 vector (pHSS). The nucleotide sequence of the HIV-1 -1 PRF sequence is displayed in the center, preceded by a short leader sequence. The zero frame amino acid sequence is displayed above the nucleotide sequence, while the amino acid sequence achieved through a -1 PRF event is shown below in italics. (B) Predicted amino-acid sequence of the frameshifted gag-pol peptide. The bracket line between codon 9 and 10 denotes the position of the predicted trans-frame peptide junction. (C) Radioactivity profiles of automated Edman degradation of pHSS-encoded protein synthesized in the presence of [³⁵S]methionine and either [³H]phenylalanine (left panel), [³H]leucine (center panel) or [³H]arginine (right panel). Figure derived from Jacks and Varmus 1988¹¹⁰.

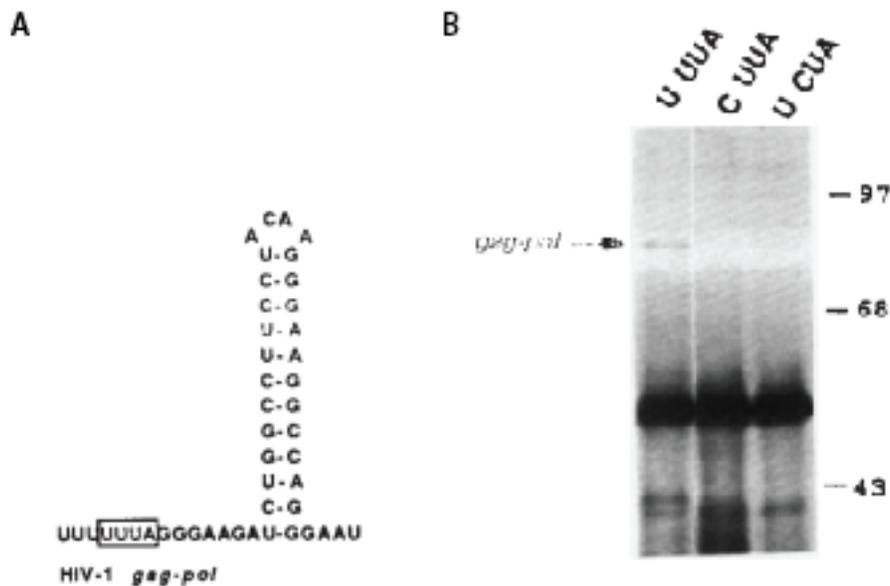


Figure 13. Efficient HIV-1 -1 PRF occurs at a designated slippery site. (A) Diagram of HIV-1 gag-pol junction sequence with highlighted UUUU “slippery” sequence and downstream stimulatory hairpin element. (B) *In vitro* translation products of an M13 DNA vector harboring the cloned in sequence from 13A that were immunoprecipitated with anti-p25^{gag} antiserum and separated through 10% SDS-PAGE. Lanes 2 and 3 contain samples with U to C point mutations in the highlighted UUUU sequence in 13A. The expected position of the gag-pol band is indicated on the left. Molecular weight positions are shown on the right. Figure is taken from Jacks and Varmus 1988¹¹⁰

The modern canonical model of a -1 PRF signal consists of a heptameric “slippery” sequence N NNW WWH, where spaces denote 0 frame codon arrangement, N denotes identical nucleotides, W denotes identical weak bases adenine or uracil, and H is not guanine, an optional spacer region and a highly structured downstream region that is typically either a hairpin stemloop¹¹² or a pseudoknot^{113,114}. These elements function in concert as a kinetic trap for elongating ribosomes, a fraction of which will pause over the slippery site¹¹⁵. Per the “simultaneous slippage” model, the non-wobble bases of tRNAs occupying

the A and P sites of the ribosome shift one nucleotide in the 5' direction on the slippery site (Figure 14)⁵⁶. When the ribosome resumes translation, it does so in the -1 frame, permitting the translation of an alternative polypeptide¹¹⁶.

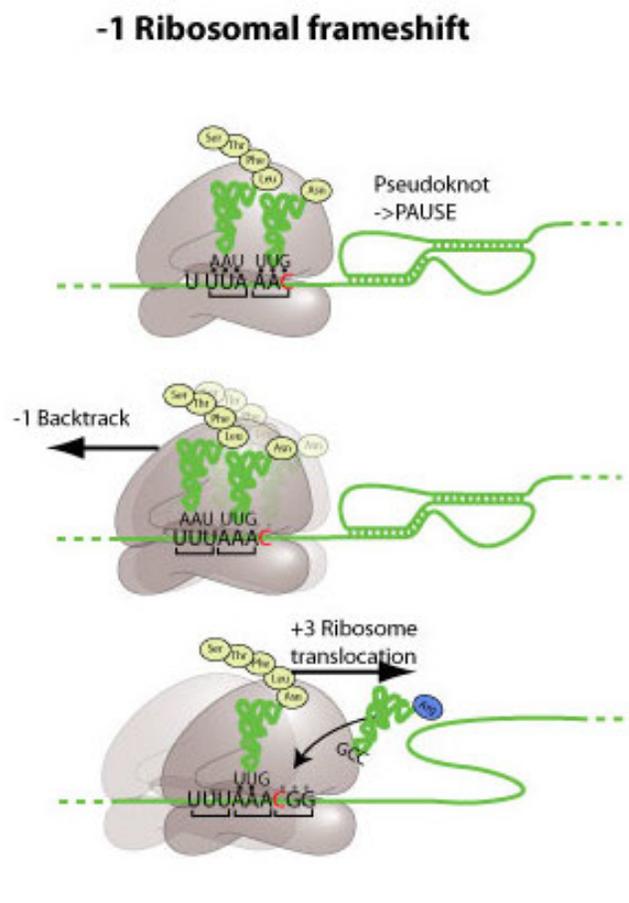


Figure 14. Canonical mechanism of -1 PRF. A canonical -1 PRF signal is comprised of a N NNW WWH slippery site, short spacer region, and downstream stimulatory element. These components work in concert as a kinetic trap for elongating ribosomes, a fraction of which will pause over the slippery sequence. Changes in kinetic partitioning result in the A and P site tRNAs to reposition one nucleotide in the 5' direction, allowing translation in the -1 frame when ribosomal elongation resumes. Image from ViralZone⁵⁶.

Concurrent research to that of RSV reported an alternative mechanism for *gag-pol* polypeptide synthesis in other retroviruses. In retroviruses such as feline and murine leukemia virus (FeLV, MuLV, respectively), *gag* and *pol* were shown to share the same reading frame^{117,118}. Here, the amber termination codon serves as the junction point between the two genes, the readthrough of which was required to synthesize *gag-pol*. Experiments with beta-galactosidase reporter vectors demonstrated that efficient termination readthrough is predicated on the presence of a downstream stimulatory element similar to that of -1 PRF utilizing retroviruses (figure 15)¹¹⁹.

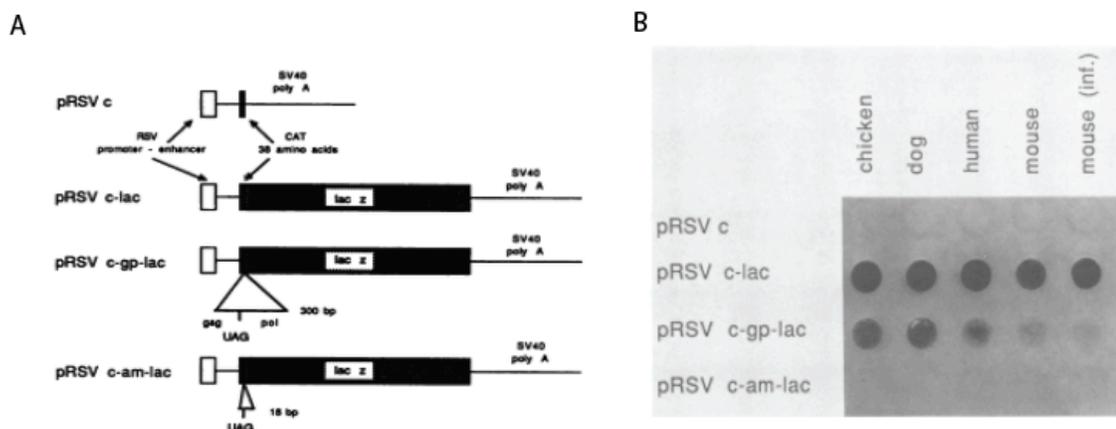


Figure 15. Amber TCR is required to express gag-pol in some retroviruses. (A) Experimental schematic for beta-galactosidase constructs that assay efficient TCR of a type C murine retrovirus (AK). Constructs include an RSV promoter-enhancer region followed by a minimal leader sequence as a negative control (pRSV c), an unaltered LacZ gene as a positive control (pRSV c-lac), 300 bases of the AK gag-pol sequence inserted 5' of LacZ (pRSV c-gp-lac), or a single amber stop codon 5' of LacZ (pRSV c-am-lac). (B) Monoclonal antibody detection of β -gal following transfection into vertebrate cells and spotting of extracts onto nitrocellulose. Images derived from Panganiban, 1988¹¹⁹.

With minor differences, viral TCR signals share a similar collection of cis-acting elements described in -1 PRF signals. In this instance, the ribosome pauses over a termination codon, where the downstream stimulatory element is hypothesized to prevent association of the eRF1/3 complex. This increases the probability of the A site erroneously accepting a near cognate tRNA in the place of eRF3, permitting the ribosome to continue translating the current reading frame (figure 16)⁵⁶. Numerous studies have also shown that the likelihood of TCR recoding is heavily influenced by the specific stop codon identity, with the less common amber and opal (UGA) stop codons more prominently associated with these signals¹²⁰. Other factors, such as the identity of the nucleotides directly 3' of the stop codon have also been implicated as contributors to leaky termination¹²¹.

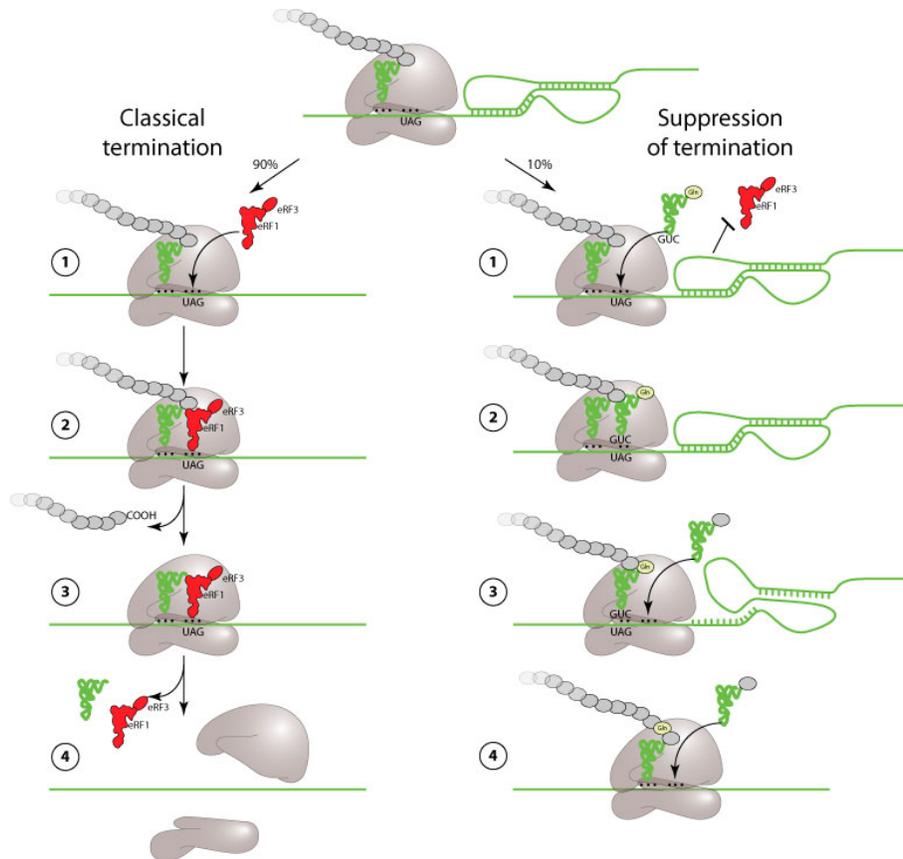


Figure 16. Canonical mechanism of TCR. Schematic of how *cis*-acting elements on the RNA transcript can facilitate efficient TCR. Under classical termination circumstances a ribosome will pause over a termination codon, enabling the incorporation of the eRF1/3 complex and subsequent dissociation of the ribosomal subunits from the transcript. A TCR signal contains a highly structured downstream element that impedes the association of the eRF1/3 complex and increases the likelihood that the A-site vacancy will instead be filled by a near-cognate tRNA. Should this occur, translation will resume on the 0 frame, allowing for an extended peptide sequence. Image obtained from ViralZone⁵⁶.

The efficiency of these recoding signals is a crucial component of viral gene expression. Seminal research on this subject was conducted in the totivirus L-A, a small double-stranded RNA virus that infects yeast. *Gag-pol* frameshifting in a manner similar to retroviruses has been demonstrated in L-A at a recoding efficiency of ~1.9%¹²². As the L-A capsid is organized into a simple icosahedron

of 60 subunits, the implication here is that L-A synthesizes *gag-pol* at an efficiency that allots one fusion protein for every capsid¹²³. When the L-A slippery sequence was experimentally mutated to alter -1 PRF efficiency above or below 1.9% a marked decrease in viral propagation was observed (figure 17)¹²⁴. Similar results were observed when the L-A slippery sequence was replaced with a +1 PRF slippery motif derived from the yeast retrotransposon Ty1, demonstrating that the importance of the stoichiometric ratio between *gag* and *gag-pol* is independent of the recoding mechanism¹²⁴. An imbalance in the *gag* to *gag-pol* ratio was similarly demonstrated to impair the replication of HIV-1¹²⁵, raising the possibility that -1 PRF represents a potential target for novel therapeutic exploitation.

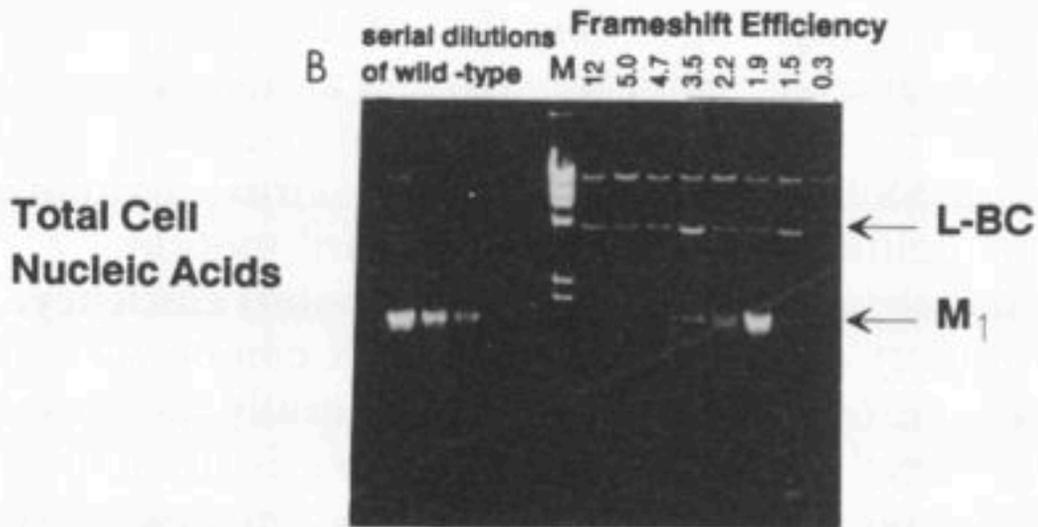


Figure 17. Altering L-A -1 PRF efficiency diminishes propagation of M₁ satellite virus. Effect of altered -1 PRF efficiency on the viability of L-A, measured by the consequences to M₁ propagation. Plasmids with cloned slippery site mutations of L-A -1 PRF were introduced alongside M₁ into yeast strain 3063. As 3063 harbors a *mak10-1* mutation, L-A is lost while M₁ is stably supported by a wild-type L-A reporter (-1 PRF = 1.9%). Nucleic acids extracted from these strains were run through a 1.5 agarose gel stained with ethidium bromide. Frameshift efficiencies of mutant L-A (determined elsewhere via beta-galactosidase experiments) are shown on the top. Band position of M₁ is indicated on the right. L-BC, a virus unrelated to L-A or M₁, is also shown. Figure adapted from Dinman and Wickner, 1992¹²⁴.

While much of the previous text has contextualized recoding events as the regulatory junction point between structural and enzymatic viral genes, it is important to note that these mechanisms serve a more modular function for viral gene expression. For example, research into SARS-CoV has identified a -1 PRF signal that serves as a division point between the early stage genes required for host cell takeover, and the ORF harboring genes required for viral replication¹²⁶.¹²⁷ Here too, the proper stoichiometric expression of the two ORFs is critical.

Experimental mutations to the SARS-CoV slippery site that ablated -1 PRF were correlated with significantly diminished production of viral genomic and sub-genomic RNAs (figure 18A)¹²⁸. Mutations that destabilized the downstream three-stemmed pseudoknot serving as the stimulatory element for SARS-CoV achieved similar results (figure 18B,C)^{129,130}.

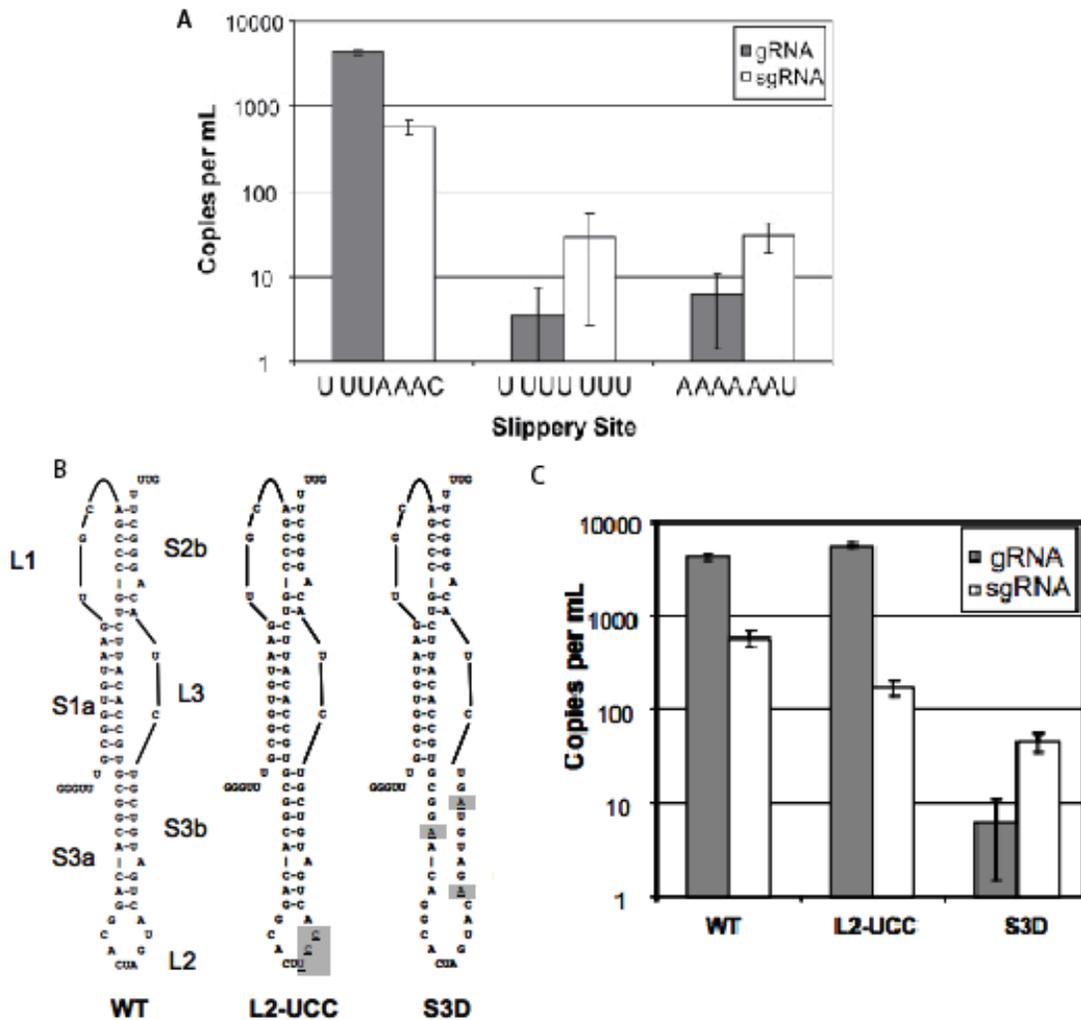


Figure 18. Alterations to efficient -1 PRF diminish genomic and subgenomic RNA production in SARS-CoV. (A) Relative abundance of SARS-CoV genomic (gRNA) and subgenomic (sgRNA) RNA following mutations to the slippery site. Taqman analysis was used to measure the abundance of viral RNA molecules compared to a reference RNA transcribed from a WT SARS-CoV replicon. (B) Schematic of the stimulatory pseudoknot of WT SARS-CoV. Two additional mutations –L2-UUC and S3D– are shown alongside. Shaded areas denote mutated areas, and the structures have subsequently been modified to reflect predicted changes from those mutations. (C) Relative abundance of SARS-CoV gRNA and sgRNA RNA consequent to pseudoknot mutations. Taqman analysis was used to measure the abundance of viral RNA molecules compared to a reference RNA transcribed from a WT SARS-CoV replicon. Figures jointly obtained from Plant *et al.* 2010 (A) and Plant *et al.* 2013 (B, C)^{128,129}.

Translational recoding, then, can be distilled into a few unifying principles. *Cis*-acting elements on the RNA transcript serve as a kinetic trap for elongating ribosomes¹³¹, stalling them over a recoding sequence, be it a slippery sequence or uncommon stop codon. Fluctuations in kinetic partitioning during the ribosomal pause increase the probability of a recoding event, resulting in the translation of alternative polypeptides at precise stoichiometric ratios.

A growing amount of research suggests that the conformational dynamics of -1 PRF stimulatory elements are highly influenced by the helicase activity of elongating ribosomes. Indeed, optical tweezer experiments conducted on the HIV-1 -1 PRF signal revealed several different refolding conformations after ribosomal perturbation, ranging from a simple hairpin stemloop to a pseudoknot-like triplex structure¹³². Elements upstream of the slippery site have also been implicated in exerting regulatory control over the central PRF structure¹³³. Additional research, particularly in plant viruses, has reported long range interactions from elements elsewhere on the viral genome that can further influence the conformational dynamics of recoding signals^{134,135,136,137}. While the full scope and function of these interactions are poorly understood, it is apparent that the factors that govern programmed translational recoding extend beyond the immediate structure.

Two separate instances of translational recoding signals have been recorded in the alphavirus genome. In the non-structural ORF, a widely conserved opal termination codon has been identified at the 3' end of the nsp3 gene, the readthrough of which is required to express the alphavirus replicase nsP4¹³⁸. The

elements that govern this readthrough are currently poorly understood, and only limited computational predictions currently exist to describe the nature of the downstream stimulatory elements for this signal (figure 19)¹³⁹. However, research has shown that mutating the opal stop to an arginine codon is correlated with attenuated pathogenicity in old world alphaviruses (figure 20)¹⁴⁰, again demonstrating the optimal stoichiometric gene expression ratios enforced by these signals. Similar mutations have also been associated with decreased infectivity in mosquito vectors, suggesting that the presence of the opal stop codon between nsP3 and nsp4 confers important fitness advantages outside of the context of host cell infection as well¹⁴¹.

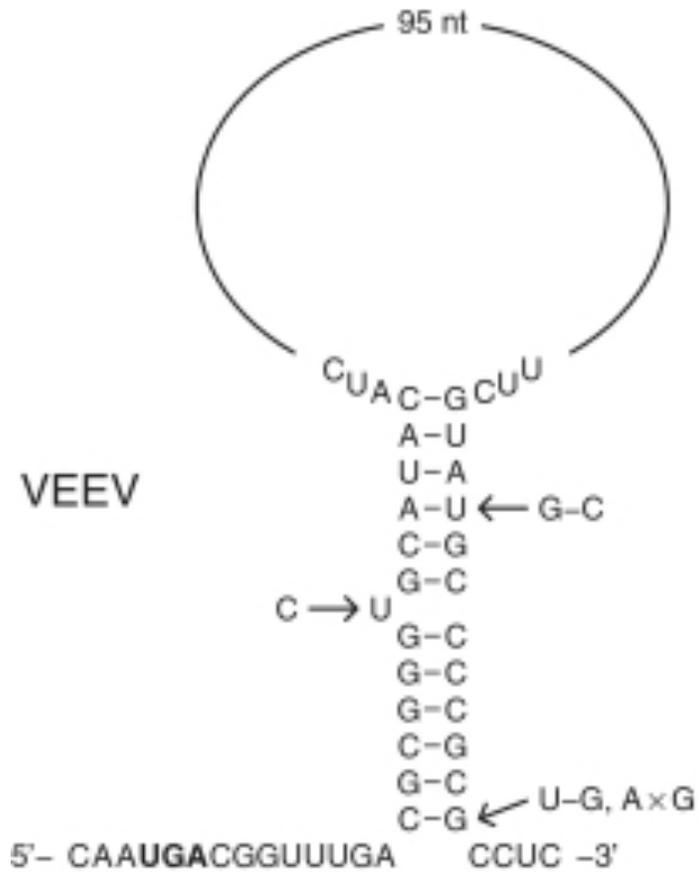


Figure 19. Computational prediction of VEEV TCR signal. Predicted secondary structure of VEEV TCR signal. Position of the amber termination codon is shown in bold. Sequences variations of the stem are shown with arrows. Remaining 95 nucleotides after the first stem are abstracted with a circle. Image derived from Firth *et al.* 2011¹³⁹.

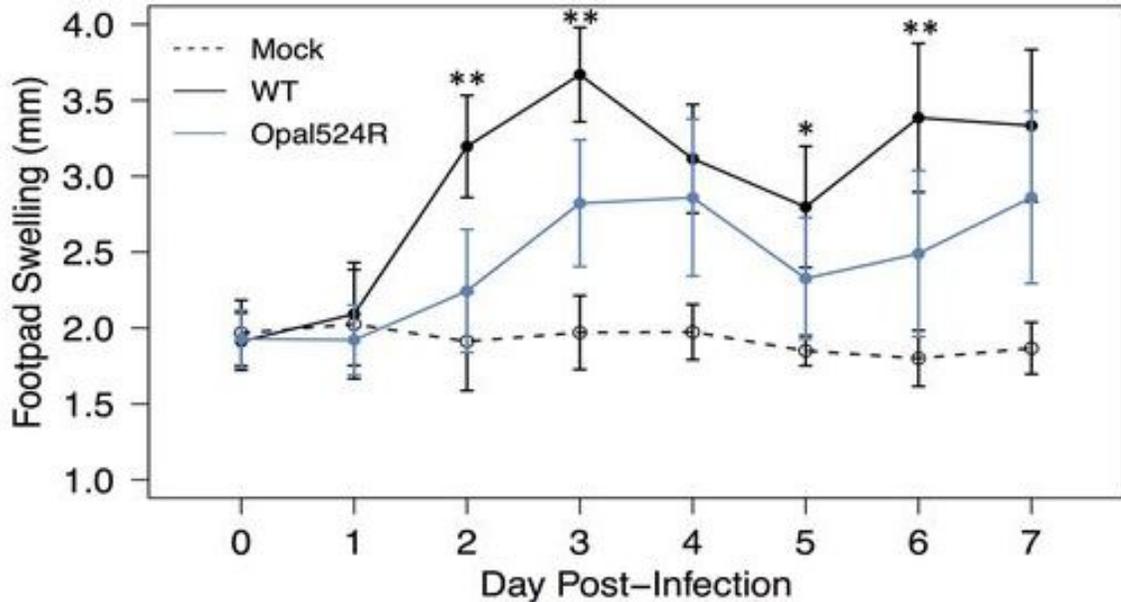


Figure 20. Arginine substitution of the opal stop codon attenuates CHIKV infection in mice. C57BL/6J mice were infected in the rear left footpad with either a mock treatment, 100 PFU of WT CHIKV, or 100 PFU of CHIKV where the opal termination codon has been mutated to AGA arginine (Opal524R). N = 6, 9, and 10, respectively. Pathology is measured as footpad swelling in mm over the next 7 days post infection. Image taken from Jones *et al* 2017¹⁴⁰.

A -1 PRF signal has also been identified in the subgenomic ORF, at the 3' end of the 6K gene. 6K is a small, hydrophobic, cysteine-rich protein that is thought to be incorporated as ion channels into the envelopes of mature alphavirus particles¹⁴². Instances of -1 PRF result in a unique C-terminally extended, 8.4 kDa variant of 6K known as “*Trans-Frame*”, or TF (figure 21)¹⁴³. Notably, downstream protein E1 is not translated following a TF frameshift event. While the precise function of TF is not known, research has shown that it too is anchored by its cysteine residues into the envelope of maturing alphavirus particles¹⁴⁴. Both 6K and TF appear to be crucial for virus budding, as research that limited expression exclusively to either protein led to diminished viral

replication and markedly weaker structural integrity of world world alphaviruses compared to that of wild-type particles^{144,145,146,147}. However, it has also been demonstrated that TF is subject to unique posttranslational modifications compared to 6K¹⁴⁴, suggesting additional functional differences¹⁴⁸. Alphavirus -1 PRF is localized at a universally conserved UUUUUA slippery region. A diverse array of downstream stimulatory structures have been computationally predicted downstream of the slippery site¹⁴⁹, suggesting either that the 6K/TF stoichiometric needs vary greatly between alphaviruses, or that a minimum threshold of the frameshifted product is generated in all cases¹⁴⁸. Crucially, however, the majority of alphavirus signals in the literature have only been characterized in a computationally predictive capacity with a predominate focus on the old world groups.

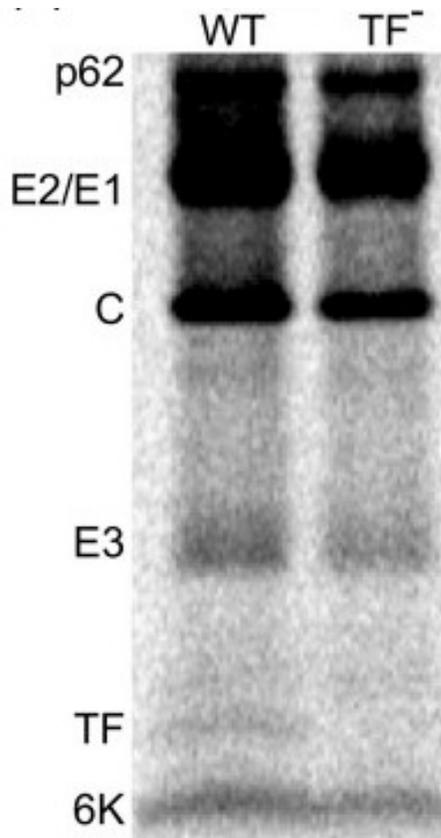


Figure 21. Discovery of alphavirus -1 PRF product TF. A -1 PRF signals was identified at the end of the alphavirus 6K gene. The result of this frameshift is an 8.4 kDa product dubbed “*Trans-Frame*” or TF. SFV-infected BHK cells were labeled with [³⁵S]Met/Cys and lysates were analyzed via SDS-PAGE. Wild type SFV is presented along side a mutant that prevents expression of TF (TF⁻). Band identities are presented on the left. Image derived from Firth *et al.* 2008¹⁴³.

Targeting Translational Recoding in Alphaviruses

This work seeks to investigate the potential application of recoding signals as a therapeutic target against alphaviruses. Here, we performed a thorough genetic and structural investigation of the encephalitic alphavirus -1 PRF signals, which had previously only been characterized through predictive computational methodologies. These signals were capable of promoting efficient -1 PRF in several mammalian cell lines, but were intolerant of mutations that ablated

canonical frameshifting mechanisms. Chemical probing experiments elucidated the presence of hairpin elements immediately 3' of the conserved slippery sequence that served as the minimum required stimulatory structure for efficient -1 PRF. A deeper analysis into the VEEV -1 PRF signal revealed that ablation of the slippery site inhibits *in-vitro* replication kinetics and attenuates neuropathology of the virus in mouse models. Crucially, a followup experiment revealed the presence of neutralizing antibodies in the sera of mice subcutaneously infected with the VEEV -1 PRF mutant, which provided 100% protection against subsequent challenge with a lethal dose of wild type VEEV. These results suggest that -1 PRF mutants could serve as the foundation for live-attenuated vaccine models against VEEV.

The research pipeline used to characterize VEEV -1 PRF signals was then applied to CHIKV, selected as a representative old world alphavirus due to its global prevalence as an emerging pathogen. We identified high levels of sequence conservation in CHIKV TCR and -1 PRF signals consolidated from geographically distinct regions. Similar to the work conducted in VEEV, both signals promoted efficient termination codon readthrough and -1 frameshifting, respectively, in mammalian cell lines. Chemical probing analysis elucidated large hairpin stemloop structures downstream of the TCR opal stop and the -1 PRF slippery site, likely serving as the stimulatory structures for both recoding events. Interestingly, none of the polymorphisms identified between the geographic isolates of CHIKV had any significant effect on the recoding efficiency or higher

order structure of either recoding signal, implying that mutations that ablate either signal's functionality could have broad-range applicability against the virus.

Finally, we consider the possibility that the attenuated pathogenesis observed in the VEEV -1 PRF mutant implies that recoding signals constitute a broadly applicable therapeutic target against all viruses that employ them to regulate gene expression. A collection of recoding signals from alphaviruses, flaviviruses and retroviruses have been genetically and structurally characterized with the research strategies described above. While the completion of this phase extends beyond the scope of this body of work, it will hopefully lay the foundation for extended work into the field of programmed translational recoding signals as a potential Achilles heel to the pathogenic competence of emerging viruses.

Chapter 2: Ablation of programmed -1 frameshifting in Venezuelan equine encephalitis virus results in attenuated neuropathogenicity¹⁵⁰

Joseph A. Kendra^{1*}, Cynthia de la Fuente^{2*}, Ashwini Brahms², Caitlin Woodson², Todd M. Bell², Bin Chen¹, Yousuf A. Khan¹, Jonathan L. Jacobs³, Kylene Kehn-Hall^{2#} and Jonathan D. Dinman^{1#}

¹Department of Cell Biology and Molecular Genetics, University of Maryland, College Park Maryland 20742

²National Center for Biodefense and Infectious Diseases, School of Systems Biology, George Mason University, Manassas, Virginia, 20110

³MRIGlobal, Global Health Security, Rockville, Maryland, 20850

* Authors contributed equally

Work adapted from Journal of Virology, February 2017, **DOI:** 10.1128/JVI.01766-16

Abstract

The alphaviruses, Venezuelan, eastern and western equine encephalitis viruses (VEEV, EEEV, WEEV) are arthropod-borne (+) RNA viruses that are capable of causing acute and fatal encephalitis in many mammals including humans. VEEV was weaponized during the Cold War, and is recognized as a select agent. Currently, there are no FDA approved vaccines or therapeutics for these viruses. The spread of VEEV and other members of this family due to climate change-mediated vector range expansion underscore the need for research aimed at developing medical countermeasures. These viruses utilize programmed -1 ribosomal frameshifting (-1 PRF) to synthesize the viral trans-frame protein (TF), which has previously been shown to be important for neuropathogenesis in the related Sindbis virus. Here, the alphavirus -1 PRF signals were characterized revealing novel -1 PRF stimulatory structures. -1 PRF attenuation mildly affected the kinetics of VEEV accumulation in cultured cells, but strongly inhibited its pathogenesis in an aerosol infection mouse model. Importantly, the decreased viral titers in the brains of mice infected with the mutant virus suggests that the alphavirus TF protein is important for passage through the blood-brain barrier, and/or for neuroinvasiveness. These findings suggest a novel approach toward development of safe and effective live attenuated vaccines directed against VEEV and perhaps other closely related -1 PRF utilizing viruses.

Importance

Venezuelan equine encephalitis virus (VEEV) is as a select agent that has been weaponized. This arthropod-borne (+) RNA virus causes acute and fatal encephalitis in many mammals including humans. There is no vaccine or other approved therapeutic. VEEV and related alphaviruses utilize programmed -1 ribosomal frameshifting (-1 PRF) to synthesize the viral trans-frame protein (TF), which is important for neuropathogenesis. -1 PRF attenuation strongly inhibited VEEV pathogenesis in mice, and viral replication analyses suggest that TF protein is critical for neurological disease. These findings suggest a new approach toward the development of safe and effective live attenuated vaccines directed against VEEV and other related viruses.

Introduction

Venezuelan equine encephalitis virus (VEEV) belongs to the family *Togaviridae*, genus alphavirus, which is further subdivided into old world and new world alphaviruses based on geographical distribution. The old-world alphaviruses include sindbis virus (SINV), Semliki Forest virus (SFV), and chikungunya virus (CHIKV): these generally cause diseases resulting in fever, rash, and arthritic disease. The new world alphaviruses are categorically encephalitic and include eastern and western equine encephalitis viruses (EEEV and WEEV) in addition to VEEV. VEEV, EEEV, and WEEV infections in humans result in 1%, 50-78%, and 3-7% mortality, respectively⁷¹. Mortality rates in horses are overall more severe at 20-80% for VEEV, 70-90% for EEEV, and 3-50% for

WEEV⁷¹. There are no FDA-licensed vaccines or therapeutics for VEEV, EEEV, and WEEV infections, underscoring the need to investigate the molecular biology and pathogenic mechanisms of these pathogens.

The small genomes of RNA viruses limit their coding capacity. In response, they can expand their coding capacity by utilizing alternative splicing¹⁵¹, RNA editing¹⁵², leaky ribosomal scanning¹⁵³ and programmed ribosomal frameshifting (PRF)^{153,154,155}. In -1 PRF, the ribosome encounters a *cis*-acting element, termed the frameshift signal, which causes it to stall on the mRNA over a special 'slippery sequence'. The ribosome is then directed to shift one nucleotide backward, allowing it to synthesize a protein with an alternative C-terminal peptide sequence¹⁵⁴. Typical -1 PRF signals are composed of three modules arranged in the following 5' to 3' direction: a heptameric "slippery site", a short spacer, and a strong RNA secondary structural element, often an mRNA pseudoknot¹¹⁵. Multiple RNA viruses including West Nile virus, human immunodeficiency virus, and severe acute respiratory syndrome coronavirus utilize -1 PRF^{153,155,130}.

Alphaviruses were bioinformatically predicted to use -1 PRF to generate the trans-frame protein (TF), which shares its N-terminal region with the 6K protein¹⁴⁹. Mass spectrometric analyses were used to confirm TF production by multiple alphavirus family members, including SFV, SINV, and CHIKV^{149,147}. A Sindbis virus (SINV) based molecular genetics study examining the consequences of altering the production, size, or sequence of TF revealed reduced levels of SINV viral release from both mammalian and mosquito cells,

without influencing genomic replication, specific infectivity, or migration of the envelope protein to the cell surface, suggesting that TF was likely involved in viral assembly¹⁴⁷. Moreover, SINV with mutated TF was attenuated in a SINV neuropathogenic mouse model. To date there have been no studies on the role and importance of TF in the new world alphaviruses. The current study characterizes the frameshifting signals in new world alphaviruses (VEEV, EEEV, and WEEV). Standard molecular genetics analyses are consistent with the prediction that these viruses utilize -1 PRF to synthesize their TF proteins. VEEV harboring a silent protein coding mutation that attenuates -1 PRF activity (VEEV_{PRFm}) displayed mildly decreased viral replication kinetics *in vitro*. Strikingly however, mice infected with VEEV_{PRFm} showed dramatically increased survival and decreased clinical signs of disease as compared to mice infected with WT VEEV. These data demonstrate that frameshifting is a critical mechanism utilized by alphaviruses to encode TF, which is important for pathogenesis, and suggest -1 PRF attenuation as a general strategy for exploring rational development of live attenuated vaccines.

Materials and Methods

Computational Prediction of Viral PRF Signals

Accession numbers of virus genomes predicted to contain programmed -1 PRF signals were imported from NCBI (<http://www.ncbi.nlm.nih.gov>) into the Dinman lab's -1 PRF database (<http://prfdb.umd.edu>)¹⁵⁶. PRFdb search algorithms were used to identify *N NNW WWW* slippery site sequences in the genome, as well as the alternate polypeptide sequences consequent to a

recoding event. Feynman diagrams of the predicted downstream stimulatory structures were generated using folding algorithms NUPAK¹⁵⁷ and RNAfold¹⁵⁸ as guides for subsequent cloning.

Dual reporter plasmid construction and bacterial transformation

All dual reporter plasmids were adapted from the dual luciferase read-through control plasmid, pLuci (pJD175f)¹⁵⁹. The multiple cloning site of pJD175f was digested at *Sal* I and *Sac* I restriction sites using the respective Fast-DigestTM restriction enzymes (Thermo-Fisher). Restriction digest products were separated via 1% agarose gel electrophoresis, visualized by ethidium bromide staining and UV detection and isolated via NucleoSpin® Gel and PCR Cleanup Kit (Macherey-Nagel). Experimental plasmid inserts containing putative frameshift signals were generated as gBlocks (IDT) using complementary oligonucleotides and cloned into the linearized plasmid backbone via an In-Fusion® HD Cloning kit (Clontech Laboratories). Specifically, the following virus-derived sequences were cloned into pJD175f: a 126 nt EEEV-derived sequence beginning at nt 9961; a 117 nt VEEV-derived sequence beginning at nt 9970; and a 92 nt WEEV-derived sequence beginning at nt 9927. Assembly products were transformed into StellarTM competent *E. coli* cells (Clontech Laboratories) and spread onto LB agar plates containing 50 µg/ml carbenicillin. Positive clones were verified by DNA sequencing (Genewiz®). Plasmid and primer sequences available upon request. Additionally, a series of frameshift reporter negative controls based on these three clones were constructed using oligonucleotide directed site specific. These consisted of (a) insertion of a 0-frame UAA

termination codon immediately 5' of the slippery sites in each plasmid (denoted 5'Ter), (b) silent coding mutagenesis of the U UUU UUA slippery sites to G UUC UUG (named ssM), and (c) a UAA termination codon in the -1 frame inserted after the viral sequences (denoted 3' -1 Ter).

Cell Culture

HeLa (# CCL-2), U-87 MG (# HTB-14), and Vero (# CCL-81) cells were obtained from ATCC (Manassas, VA). Cells were cultured in Dulbecco's Modified Eagle Media (DMEM) supplemented with 1% L-glutamine and 10% heat inactivated fetal bovine serum (FBS) and maintained in a humidified 37°C incubator with 5% CO₂. Cycling AP-7 rat olfactory bulb neuronal cells were cultured in DMEM supplemented with 1% L-glutamine and 10% FBS at 33°C and 10% CO₂¹⁶⁰. For *in vitro* differentiation, cycling AP-7 cells were plated in 6-well plates at a seeding density of 2.0 x 10⁵ cells per well. The next day, media was changed to DMEM containing 1% L-glutamine, 1% penicillin/streptomycin, 10% FBS, 2 µg/mL insulin, 40 µM dopamine hydrochloride, and 100 µM ascorbic acid. Cells were maintained at 39°C and 5% CO₂ for 5 to 6 days before infection.

Plasmid and siRNA transfection

HeLa or U-87 MG cells were seeded 0.6 x 10⁵ cells per well into 24 well plates in 0.5 ml of DMEM enhanced with 1% L-glutamine, 15% FBS and 1X penicillin/streptomycin at 37° C and 5% CO₂. Following a 24 hour incubation, control and experimental dual luciferase reporter plasmids were then transfected into cells using the Fugene HD[®] transfection kit at a 3:1 transfection reagent to

DNA ratio. siRNAs directed against human DGCR8, Exportin 5, Argonaute 1, Argonaute 2 or scrambled sequences were transfected into U87-MG or HeLa cells as previously described¹⁶¹.

Assays of programmed -1 ribosomal frameshifting

Frameshifting efficiency of the experimental reporter plasmids was assayed as previously described¹⁶² using a Dual Luciferase Reporter Assay System kit (Promega). Twenty-four hours post transfection, cell culture media was aspirated and cells were rinsed twice with 1x phosphate buffered saline before disruption with 1X passive lysis buffer. Cell lysates were assayed in triplicate in a 96 well plate. Firefly and *Renilla* luciferase activity was quantified using a Glowmax[®] 96 Microplate Luminometer (Promega).

Chemical modification analysis of -1 PRF promoting RNA structural elements

mRNAs were structurally assayed using SHAPE^{163,164}. DNA templates for mRNA secondary structure analysis were generated by PCR amplification using a Dreamtaq[™] DNA Polymerase kit (ThermoFisher[™]). Forward and reverse primers for the *Renilla* and firefly regions on reporter plasmid pJD175f were used to amplify the inserted PRF sequence and attach a T7 promoter sequence and Kozak sequence at the 5' end of the amplicons. Amplicons were isolated by agarose gel purification and *in vitro* transcribed RNA was generated using a T7 MEGAscript kit[®] (Life Technologies[™]). *In vitro* transcripts were purified using a MEGAclean[®] cleanup kit (Life Technologies[™]) and the quality of the RNA transcripts were assessed by agarose gel electrophoresis. Nine pmols of the

PRF RNA templates were denatured at 65°C for four minutes and refolded at 37°C for 20 mins in 5x folding buffer (400 mM Tris HCL pH 8.0, 800 mM NH₄Cl, 55 mM MgOAc). Probing of flexible bases in the RNA transcripts was conducted through N-methylisatoic anhydride (NMIA) acylation of unprotected 2'-OH groups. Primer extension with $\gamma^{32}\text{P}$ radiolabeled probes and reverse transcription was carried out as reported elsewhere^{163,164}. cDNA products were separated through an 8% urea gel and visualized on a phosphoimager. Visual clarity of gel images was adjusted on Adobe Photoshop Lightroom 5.

3-dimensional structural modeling of the tVEEV -1 PRF stimulatory element

All-atom models were generated using the MC-Fold and MC-sym pipeline programs¹⁶⁵. Initially, the RNA sequence of tVEEV was imported into MC-Fold to generate a series of secondary structures. Over 1000 structures were explored in total and the top 20 were selected for further consideration based on their energetic scores. Among these, the one whose secondary structure was most consistent with the SHAPE experiments was submitted to MC-Sym for 3D modeling. 200 structures were subsequently generated, and these were subjected to energy minimization and solvent refinement, yielding 15 best models. The highest scoring model was selected to represent the predicted 3-dimensional structure of tVEEV shown in Figure 23E.

Introducing -1 PRF mutations into infectious VEEV clones

Synonymous substitutions were introduced by overlapping PCR extension using standard techniques to disrupt the -1 PRF signals in the TC83 and TrD

genomes. The silent slippery site mutations consisted of the following changes - T9964G, T9967C, A9970G – to change the U UUU UUA slippery sites to G UUC UUG within the pTC83 (VEEV) and pV3000 (TrD) plasmids. There are five amino acid changes between the TC83 strain as compared to TrD which all lie within the structural coding region: four in E2 (K7N, H85Y, T120R, V192D, T296I) and one in E1 (L161I^[105,166]). Furthermore, the V3000 clone of TrD utilized for this study also encodes for two additional changes within E2 (one previously published, N239I¹⁶⁷; and one unpublished, E323G). All plasmid constructs were verified by restriction enzyme digestion and sequencing. Plasmid and primer sequences are available upon request.

VEEV viral stocks

Viral stocks were produced from electroporation of *in vitro* transcribed viral RNA generated from either the pTC83 plasmid¹⁰⁵, the pV3000 plasmid (TrD¹⁶⁸) or PRF mutants pTC83_{PRFm} and pV3000_{PRF} plasmids. In brief, the viral cDNA was linearized with restriction enzyme and then purified using the MinElute PCR Purification kit (Qiagen) according to manufacturer's directions. Capped RNAs were synthesized using the MEGAScript kit (Invitrogen) with a 2:1 ratio of cap analog [^{m7}G(5')ppp(5')G NEB] to GTP and treated with DNase I supplied with the kit. RNA was then isolated with the RNeasy Mini kit with a second DNase I on-column digestion (Qiagen). The RNA integrity and concentration were determined by gel electrophoresis and absorbance at 260nm, respectively. *In vitro* transcribed viral RNAs were electroporated into BHK-J cells utilizing a 2mm gap cuvette (BTX ECM 630 exponential decay wave electroporator; Harvard

Apparatus, Holliston, MA). After trypsinization, cells were washed twice and resuspended in cold Dulbecco's phosphate-buffered saline without $\text{Ca}^{2+}/\text{Mg}^{2+}$ (D-PBS; RNase-free) at 1.25×10^7 cells/ml. An aliquot of the cell suspension (400 μl) was mixed with 1 μg of RNA transcripts, placed into the cuvette, and pulsed once at 860V, 25 μF capacitance, and 950 Ω resistance. Cells were allowed to recover for 5 min at room temperature and resuspended in complete minimal essential medium (MEM; Gibco-Invitrogen). Cells from three replicate electroporations were plated in three 75 cm^2 culture flasks for virus production. On the next day (~12 hours post electroporation [hpe]), transfection media was replaced with fresh MEM. Media supernatants were harvested at several timepoints, pooled and stored at 4°C. After the last collection, supernatants were then filtered (0.2 μM), aliquoted, and stored at -80°C. Viral titers were determined by plaque assay on Vero cells.

Analysis of viral kinetics

VEEV RNA replication as well as infectious viral titers were determined within Vero cells. Vero cells (seeded in 12-well plates) were infected at a multiplicity of infection (MOI) of 1 for 1 hour. After the inoculum was removed, cells were washed twice with D-PBS, and cultured further in complete media. At 3, 6, 9, 18 and 24 hours post infection (hpi) supernatants were collected and cells were washed once with D-PBS and lysed in Trizol-LS (TC83 viruses) or RLT buffer (RNeasy kit; TrD viruses). Both sample sets were stored at -80°C until they could be further processed. Infectious virus titers were determined by plaque assay on Vero cells. Differentiated AP-7 (dAP-7) cells were infected with either

TrD or TrD_{PRFm} at an MOI 1.0 and maintained at differentiation conditions throughout the experiment. Supernatants and RNA lysates were harvested at 3, 9, 18, 24, and 48hpi.

Total cellular RNA was isolated from Trizol-LS lysates utilizing Direct-zol™ RNA MiniPrep kit (Zymo Research, Irvine, CA) or RNeasy Mini kit (Qiagen) according to manufacturer's directions. RNA quality and concentration was analyzed by gel electrophoresis and absorbance at 260nm, respectively. For TrD virus containing samples, High-Capacity RNA-to-cDNA™ Kit (Thermo Fisher) was used to generate complementary (c)DNA. Quantification of viral RNA was determined by quantitative reverse transcription PCR (qRT-PCR; TC83) or qPCR (TrD) using the StepOnePlus Real-Time PCR System (Applied Biosystems). Primer-pairs (forward 5' TCTGACAAGACGTTCCCAATCA 3', reverse 5' GAATAACTTCCCTCCGACCACA 3') and Taq-man probe (5' 6-carboxyfluorescein-TGTTGGAAGGGAAGATAAACGGCTACGC-6-carboxy-N,N,N',N'-tetramethylrhodamine-3') for nucleotides 7931–8005 of VEEV TC-83 were originally described previously¹⁶⁹. TC83 reactions were assembled using the RNA UltraSense™ One-Step Quantitative RT-PCR System (Invitrogen) and absolute quantification was calculated based on the threshold cycle (Ct) relative to the standard curve. For TrD reactions, TaqMan® Gene Expression Master Mix (Thermo Fisher) was used and relative quantification to the 3 hpi timepoint was calculated based on the 2-DDCt method using 18S rRNA¹⁷⁰. Undetermined Ct values were given the value of 40 for analysis.

Animal experiments

Six to eight week old female BALB/c mice were obtained from Harlan Laboratories. Groups of 35 mice were infected with VEEV-TrD or VEEV-TrD_{PRFm} using Biaera's AereoPm System, whole body chamber and a three jet Collision nebulizer. They were exposed to 1×10^5 pfu/ml of VEEV-TrD or VEEV-TrD_{PRFm} for 10 min. Hank's Balanced Salt Solution (HBSS) plus 1% FBS was used for viral aerosol. Ten animals from each group were observed for survival over the course of twenty-one days. Five animals from each group were euthanized on Days 2, 4, 6, 8 and 10 post infection to determine the kinetics of disease in the mouse system. Serum, spleen, and brain were collected from each animal. Organs were homogenized using Omni Bead Ruptor 4 (Omni International) and then centrifuged at 10,000 rpm for 10 min. Supernatants were analyzed by plaque assays to determine viral titers. All VEEV TrD experiments were performed in animal bio-safety level 3 (ABSL-3) facilities, in accordance with the National Research Council's *Guide for the Care and Use of Laboratory Animals*¹⁷¹ and under GMU IACUC protocol number 0331.

Statistical analysis

Frameshifting efficiencies and statistical analyses were calculated as previously described¹⁷². Frameshifting assays were independently repeated a minimum of three times as technical triplicates. Data was normally distributed and statistical analyses were conducted using a Student's t-test. Statistical analysis for viral kinetics was performed using Prism 6 (GraphPad). Multiple unpaired t-test analysis of the titer and RT-qPCR datasets with the Holm-Sidak

correction was applied with the assumption that all comparisons had the same scatter.

Results

Alphavirus-derived sequences promote efficient levels of -1 PRF

The alphavirus genomes consist of a single (+) RNA that harbors two open reading frames (ORFs) (Fig. 22A). The 5' ORF encodes four non-structural proteins, while five structural proteins are encoded by the 26S subgenomic RNA. These two ORFs are translated as polyproteins which are proteolytically processed into the mature proteins (reviewed in ¹⁷²). In two old world alphaviruses, SINV and CHIKV, -1 PRF events in the 6K gene result in production of 8.4 kDa *trans*-frame (TF) protein¹⁴⁷. Similar signals are predicted to be located near the 3' end of the sequence encoding the 6K protein in the New World alphaviruses¹⁴³, and the 8.4 kDa TF protein is produced consequent to -1 PRF events. Strategies for cloning the predicted -1 PRF signals from EEEV, VEEV and WEEV into dual luciferase reporters¹⁵⁹ were determined by identifying their conserved 5' slippery sites and *in silico* RNA folding of 3' sequences¹⁵⁶. Information pertaining to these clones is shown in Figure 22B. Figure 22C shows that all of these sequences promoted efficient levels of -1 PRF in both HeLa and U-87 MG cell lines. These findings are in general agreement with prior measurements of VEEV and EEEV frameshifting efficiencies measured by dual-luciferase assays¹⁴⁹. To rule out the possibilities that these sequences harbor IRES activity, cryptic promoters or cryptic splice sites, a series of mutants harboring slippery site inactivation mutations (ssM), in-frame termination codons

prior to the slippery sites (5' Ter), and -1 frame termination codons 3' of the virus-derived sequences (3' -1 Ter) were tested. Dual luciferase reporter assays ascertained that all of these mutations significantly reduced apparent -1 PRF activity (Figure 22D), indicating that these sequences encode *bona fide* -1 PRF signals. A prior study demonstrated that -1 PRF can be naturally regulated in cells by microRNAs (miRNAs)¹⁶¹. To investigate this possibility, miRNA processing or export was inhibited by siRNA knocking down expression of DGCR8, Exportin5 or Ago2 in U87-MG cells, and the effects on VEEV-mediated -1 PRF were subsequently assayed. As positive controls, -1 PRF mediated by the HIV-1 and CCR5 -1 PRF signals were also assayed. Results from these experiments suggest that, unlike HIV-1 and CCR5, VEEV -1 PRF is not regulated by miRNAs in this cell type (Figure 22E). Similarly, siRNA knockdown of Ago1 in HeLa cells did not significantly affect -1 PRF promoted by the VEEV, EEEV, or WEEV sequences in HeLa cells (data not shown).

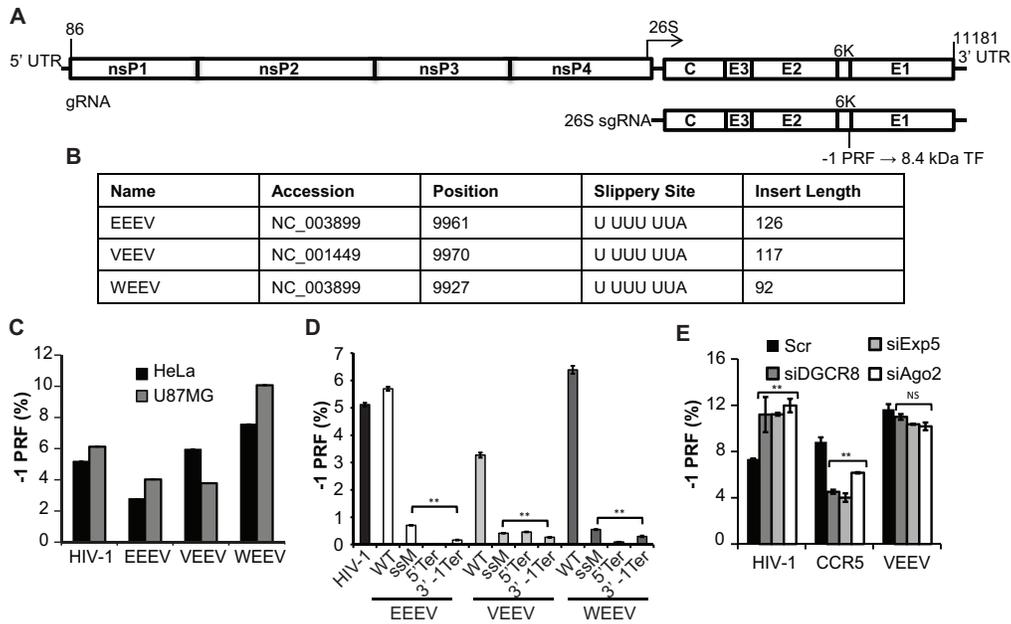


Figure 22. Identification and monitoring of -1 PRF in the alphaviruses. (A) General schematic of the alphavirus genome and subgenomic transcript. The specific nucleotide numbers are given for the boundaries of the VEEV untranslated regions (UTRs). This family of viruses harbors -1 PRF signals toward the 3' end of the 6K structural protein-encoding mRNA. The resulting frameshifting product is an 8.4-kDa protein called the *trans*-frame (TF) protein. gRNA, genomic RNA; sgRNA, subgenomic RNA. (B) Accession numbers of the sequences from which the predicted -1 PRF signals were cloned, nucleotide positions of the first base of the predicted slippery sites, the slippery site sequences (the 0 frame is indicated by spaces), and the lengths of inserts (in numbers of nucleotides) initially tested for -1 PRF activities. (C) The predicted -1 PRF signals derived from EEEV, VEEV and WEEV were cloned into the dual-luciferase reporter plasmid pJD175f, and their ability to promote efficient levels of frameshifting was measured in both HeLa cells and U-87 MG astrocyte cells. The -1 PRF signals from HIV-1 was employed as a positive control. (D) Site-directed mutagenesis was utilized to genetically validate the -1 PRF activities of the EEEV-, VEEV-, and WEEV-derived sequences in U-87 MG cells. WT, wild type; ssM, silent mutations of the slippery sites; 5' Ter, 0-frame termination codons introduced upstream for the slippery sites; 3'-1 Ter, termination codons inserted in the -1 reading frame '3 of the virus-derived sequences. (E) miRNA processing or export was inhibited by siRNA knockdown of DGCR8 (siDGCR8), Exportin 5 (siExp5), or Argonaute 2 (siAgo2) in U-87 MG cells, and the rates of -1 PRF promoted by the indicated sequences were monitored. Control samples were transfected with siRNAs harboring scrambled (Scr) sequences. For assays of -1 PRF, a minimum of three or more biological replicates was performed in triplicate until biological significance was achieved, as previously described¹⁷². Bars represent standard errors of the means. **, $P < 0.01$; NS, not significant.

Efficient alphavirus -1 PRF is stimulated by stem-loop mRNA structural elements

In classic -1 PRF signals, RNA structural elements located immediately 3' of heptameric slippery sites serve as kinetic traps to stimulate frameshifting by stalling elongating ribosomes over the slippery site¹¹⁵. Typically, these are mRNA pseudoknot structures, although -1 PRF can be stimulated by simple stem loop structures as well¹⁷³. Selective 2' hydroxyl acylation and primer extension (SHAPE)¹⁷⁴ was employed to characterize the nature of the EEEV, VEEV and WEEV -1 PRF stimulating elements (Figure 23A-C). Analysis of the SHAPE data revealed the presence of tandem stem-loops rather than RNA pseudoknot structures for all three of the viral sequences. Dual luciferase reporter assays of 3' truncation mutants (tEEEV, t2EEEV, tVEEV and tWEEV) revealed that only the slippery site proximal stem loops are required to promote efficient rates of -1 PRF (Figure 23D). Guided molecular dynamics simulations of the tVEEV element suggests that this folds into a novel V-shaped structure comprising three stems labeled Sa, Sb and Sc (Figures 23B, 23E). How this may promote efficient levels of -1 PRF is discussed below.

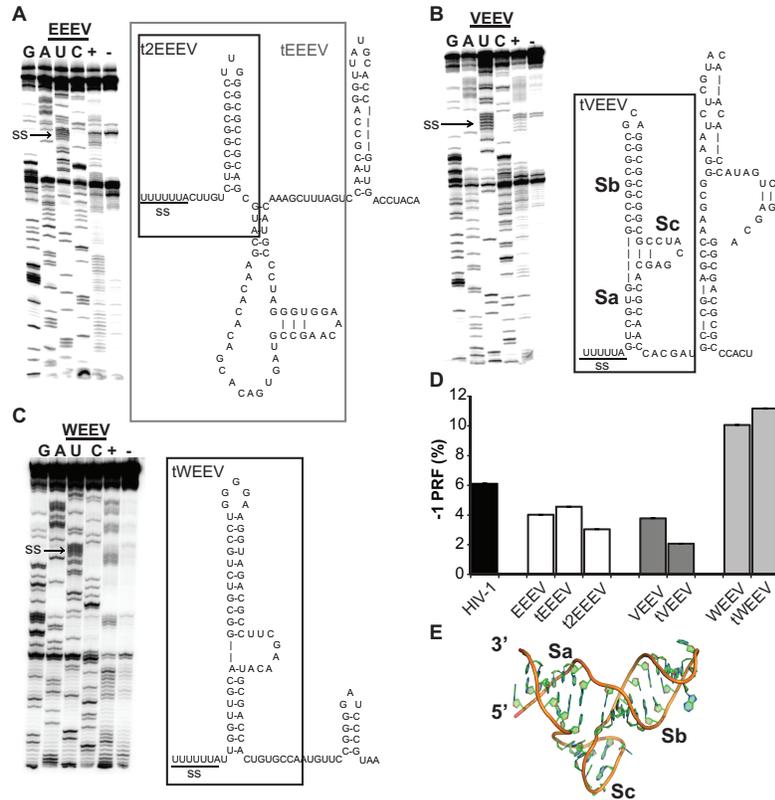


Figure 23. Structural analysis of alphaviral -1 PRF signals. The EEEV (A), VEEV (B), and WEEV (C) -1 PRF signals were chemically resolved using selective 2'-hydroxyl acylation and primer extension (SHAPE). RNAs for the -1 PRF signals of each alphavirus were transcribed from their corresponding reporter plasmids, RNA samples were treated with *N*-methylisatoic anhydride (NMIA), primer extension reactions were carried out, the reaction mixtures were separated through 8% urea-PAGE, and ³²P cDNA products were visualized using a Typhoon phosphorimager (left). (Right) RNA structures deduced from the SHAPE data. The 3' truncation mutants are indicated in the boxed regions. (D) Identification of the minimal structures required to promote efficient levels of -1 PRF (defined as >1%). The truncated sequences in panels A. to C. were cloned into pJD175f dual-reporter plasmids and assayed for their ability to promote efficient levels of frameshifting in U-87 MG cells. For assays of -1 PRF, a minimum of three or more biological replicates was performed in triplicated until statistical significance was achieved, as previously described¹⁷². Bars represent standard errors of the means. (E) Predicted three-dimensional structure of the tVEEV -1 PRF stimulatory element.

Ablation of -1 PRF mildly decreases VEEV virus production in vitro

Prior studies have suggested that viral -1 PRF signals have evolved to promote frameshifting at very precise rates, and that changes in -1 PRF efficiencies have detrimental effects on virus propagation^{125,128,175}. To test the importance of -1 PRF on virus propagation in cultured cells, a silent protein coding change was introduced into the VEEV infectious clones for the TC83 vaccine strain and the highly pathogenic Trinidad Donkey (TrD)¹⁶⁸ strain to create pTC83_{PRFm} and pTrD_{PRFm} (Figure 24A). Vero cells were infected with TC83, TC83_{PRFm}, TrD, or TrD_{PRFm} as described in the Materials and Methods, and samples were collected for analysis at 3, 6, 9, 18 and 24 hours post infection (hpi). Surprisingly, although statistically significant effects were observed at some of the later time points, ablation of -1 PRF within the TC83 backbone had minimum effects on virus titers and viral RNA accumulation (Figure 24B and 24E). Similar results were observed with C6/36 mosquito cells (data not shown). Disruption of the -1 PRF signal within TrD resulted in decreased viral titers (~1.5 logs) starting at 9 hpi (Figure 24F). Viral RNA levels were not affected until 18 hpi, which is consistent with a defect in viral assembly (Figure 24C). Replication analysis was also performed in dAP-7 rat neuronal cells, which represent a more relevant physiological model of VEEV infection. Viral RNA levels did not differ between TrD and TrD_{PRFm} (Figure 24D), but a decrease in viral titers was observed with TrD_{PRFm} (Figure 24G), supporting the notion that TF protein plays a role in viral assembly. dAP-7 cells have exited the cell cycle whereas Vero cells are cycling, which may be one reason for the observed differences in viral RNA

levels between the two cell types. With several substitutions within E2 and E1 between TrD and TC83 (as noted within the Materials and Methods) this may indicate that there is less dependence on the -1 PRF signal for the attenuated TC83 strain *in vitro*.

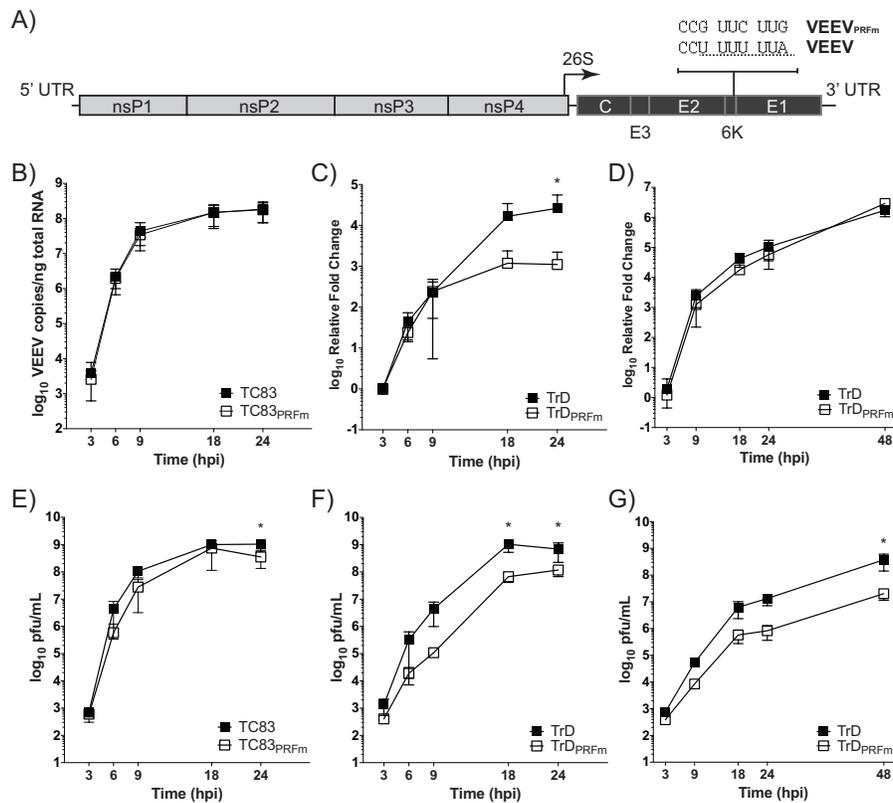


Figure 24. Ablation of -1 PRF decreases VEEV release *in vitro*. (A) Schematic diagram indicated silent coding nucleotide substitutions ablating the -1 PRF signals in the VEEV TC83 and TrD infectious clones (VEEV_{PRFm}). (B to G) Replication kinetics of the TC83 and TC83_{PRFm} in Vero cells (B and E), TRD and TRD_{PRFm} in Vero cells (C and F), and TrD and TrD_{PRFm} in dAP-7 cells (D and G) infected at an MOI of 1. (B to D) Total RNAs were isolated from cells, and the absolute quantities for (TC83-infected cells) and relative quantities (for TrD-infected cells) of VEEV genomic copies were determined by qRT-PCR. For TrD-infected cells, relative fold change values were calculated using the number of genomic copies of the respective virus detected at 3 hpi. Data are plotted as means with standard deviations ($n = 9$ experiments for TC83-infected cells, $n = 6$ experiments for TrD-infected Vero cells, $n = 6$ experiments for TrD-infected dAP-7 cells). For all data sets, a multiple unpaired t test analysis with the Holm-Sidak correction was applied to each time point. Bars represent standard deviations. *, $P < 0.05$.

Ablation of -1 PRF strongly attenuates VEEV pathogenesis

Prior studies in which -1 PRF was ablated in flaviviruses revealed that the NS1' frameshift PRF product is required for neuroinvasion and replication in both avian and insect hosts^{176,177}. Similarly, deletion of the 6K gene reduced the pathogenesis of the Ross River alphavirus in mice¹⁷⁸. To determine the importance of the -1 PRF signal for VEEV pathogenesis, mice were exposed to 1×10^5 pfu/ml of VEEV-TrD or VEEV-TrD_{PRFm} for 10 min via aerosol route. Two groups of mice were followed for 21 days in order to assess survival, while others were sacrificed over the course of infection to assess viral kinetics *in vivo*. Kaplan-Meier survival analysis shows that ablation of -1 PRF had a strong negative effect on VEEV-induced mortality (Figure 25A). VEEV TrD infected mice succumbed to infection beginning at 8 days post infection (dpi) with all mice succumbing by 13 dpi. In contrast, 70% of VEEV TrD_{PRFm} infected mice survived infection. Weight loss (Figure 25B) and clinical symptoms of disease (Figure 25C) were less severe and delayed in VEEV TrD_{PRFm} infected mice as compared to VEEV TrD infected mice. In parallel experiments, virus titers were monitored in the serum, spleens and brains of infected mice every two days for 10 days total. In all of the VEEV TrD infected mice, virus was detected in the blood and spleen early after infection (2 and 4 dpi) and cleared by 6dpi (Figure 25D). In contrast following infection with VEEV TrD_{PRFm}, virus was only detectable in the blood and spleen in 50% of the mice at 2 dpi, and in only 60% of the mice at 4 dpi. Virus was also detected in the spleen at 6 dpi in 80% of the VEEV TrD_{PRFm} infected mice. Plaque assays of brains revealed the presence of high levels of virus in the

VEEV TrD infected mice at all time points tested. In contrast, virus was not detectable in the VEEV TrD_{PRFm} infected mice until 4 dpi, and was cleared in 80% by 10 dpi. These results indicate that the dissemination of VEEV TrD_{PRFm} is altered *in vivo* resulting in less viral replication within the brain and overall decreased pathogenesis.

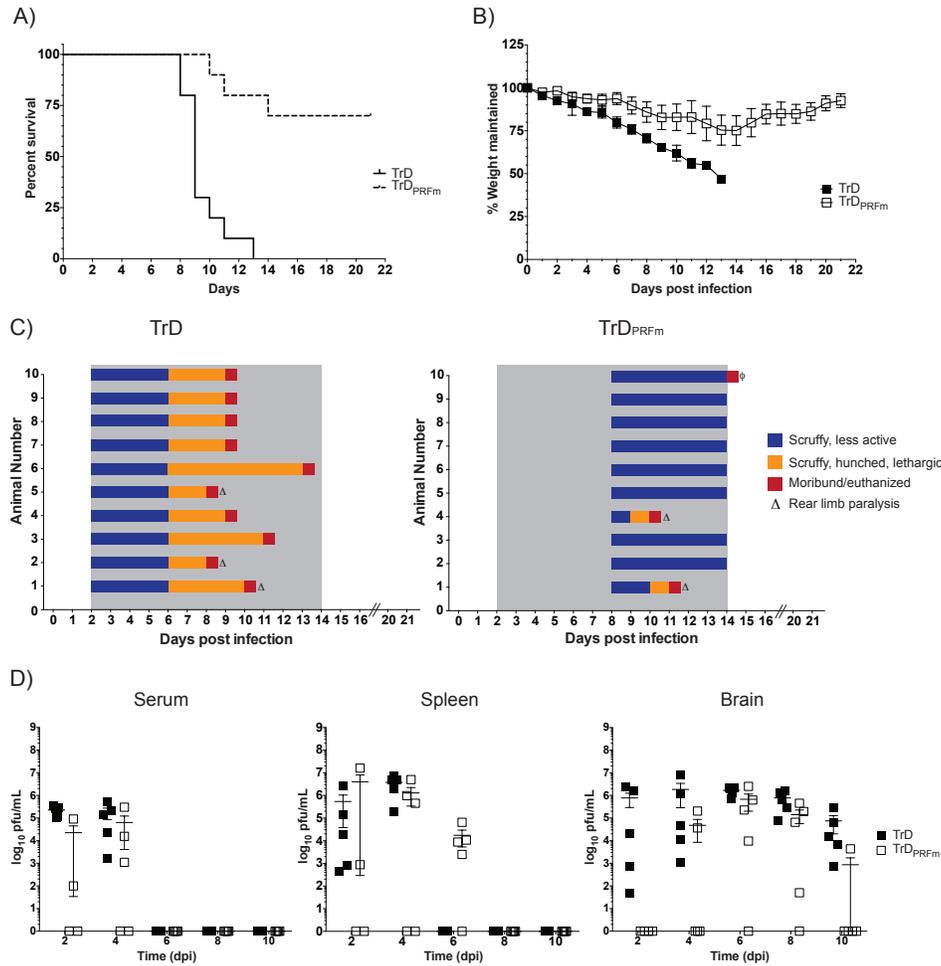


Figure 25. Ablation of -1 PRF strongly attenuates VEEV pathogenesis. (A) BALB/c mice were infected with VEEV TrD or VEEV TrD_{PRFm} by aerosol exposure. Animals were monitored for 21 days post-challenge and survival curves were determined. The data plotted represent those for 10 animals per group. (B) Mice were monitored for weight loss daily over 21 days. The percentage of weight maintained (relative to the starting weight) was determined. The data plotted represent the mean values and standard deviations for 10 animals per group. (C) Mice were also monitored at least daily for clinical symptoms of disease over 21 days. Data are plotted per animal per day. The gray-shaded area indicates the time frame when clinical disease was observed in VEEV TrD-infected mice. ϕ , one animal had to be euthanized due to self-mutilation. Necropsy indicated no signs of disease in this mouse. (D) Mice were infected as described in the legend to panel A and were sacrificed at 2, 4, 6, 8, and 10 dpi. Brain, spleen and serum were harvested. Viral titers were determined by plaque assays. The data plotted represent means and standard errors of the means for five animals per condition. Filled and open squares, VEEV TrD and VEEV TrD_{PRFm}, respectively. Samples without detectable plaques were plotted as 1 PFU/ml.

Discussion

'Canonical' -1 PRF signals are described as being composed of three elements arranged from 5' to 3' in the following order: a heptameric *NNW* *WWW* slippery site, a vaguely defined short spacer segment, and a downstream structural element which is typically a variation of an mRNA pseudoknot (reviewed in ¹⁷⁹). Examples in which efficient -1 PRF is stimulated by stem-loop structures are rare, the most well-documented of which is the HIV-1 frameshift signal^{110,180,181}. However, while the HIV-1 stem loop is sufficient to promote efficient -1 PRF, local genomic RNA secondary structure influences -1 PRF efficiency¹⁸² suggesting that more complex and dynamic interactions between the ribosome and -1 PRF stimulating RNA structural elements are involved. The novel predicted tVEEV structure (Figure 23E) may provide an explanation for how these structural elements may stimulate -1 PRF. Here, the internal loop results in a structure in which two of the stems (labeled Sb and Sc) abut one another. We suggest that, as this structure enters the ribosomal mRNA entry tunnel, Sb and Sc are compressed toward one another, bringing their phosphodiester backbones in close proximity. At a certain point, the ensuing charge repulsion may drive the decompression or "opening-up" of the structure, resulting in the backward movement of the mRNA relative to the ribosome, i.e. a -1 frameshift. We envision that emerging computational and single-molecule experimental platforms can be applied to test this novel mechanistic model of -1 PRF. Interestingly, a prior study indicated that a VEEV deletion mutant harboring only 32 nt of sequence (predicted to retain the Sb structure while eliminating the

Sa and Sc structural elements) stimulated -1 PRF approximately twice as well as VEEV with the native sequence¹⁴⁹. As discussed in that work, these findings indicate that a diverse array of 3' RNA structures are capable of promoting efficient levels of -1 PRF.

Programmed -1 ribosomal frameshifting was first discovered in retroviruses, where it directs the synthesis of Gag-pol polyproteins¹⁸³. Subsequent studies using retroviruses¹⁸⁴ and totiviruses¹²⁴ demonstrated that changes in -1 PRF efficiency affects virus production. From this, a bioeconomics model emerged in which -1 PRF rates are optimized to maximize virus particle assembly by ensuring the synthesis of the correct stoichiometric ratios of the structural Gag-derived proteins to the enzymes encoded by the Pol open reading frame (ORF) (reviewed in ¹⁷⁵). These findings engendered interest in targeting -1 PRF for antiviral therapeutics (reviewed in ¹⁸⁵). In parallel, early studies also examined -1 PRF signals in virus families where they do not occur between open reading frames encoding structural and enzymatic open reading frames (ORFs), e.g. in coronaviruses and luteoviruses. The finding that the Barley Yellow Dwarf virus uses -1 PRF as a developmental switch, from initial translation of non-structural proteins to those involved in viral genome replication, represented an expansion of our understanding of the utility of this molecular mechanism¹⁸⁶. Similarly, research in coronaviruses revealed that -1 PRF also serves as a switch, in this case from expression of immediate early non-structural proteins that are implicated in modulating the innate immune response, to the next developmental step of the viral program, expression of viral replication machinery

(reviewed in references^{127,187}). Nonetheless, altering -1 PRF efficiency in SARS-CoV severely impacted its infectivity in tissue culture, reinforcing the idea of -1 PRF as an antiviral therapeutic target¹²⁸. In flaviviruses the viral (+) RNA genome encodes a single ORF in which the structural proteins are encoded by the 5' third of the genome, and the 3' two-thirds encode the non-structural proteins. In these viruses, the location of the -1 PRF signal in the first non-structural gene (NS1) has been proposed to ensure production of large amounts of structural proteins for virus particle assembly, and smaller amounts of the non-structural proteins¹⁵⁵. Interestingly, lower rates of -1 PRF correlate with decreased pathogenicity in West Nile Virus¹⁵⁵, and production of the NS1' frameshift product is critical for neuroinvasiveness in West Nile and Japanese encephalitis viruses^{176,188}. In these viruses, the NS1' protein is thought to be important for virion assembly^{176,189}. Additionally, -1 PRF has now been demonstrated to be used to control the expression of a large fraction of cellular genes in eukaryotes by functioning as to control mRNA stability (reviewed in¹⁹⁰). Thus, we suggest that -1 PRF is an ancient, basic biological regulatory mechanism that has been evolutionarily selected for numerous end uses. On a final note, the observation that -1 PRF in alphaviruses does not appear to be regulated by host cellular miRNAs is not surprising from an evolutionary point of view. Like many RNA viruses, their evolution has followed a generalist strategy favoring the ability to produce acute, high yield infections in a wide range of host organisms as opposed to viruses that have co-evolved with a single or closely related host species over a long period of time. Thus, a HIV-1 with its long co-evolutionary history with the Great Apes

would be expected to utilize host-encoded miRNAs, while the alphaviruses would not.

While related to flaviviruses, the alphavirus genomes are arranged such that the non-structural proteins are located in the 5' ORF, while the structural protein genes are in a separate 3' ORF, and are expressed from the 26S subgenomic RNA (see Figure 22A). In alphaviruses, production of the 8.4 kDa TF protein may have two consequences. First, because E1 is a structural protein, -1 PRF may play a role in virion assembly by controlling E1 expression levels, and thus altered E1 production could negatively interfere with virion assembly^{147,191}. The data presented in Figure 3 partially supports this model, as ablation of the -1 PRF signal resulted in decreased viral production/release *in vitro*. This small decrease in virion production may provide just enough of a difference to enable the host to mount an effective immune response, as evidenced by the longer viral residence times of TrD_{PRFm} in the spleen (Figure 25D). Alternatively, TF itself may have a biological role separate from viral particle assembly. This is supported by the observation that -1 PRF inhibition attenuated VEEV pathogenicity and altered viral spread in mice. Consistent with the flavivirus NS1' protein literature, the observation that TrD_{PRFm} virus promoted decreased viral titers in the brains of infected mice suggest that the VEEV TF may be important for passage through the blood-brain barrier, and/or for neuroinvasiveness^{147,176,188}. It is important to note that these two options are not mutually exclusive, in that decreased TF expression and the accompanying increase in E1 levels may influence viral assembly/release, while at the same

time impact an as yet unidentified role of TF. There have been no studies to date examining the interactions between TF and other viral or cellular proteins. Such analyses would lend great insight into the role of TF and will be the focus of future studies. It is also worth noting that attenuation of TC83 is due a substitution at position 120 of the E2 glycoprotein¹⁰⁵; thus it is possible that the differences observed with the PRF mutant in TC83 as compared to TrD may be due to disruption of an interaction between E2 and TF.

The development of VEEV as a biological weapon in the U.S. and former USSR, and a documented history of over 150 cases of serious laboratory infections by VEEV⁷¹ led to it being included as a select agent by the government of the United States of America. As noted above, the FDA has not approved any vaccines or therapeutics for the equine encephalitis viruses. The attenuated vaccine strain, TC83, was generated in the early 1960's by serially passing VEEV 83 times in guinea pig heart cells¹⁹². TC83 poses a high risk for reversion due to the fact that it harbors only two attenuating mutations¹⁰⁵, and can also be transmitted by mosquito vectors¹⁹³. Because of these risks, coupled to its demonstrated ability to cause mild- to severe flu-like symptoms in approximately 25% of volunteers, and promoted seroconversion in only 80% of volunteers¹⁹⁴, TC83 has only limited utility for use in humans and its use is limited to laboratory personnel and military at risk of contracting the virus¹⁹⁵. More recent live attenuated vaccine candidates are based on the VEEV TrD infectious clone used in the current study. These include insertion of specific point mutations or a mutation in the PE2 cleavage-signal combined with a mutation that rescues E1

gene function. The resulting V3526 strain is safe and immunogenic in non-human primates and mice, and has a lower risk for mosquito transmission (reviewed in ¹⁹⁵). The finding in this study that attenuation of -1 PRF strongly attenuated VEEV neuropathogenicity of the virus represents a promising new avenue of inquiry toward the development of safe and effective live attenuated vaccines directed against VEEV and perhaps other -1 PRF utilizing members of the *Togavirus* and *Flavivirus* families. In addition, novel -1 PRF stimulating mRNA elements identified here may also serve as targets for small molecule therapeutics directed against these viruses.

Acknowledgements

The content of this article does not necessarily reflect the position or the policy of the federal government, and no official endorsement should be inferred. The authors would like to thank Dr. Pamela Glass, United States Army Medical Research Institute of Infectious Disease, for providing the TrD molecular clone and Dr. Ilya Frolov, University of Alabama, for providing the TC83 molecular clone. The authors also thank Dr. Diane Griffin, Johns Hopkins Bloomberg School of Public Health, for providing the AP-7 rat neuronal cells. This work was funded through a Defense Threat Reduction Agency grant, HDTRA1-13-1-0005. JAK was supported by NIH Institutional Training Grant 2T32AI051967-06A1.

Chapter 3: Functional and structural characterization of the Chikungunya virus translational recoding signals

Joseph A. Kendra^{‡1}, Vivek M. Advani^{‡§.2}, Bin Chen¹, Joseph W. Briggs^{‡.3}, Jinyi Zhu[§], Hannah J. Bress[§], Sushrut M. Pathy[§] and Jonathan D. Dinman^{‡*}

From the [‡]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park MD, 20742, [§]First-Year Innovation & Research Experience Program, University of Maryland, College Park MD, 20742

Running title: *Chikungunya virus recoding signals*

¹Both authors contributed equally to this work

This manuscript has not been published at the time of this dissertation. It was submitted to the Journal of Biological Chemistry as of August 28th, 2018.

Abstract

Climate change and human globalization have facilitated the rapid spread of mosquito-borne diseases to naïve populations. One such emerging virus of public health concern is Chikungunya virus (CHIKV), a member of the *Togaviridae* family, genus alphavirus. CHIKV pathogenesis is predominately characterized by acute febrile symptoms and severe arthralgia, which can persist in the host long after viral clearance. CHIKV has also been implicated in cases of acute encephalomyelitis and vertical transmission has been reported. Currently, no FDA approved treatments exist for this virus. Here, we report the molecular and structural characterization of two CHIKV translational recoding signals: a termination codon readthrough (TCR) element located between the non-structural protein 3 and 4 genes, and a programmed -1 ribosomal frameshift (-1 PRF) signal located toward the 3' end of the CHIKV 6K gene. Efficient TCR and -1 PRF were validated and genetically characterized using dual-luciferase and immunoblot assays in HEK293T and U-87 MG mammalian cell lines. Analyses of RNA chemical modification data (SHAPE) revealed that CHIKV -1 PRF is stimulated by a tightly structured, triple stem hairpin element, consistent with previously investigated alphaviruses, and that the TCR signal is composed of a single large multi-bulged hairpin element. These findings illuminate the roles of RNA structure in translational recoding and provide critical information toward creating of live-attenuated vaccine design.

Introduction

The spread of mosquito-borne viruses has been accelerated by climate change and advances in globalization²⁷. One such example is Chikungunya virus (CHIKV), a member of the *Togaviridae* family, genus *alphavirus*. First identified during an outbreak in Tanzania in 1952⁶³, CHIKV was implicated in large-scale outbreaks in Africa and Asia in 2004¹⁹⁶. Its geographic spread has since encompassed Europe, Australia, the Pacific Islands and the Americas¹⁹⁷. “Old-World” alphaviruses including CHIKV, Sindbis virus (SINV), Semliki Forest Virus (SFV) and O’Nyong’Nyong virus (ONNV) are endemic to Africa and central Asia, and infections are characterized by fever, rash, and arthritic disease. The pathogenesis of CHIKV is predominately characterized by an incubation period of 3-7 days, followed by acute, febrile illness and severe arthralgia⁵⁸. Most patients recover within two weeks, but complications can result in debilitating sequelae persisting for years after viral clearance. Recent reports have implicated CHIKV in cases of acute encephalomyelitis^{198,199} similar to that of “New World” Equine Encephalitis alphaviruses, and the capability for vertical transmission during pregnancy or birth^{200,201} has been demonstrated. Currently, no FDA approved treatments for CHIKV have been approved for civilian use.

Many RNA viruses employ translational recoding mechanisms to expand the coding capacity of limited genome space and to optimize the stoichiometric expression of critical proteins¹¹⁵. Two recoding signals have been documented in alphaviruses: those promoting termination codon readthrough (TCR) and programmed -1 ribosomal frameshifting (-1 PRF). The CHIKV TCR is located at

an opal (UGA) termination codon in the genomic RNA beginning at nucleotide position 5656, which marks the boundary between the nsP3 and nsP4 genes¹³⁸. A closely associated downstream stimulatory element is thought to prevent efficient association of the eRF1/eRF3 complex with the ribosomes stalled at this codon, increasing the likelihood of a TCR event¹³⁹. The specific stop codon identity for this signal is critical for optimal alphavirus functionality, as substitutions to amber or ochre stop codons or an arginine readthrough have been associated with lowered transmission in mosquito vectors and significantly attenuated pathogenesis^{140,202,141}. The -1 PRF signal is located in the subgenomic RNA that encodes a polyprotein that is subsequently processed into structural proteins. -1 PRF occurs at a conserved U UUU UUA slippery site sequence near the 3' end of the 6K gene, resulting in production of a *trans*-frame (TF) product¹⁴³. TF then undergoes unique post-translational modifications respective to 6K and is integrated into the envelopes of mature virions¹⁴⁴. While the biological function of TF is poorly understood¹⁴⁸, it has been hypothesized to function as an ion channel, similar to its 6K counterpart¹⁴². A study with SINV mutants that disrupted production of TF resulted in reduced virus production in mammalian and insect cell lines that was independent of genome replication, particle infectivity or envelope protein transport to the cellular membrane. These findings suggest potential roles for TF in virus particle assembly and budding. Notably, SINV mutants with disrupted TF production strongly attenuated the virus in mouse models¹⁴⁷. Similarly, ablation of -1 PRF in VEEV, a new world alphavirus, had minimal effects on viral replication while abrogating

neurovirulence in mouse models¹⁵⁰. These findings suggest that ablation of alphavirus recoding elements could serve as the foundation for live-attenuated vaccines for this virus family.

This study describes the genetic and structural characterization of the CHIKV TCR and -1 PRF signals. We demonstrate that the bioinformatically predicted CHIKV recoding signals promote efficient levels of translational recoding in several mammalian cell lines. The observed recoding efficiencies were genetically verified through mutations used to ablate the respective TCR and -1 PRF mechanisms. Analyses of chemical modification assays identified the presence of well-organized hairpin structures downstream of the respective recoding sequences, the structures of which were tolerant of geographically isolated polymorphisms. Key mutations were elucidated that significantly ablated recoding functionality were identified for both signals, laying a foundation for follow-up attenuated live virus experiments in animal models.

Materials and Methods

Computational prediction of CHIKV recoding signals

CHIKV sequences were imported from NCBI (<http://www.ncbi.nlm.nih.gov>) into the Dinman lab frameshifting database (<http://prfdb.umd.edu>)¹⁵⁶. The conserved opal stop codon at the 3' end of the NSP3 gene, and the UUUUUUA slippery site in the 6K gene, were used to identify the sequences harboring the CHIKV TCR and -1 PRF stimulatory elements respectively. The RNA folding algorithm NUPACK¹⁵⁷ was employed to model potential downstream RNA

secondary structures of these recoding signals, serving as guides for subsequent cloning into reporter vectors.

Dual reporter plasmid construction

Translation recoding rates were monitored using pJD2257 based dual luciferase vectors, a modification of pSGDluc²⁰³ in which a duplicate *Bam* HI and *Sal* I restriction site was eliminated, leaving unique sites located between the *Renilla* and firefly luciferase coding sequences. DNA sequences harboring the CHIKV recoding signals containing 5' and 3' overlapping sequences with pJD2257 were designed *in silico* and synthesized by Genewiz (Supplementary Table 1) (Gaithersburg MD). These were used to construct reporter plasmids by the DNA fragment assembly method²⁰⁴ using the In-Fusion® HD cloning plus kit from Clontech laboratories (Mountain View CA, catalog # 638911). These inserts we also cloned into control variants of pJD2257 that harbor in frame UAA termination codons either 5' or 3' of the multiple cloning site (pJD2267 and pJD2269, respectively). Plasmids and sequences are described in Supplementary Table 2, and are available upon request.

Cell Culture

Cell lines for HEK293T (catalog # CRL-3216) and U-87 MG (catalog # HTB-14) were purchased through ATCC (Manassas, VA). HEK293T cells were cultured in Dulbecco's modified Eagle medium (DMEM) (Lonza) and U-87 MG cells were cultured in Eagle's minimum essential medium (EMEM) (Lonza) supplemented

with 1% L-glutamine, 10% FBS and 1X penicillin-streptomycin to obtain complete growth media. The cells were grown at 37°C in 5% CO₂.

Plasmid Transfections

HEK293T, or U-87 MG cells were seeded at 0.6×10^5 cells in 0.5 ml per well for 24-well plates for dual luciferase transfections in appropriate complete growth media (DMEM or EMEM). Following a 24-hour incubation period, control and experimental reporter plasmids were then transfected into cells using a Lipofectamine 3000 transfection reagent from ThermoFisher (catalog# L3000015).

Assays of translational recoding

Translational recoding efficiencies measured using pJD2257 based reporter plasmids were assayed as previously described¹⁵⁰ using a dual-luciferase reporter assay system kit from Promega (catalog# E1980). At 24-hours post-transfection, cell culture media was aspirated and the cells were rinsed twice with 1X phosphate-buffered saline (PBS) before disruption with 1X Passive Lysis Buffer (PLB). Cell lysates were assayed in triplicate in a 96-well plate. Firefly and *Renilla* luciferase activities were quantified using a GloMax®-Multi Detection system from Promega (catalog# E7041).

Immunoblot Analyses

HEK293T cells were seeded at 1×10^6 cells/well in 6-well plates followed by transfections with dual luciferase plasmid vectors using the Lipofectamine 3000

transfection reagent from ThermoFisher (catalog # L3000015). Cells were harvested 48 hours post transfection and lysates were prepared using NP-40 lysis buffer (from Boston BioProducts catalog# BP-119) including protease and phosphatase inhibitors (from ThermoFisher catalog# 78440). Protein concentrations were determined and 15-20 µg of each sample was separated by SDS-PAGE (4-15% gels from Bio-Rad catalog# 4568084) and then transferred to 0.45µm nitrocellulose membranes from Fisher (catalog# LC2001). Post-transfer membranes were blocked with 5% non-fat skim milk in 1× PBST (1× PBS, 0.5% Tween20) for 2 hours at room temperature and then incubated overnight at 4°C with primary firefly luciferase (polyclonal anti-goat from Promega catalog# G7451 and Lot# 0000255518) and *Renilla* luciferase (monoclonal anti-rabbit from BosterBio catalog# MT0022 and Lot# 001812M2250) antibodies at 1:1000 dilution in 1× PBST containing 5% BSA. Blots were washed with 1× PBST and incubated for 3 hours at room temperature with goat anti-goat IgG HRP conjugate secondary antibody (from Bethyl laboratories catalog# NC9452917 and Lot# A50-201P) and anti-rabbit IgG HRP conjugate secondary antibody (from Cell Signalling catalog# 7074S and Lot# 26) in 5% non-fat skim milk 1× PBST. Immunoreactive bands were detected using LumiGLO reagent from Cell Signaling Technologies (catalog# 7003S) and visualized in Fuji, LAS-3000 imager.

RNA structure analyses

The CHIKV -1 PRF and TCR signals were structurally assayed by Selective 2'-Hydroxyl Acylation analyzed by Primer Extension (SHAPE) as previously

described with the following modifications^{205,206}. DNA templates for *in vitro* transcription reactions were generated by PCR amplification using DreamTaq DNA polymerase from Thermo Fisher (catalog # FERK1071PM). Complementary primers for the *Renilla* and *Firefly* regions were used to amplify the inserted CHIKV sequence and attach a T7 promoter sequence to the 5' end of the amplicons. *in vitro* transcription was carried out with a T7 MEGAscript kit from Life Technologies (catalog#AM1334). Transcribed RNA was purified with MEGAclear cleanup kit (Life Technologies, catalog# 1908) and the quality of the RNA samples was assessed through urea-PAGE. Modification of flexible bases with *N*-methylisatoic anhydride (NMIA) was carried out as previously described¹⁵⁰. The oligonucleotide 5'-AGGATAGAATGGCGCCGGGCC-3' was 5' labeled using g[³²P] ATP from Perkin Elmer (catalog # BLU502Z250UC) and polynucleotide kinase. The primer was annealed to modified RNA and subsequent reverse transcription (using Superscript III RT, ThermoScientific catalog# 18080044) were carried out as previously reported^{205,206}. Radioactivity of cDNA samples was standardized with a liquid scintillator prior to electrophoresis through 8% urea-PAGE (SequaGel UreaGel system, National Diagnostics catalog# EC-833). cDNA products were visualized on a Fujifilm phosphorimager. Visual clarity of gel images was adjusted with Adobe Photoshop Lightroom 5. All primers were purchased through IDT.

Three-dimensional structural modeling of CHIKV -1 PRF and TCR stimulatory elements

All-atom models were generated using the MC-Fold and MC-Sym pipeline programs¹⁶⁵. RNA sequences for the CHIKV -1 PRF and TCR signals were imported into the MC-Fold program to generate a series of secondary structures as previously described¹⁵⁰. The highest-scoring model was selected to represent the predicted 3-dimensional structure of the viral PRF signals.

Phylogenetic analyses

Accession numbers of CHIKV sequences isolated from patients of either African-Asian origin (n=14) or Caribbean origin (n=5) were considered for phylogenetic analysis. Translational recoding sequences (PRF and TCR) from these strains were aligned using Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>)²⁰⁷.

Statistical analyses

Translational recoding assays repeated in triplicate (experimental replicates) were independently repeated a minimum of three times (biological replicates) and mean technical replicate values for each independent biological replicate are shown on graphs along with standard deviation of biological replicates. Statistical analyses were conducted using a Student's t test or one-way ANOVA as appropriate using Prism, version 6, software (GraphPad).

Results

The CHIKV translational recoding sequences are highly conserved

Alphaviruses have single-stranded, positive sense RNA genomes that contain two large open reading frames (ORF) (Figure 26A)⁵⁷. The first ORF encodes a polyprotein that is proteolytically cleaved into the non-structural proteins nsP1, nsP2, nsP3 and nsP4 (nsP4 encodes the viral RNA-dependent RNA polymerase or replicase). The second ORF, accessible to the translational apparatus as a sub-genomic transcript, encodes a polyprotein which is subsequently processed into the C, E2, E3, 6K and E1 structural proteins. Programmed translational recoding signals have been identified in both ORFs. In the nonstructural ORF, the nsP3 gene ends with a UGA opal termination: thus expression of the nsP4 replicase requires a programmed termination codon readthrough (TCR) mechanism¹³⁸. In the subgenomic RNA, a programmed -1 frameshift signal (-1 PRF) located in the 3' region of the 6K viroporin gene enables synthesis of a C-terminally extended trans-frame variant dubbed TF that has been identified in the envelope of mature alphavirus particles¹⁴³. Ribosomes that have been shifted to the -1 frame to produce TF cannot translate the E1 protein.

Human globalization and climate change have facilitated the rapid global spread of CHIKV, resulting in the generation of multiple viral strains and lineages with differing degrees of pathogenicity^{208,26,25}. A prior analysis revealed considerable amounts of heterogeneity among these strains throughout their genomes²⁰⁹. Consensus sequences of patient-derived CHIKV isolates from the

Africa/Asia and Caribbean geographic regions (abbreviated as Af/As and Carib) were aligned to assess the conservation of the TCR and -1 PRF recoding regions (Figure 26B). The full collection of CHIKV isolate sequences used in this study can be found in Figure 35. The opal termination codon is universally conserved in both consensus sequences, but five polymorphisms were identified in the downstream sequence predicted to harbor the accompanying TCR stimulatory element. Similarly, full conservation of the alphaviral U UUU UUA slippery site was observed in both CHIKV -1 PRF consensus sequences, but the sequence identity of the downstream stimulatory structure diverged by only a single nucleotide. Following this analysis, it was determined that both geographic consensus sequences for the CHIKV TCR signal would be considered for subsequent experiments, while only the Af/As sequence of the CHIKV -1 PRF signal was selected as a representative.

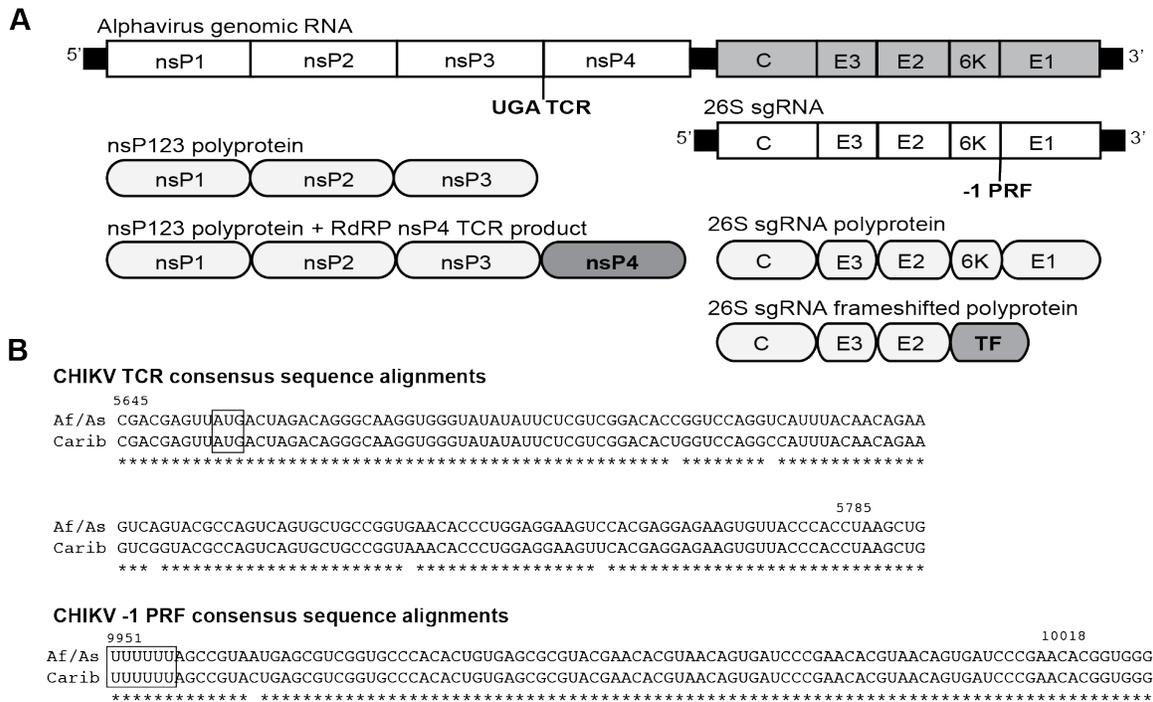


Figure 26. Identification and conservation analysis of CHIKV derived programmed translational recoding signals. (A) General schematic of the alphavirus genome and subgenomic transcript. The presence of the TCR signal 3' of nsP3 and the -1 PRF signal in the 3' region of 6K are identified. The oval chains below denote the two polyproteins that can be synthesized from both the nonstructural and structural ORFs, and the proteins that are encoded by TCR (nsP4) or -1 PRF (TF) are shaded. (B) Consensus sequence alignments of patient-derived CHIKV programmed translational recoding signals from the Af/As and Carib geographic regions. Numbers indicate the starting and ending nucleotide positions of each region under investigation. Asterisks (*) are used to denote identical nucleotides among consensus sequences.

CHIKV-derived sequences promote efficient recoding in mammalian cell lines

The Af/As and Carib variants of the CHIKV TCR signal and the singular -1 PRF signal were cloned into dual luciferase reporter vectors. The firefly luciferase ORF flanking the 3' end of the inserted sequences was adjusted such that translation could only be achieved consequent to a termination codon readthrough event or a ribosomal shift into the -1 frame, respectively¹⁷². Dual

luciferase reporters harboring the Murine Leukemia Virus (MuLV) TCR and the VEEV -1 PRF signals were employed as positive controls and a reporter with a UAA termination codon inserted 5' of the firefly luciferase ORF (5' UAA) constituted the negative control. All three CHIKV-derived sequences promoted efficient levels of translational recoding in HEK293T and U-87 MG human astrocyte-derived cell lines (Figure 27A). Specifically, average Af/As and Carib TCR efficiencies were in the range of 7.0% and 7.1% in HEK293T cells and 13.4% and 13.7% in U-87 MG cells, respectively. These values were very similar to those measured from the MuLV TCR control reporter. Notably, no significant differences in TCR efficiency were recorded between the Af/As and Carib geographic variants of CHIKV. Average CHIKV -1 PRF efficiency was lower, approximately 1.8% in HEK293T cells and 4.5% in U-87 MG cells, as compared to the VEEV -1 PRF reporter (average 2.3% in U-87MG, and 4.6% in HEK293T cells). However, all of these values were statistically significantly greater than the 5' UAA control (<1% in both cell lines). The possibility that the observed recoding efficiencies were the products of a cryptic promoter, splice site or IRES was tested with two modified vector backbones that placed an in-frame termination codon either directly 5' of the inserted test sequence or in the outgoing reading frame 3' of the insert prior to the firefly sequence. Dual luciferase assays revealed that these control vectors significantly abrogated CHIKV Af/As and Carib-mediated TCR and -1 PRF efficiencies in both cell lines, supporting the hypothesis that these are indeed legitimate viral recoding signals (Figure 27B). As an orthogonal approach to characterizing efficient recoding, the firefly and

Renilla luciferase products of transfected HEK293T cells were also visualized by immunoblot; these data effectively recapitulated the enzymatic reporter data (Figure 27C).

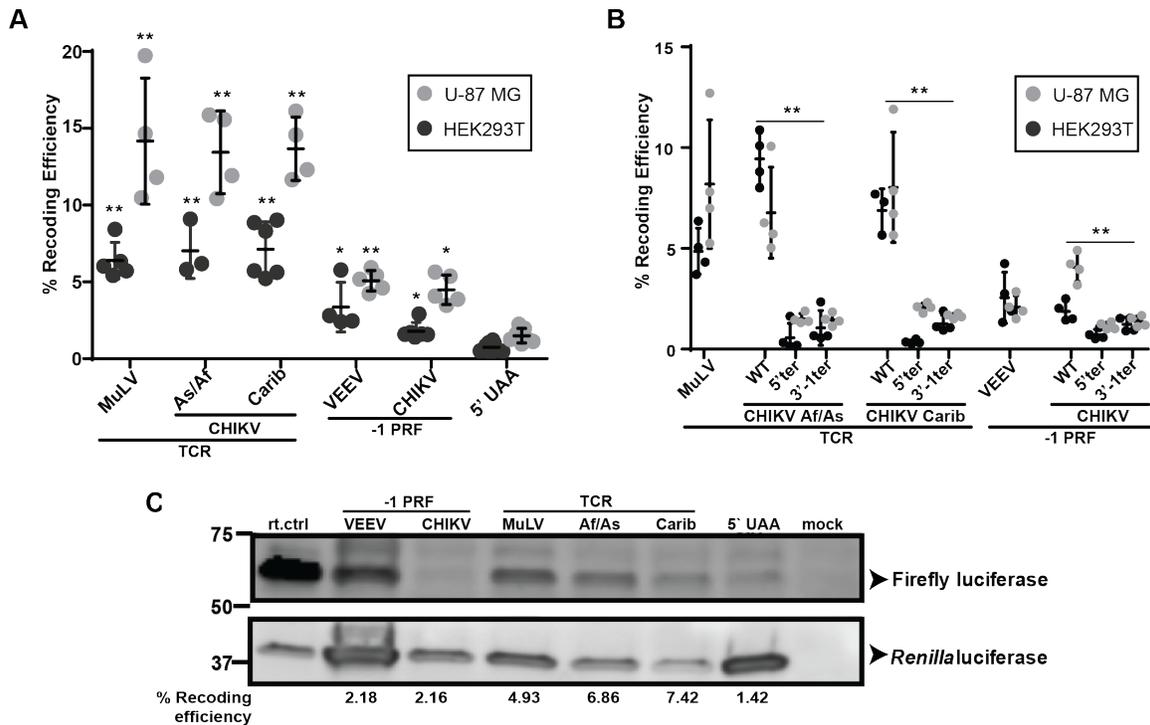


Figure 27. Characterization and validation of efficient CHIKV translational recoding in mammalian cell lines. (A) The predicted TCR and -1 PRF signals identified in Figure 26 were cloned into pJD2257 dual luciferase reporter vectors and their ability to promote efficient recoding was monitored in HEK293T and U-87 MG cells: human kidney and astrocyte cell lines, respectively. MuLV and VEEV were employed as positive controls for efficient TCR and -1 PRF, as well as a negative control consisting of a UAA termination codon in the 0 frame. (B) CHIKV recoding signals were cloned into dual reporter control vectors containing a premature termination codon either 5' of the insert sequence in the 0 frame (5'ter) or 3' of the insert in the -1 frame (3'ter). (C) Representative immunoblot of protein lysates generated from HEK293T cells transfected with the indicated reporter plasmids. The slower migrating band in the firefly luciferase probed panel corresponds to incomplete cleavage of the Stop/Go inteins that were inserted immediately 3' of the recoding signals (n=2). Dual luciferase data is presented as means with standard errors, where each point denotes a biological replicate assayed as technical triplicates. Asterisks denote statistical significance, where * $p < 0.05$; ** $p < 0.01$.

Targeted mutations alter CHIKV recoding efficiencies

Prior studies revealed that efficient -1 PRF in encephalitic alphaviruses can be abrogated by strategically targeted slippery sequence point mutations, leading to attenuated pathogenesis^{147,150}. Site-directed mutagenesis was employed to test the effects of targeted mutations of CHIKV translational recoding efficiencies in HEK293T cells. A significant decrease in efficient TCR in mammalian cell lines was observed when the termination codon was substituted with more frequently employed UAA and UAG stop codons, reducing Af/As TCR from ~10% to 3.7% and 2.7%, respectively, and 0.9% and 2.7% for Carib (Figure 28A). Similarly, silent coding mutations to the slippery site of the -1 PRF signal intended to impede 5' slippage of tRNA non-wobble bases resulted in significantly decreased frameshifting efficiency (<1.0%) (Figure 28A). Reports in the literature have also indicated that substitution of an arginine codon for the UGA codon of the CHIKIV TCR signals promoted reduced viral pathogenesis¹⁴⁰. When the TCR opal termination codon was mutated to an AGA arginine, 83.6% and 77.0% readthrough efficiencies were observed for the Af/As and Carib consensus sequences (Figure 28B). These results were replicated via Western blot as an orthogonal means of verification (Figure 28C, 28D).

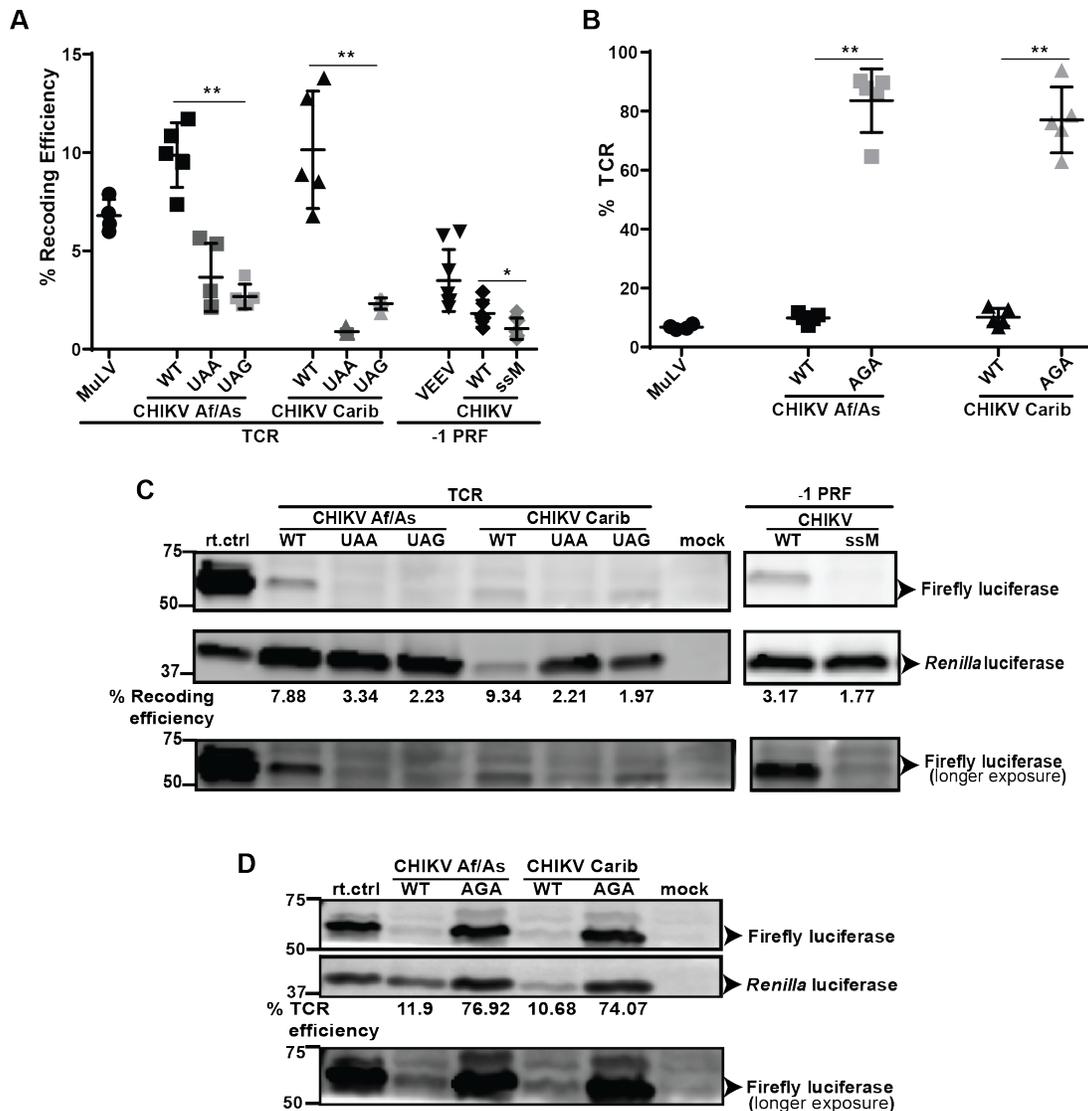


Figure 28. Mutagenesis-mediated destabilization of efficient CHIKV recoding. (A) Site directed mutagenesis was employed to replace the UGA termination codon (WT) in the CHIKV TCR signal with either a UAA or UAG stop. A slippery site mutant of the CHIKV -1 PRF signal (ssM) that replaced the U UUU UUA sequence with G UUC UUG was also created. The recoding efficiencies promoted by recoding mutants relative to their wild-type counterparts were assayed in HEK293T and U-87 MG cell lines. (B) Site directed mutagenesis was used to substitute the UGA stop codon in both CHIKV TCR sequences with AGA. The readthrough efficiency of this mutant was tested alongside the wild-type TCR sequences in HEK293T cells. (C, D) Western blot verification of the results described in A and B (n=2). Sequences from MuLV and VEEV were employed as positive TCR and -1 PRF controls where appropriate. Dual luciferase assay data is presented as means with standard errors, where each point represents a biological replicate assayed as technical triplicates. Asterisks denote statistical significance, where * $p < 0.05$; ** $p < 0.01$.

*Chemical modification analyses identify complex stem-loop elements
immediately 3' of CHIKV recoding sites*

Translational recoding signals typically require the presence of highly structured RNA stimulatory elements immediately 3' of the recoding site. These *cis*-acting elements are hypothesized to function as kinetic traps for ribosomes, facilitating the conditions favorable for TCR or -1 PRF¹¹⁵. A diverse array of 3' stimulatory structures have been reported, ranging from simple stem-loops to RNA pseudoknots¹⁴⁹. Each unique recoding signal secondary structure plays an integral role in determining the specific efficiency of translational recoding^{130,129}. Selective 2' hydroxyl acylation and primer extension (SHAPE)¹⁶⁴ was employed to resolve the *cis*-acting stimulatory elements of the CHIKV TCR and -1 PRF signals (Figure 29A-C). A large, terminally forked stem-loop structure was identified 3' of the opal termination codon for both the Af/As and Carib variants of the CHIKV TCR signal (Figure 29D, 29E). Notably, the higher order structure of this stem-loop was permissive of all five of the polymorphisms between the two geographic variants. A smaller bulged stem-loop structure was also identified as the *cis*-acting element located downstream of the CHIKV -1 PRF slippery sequence (Figure 29F, 29G). The predicted organization and folding of this element is very similar to the minimum-required structures required for efficient -1 PRF in the Equine Encephalitic virus family of alphaviruses¹⁵⁰.

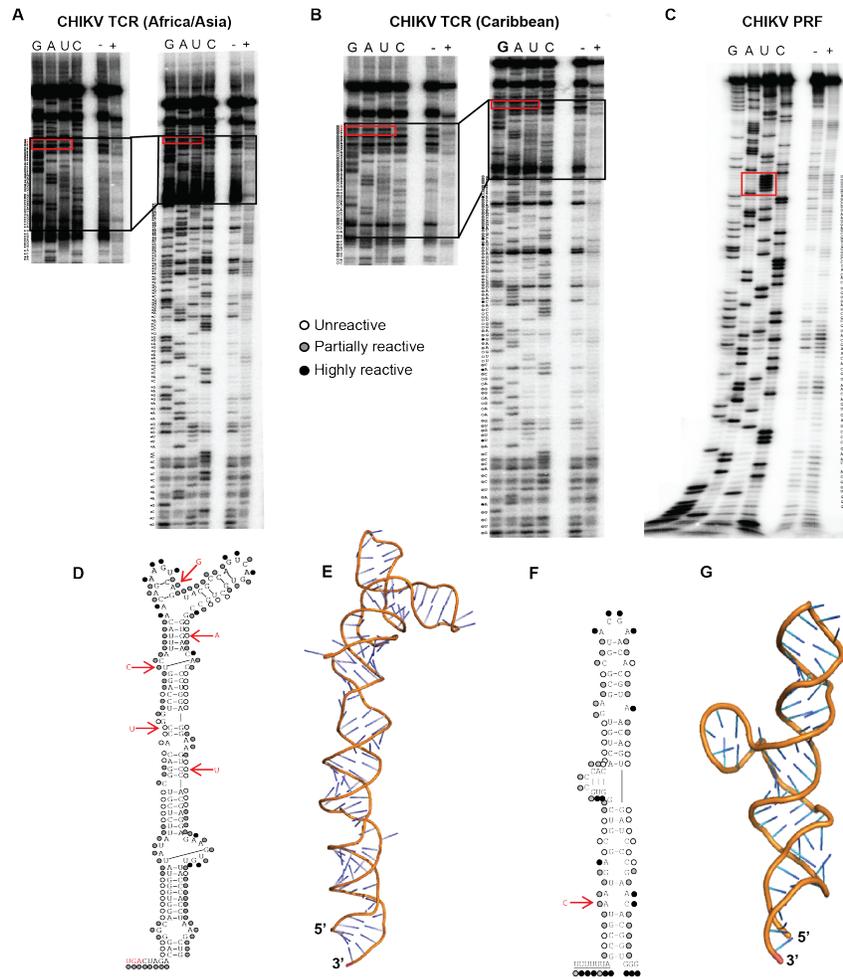


Figure 29. Structural analyses of the CHIKV recoding signals. (A-C) Stimulatory elements for Af/As and Carib CHIKV TCR and -1 PRF signals resolved through selective 2' hydroxyl acylation and primer extension (SHAPE). RNA templates containing the translational recoding sequences were transcribed from corresponding dual-luciferase reporter vectors and treated with N-methylisotoic anhydride (NMIA). Untreated RNAs were used as negative controls. $\gamma^{32}\text{P}$ radiolabeled cDNA products were separated through 8% urea-PAGE and visualized via a Fujifilm phosphorimager. Autoradiograms are annotated to indicate the respective sequencing lanes (G, A, U, C), an untreated control lane (-) and the NMIA labeled experimental lane (+). Circles denote the relative reactivity of bases, where white is the most unreactive and black is the most reactive. For added visual clarity of the CHIKV TCR gels, a longer run of the samples has been provided to further separate the 5' sequence information. (D, F) Structures of CHIKV recoding signals derived from the above SHAPE data. Circles correspond to the previously described nucleotide reactivity scale. Polymorphisms between the Af/As and Carib consensus sequences are indicated in red. (E, G) 3D models of the CHIKV TCR and -1 PRF signals were generated by Molecular Dynamics simulations and visualized in Pymol.

Discussion

Although the presence of alphavirus TCR and -1 PRF recoding signals were first identified in 1983 and 2008 respectively^{138,143}, the structural and functional characterization of these *cis*-acting control elements has only been minimally examined. The current study presents the first structural characterization of a TCR signal through direct experimentation, and reveals that the CHIKV -1 PRF signal is very similar to those characterized in the new world encephalitic alphaviruses¹⁵⁰. Specifically, the predicted CHIKV TCR signal was also shown to promote recoding in mammalian cells at a similar efficiency to that Sindbis virus, another “Old World” alphavirus¹⁴⁷. It was experimentally determined that the highly efficient TCR is dependent on the presence of the UGA opal stop codon, as amber UAA and ochre UAG stop codon substitutions resulted in diminished TCR efficiencies. A large stem-loop structure was characterized immediately 3' of the opal termination codon, likely influencing TCR activity as well. CHIKV -1 PRF was ablated by mutations predicted to disrupt canonical recoding mechanisms, and chemical probing assays identified a downstream branched stem-loop element similar to what was observed in VEEV.

A close examination of the recoding data suggest a few notable items. With regard to the TCR reporters, the apparent “readthrough” efficiency of the Arg containing mutant was ~90%, i.e. not 100% as would be expected. This reporter retains the strong TCR stimulating stem-loop element, while the control readthrough reporter does not. Given that strong RNA secondary structures can

induce ribosome pausing and dropoff^{210,211,212,213,214}, we speculate that this structural element may cause a fraction (~10%) of elongating ribosomes to drop off of the reporter mRNA, thus accounting for the observed TCR values of less than unity. Experiments to measure the fraction of ribosomes paused at this element, their pause times and dropoff rates are planned for the future. A second set of insights stems from the observation that the CHIKV -1 PRF signal promotes relatively low rates of recoding (most viral frameshifting is in the range of $\geq 5\%$, see ¹⁷⁹). Examination of sequences proximal to the -1 PRF signal did not reveal the potential to form either a larger stem-loop or a more complex structure, e.g. an RNA pseudoknot. However, it is possible that this element may be involved in long range interactions which, by rendering it more difficult to unwind, would increase ribosome pause rates and thus -1 PRF efficiency²¹⁵. For example, a -1 PRF stimulating long-distance RNA-RNA interaction was first demonstrated in Barley yellow dwarf virus (BYDV), where base-pairing between an internal bulge in the -1 PRF promoting stem-loop and a sequence nearly 4kb downstream is required to promote efficient -1 PRF^{137,186}. Similar long distance interactions involving bulges located inside of -1 PRF stimulating stem-loops have been observed in Red clover necrotic mosaic virus (RCNMV)²¹⁶, in the Pea enation mosaic virus (PEMV)²¹⁷, and in the TCR promoting elements of tobacco necrosis virus-D²¹⁸ and Turnip crinkle virus (TCV)¹³⁵. The bulged stem-loop in the CHIKV -1 PRF signal in the current study, and prior demonstration of similar -1 PRF stimulating RNA structural elements in the new-world alphaviruses¹⁵⁰ are highly reminiscent of the plant virus recoding signals, suggesting that the

alphaviral recoding elements may similarly involve long distance RNA-RNA interactions. More broadly, it is becoming clear that dynamic long range RNA-RNA interactions are critical for programming the fundamental molecular processes of (+) ssRNA viruses, including the switch from translation to transcription, cap-independent translation, genome circularization, replicase complex assembly, subgenomic mRNA synthesis, and repriming during discontinuous template synthesis^{219,220}. Many of these are characterized by base pairing interactions between the loops at the ends of stem-loop structures. In this respect, the forked distal tip of the CHIKV TCR signal may present a novel such interacting site. Thus it is not unlikely that interactions between the structured recoding elements identified in the current study and distal sequence elements are involved in the dynamic programming of the CHIKV life cycle. A deeper understanding of such interactions will require studies examining RNA structural dynamics of the gRNA, sgRNA, and (-) strand replicative intermediates.

PRF was first discovered in retroviruses, where ribosome slippage directs the synthesis of Gag-pol polyproteins¹⁸³. Subsequent studies using totiviruses¹²⁴, and later in retroviruses¹⁸⁴ demonstrated that changes in -1 PRF efficiency affect virus production. From these findings, a bioeconomics model emerged in which -1 PRF rates are optimized to maximize virus particle assembly by ensuring the synthesis of the correct stoichiometric ratios of the structural Gag-derived proteins to the enzymes encoded by the Pol open reading frame¹⁷⁵. These findings engendered interest in targeting -1 PRF for antiviral therapeutics¹⁸⁵. In parallel, early studies examined -1 PRF signals in virus families where they do

not occur between open reading frames encoding structural and enzymatic open reading frames (ORFs), e.g. in coronaviruses and luteoviruses. Altering -1 PRF efficiency in severe acute respiratory syndrome coronavirus (SARS-CoV) severely impacted its infectivity in tissue culture, reinforcing the idea of -1 PRF as an antiviral therapeutic target¹²⁸. Lower rates of -1 PRF correlate with decreased pathogenicity in West Nile Virus¹⁵⁵, and production of the NS1' frameshift product is critical for neuroinvasiveness in West Nile and Japanese encephalitis viruses¹⁷⁷. -1 PRF also presents a target for *Alphaviruses*: in mice ablation of -1 PRF attenuated the symptoms of Sindbis¹⁴⁷ and Venezuelan Equine Encephalitis Virus (VEEV)¹⁵⁰ infections. Although less well studied than -1 PRF, five published studies report that alteration of TCR also has negative impacts on virus replication^{141,221,222,223,224}. While^{221,222,223} examined the impact of TCR ablation in a retrovirus (murine leukemia virus), pertinent to this study TCR ablation was shown to reduce the pathogenicity of two *Alphaviruses*: O'Nyong'Nyong and Sindbis viruses^{141,224}. Thus, both -1 PRF and TCR represent important, yet underexploited targets for antiviral intervention. For example, ablation of -1 PRF and/or TCR signals may be incorporated into the design of attenuate live virus vaccines. In parallel, elucidation of the atomic resolution structures of these elements, and of their structural dynamics, may be useful in the design of therapeutic small molecule inhibitors.

Acknowledgements

This work was partially funded through grants from the Defense Threat Reduction Agency grant (HDTRA1-13-1-0005) and the National Institutes of

Health/National Institutes of General Medical Sciences (R01 GM117177). J.A.K. was supported by NIH institutional training grant 2T32AI051967-06A1. J.W.B. was supported by a grant from the National Institutes of Health/National Heart Lung Blood Institute (R01HL119439). V.M.A was supported by First-year Innovation and Research Experience (FIRE) program at University of Maryland, College Park. Part of this work was made possible by an undergraduate fellowship to J.Z. from the Howard Hughes Medical Institute/University of Maryland College Park.

Chapter 4: Conclusions and future directions

Programmed translational recoding is a critical, multifaceted feature of many RNA viruses. Originally identified as a means of expanding the genetic economy of limited genome size, further investigation has contextualized these signals as a regulatory element for critical genes^{108,128}. These signals are therefore under heavy selective pressure to maintain their recoding efficiency at specific stoichiometric ratios¹²⁴. Due to their necessary inflexibility, it stands to reason that these signals constitute a crucial genetic vulnerability in viruses that can be exploited for therapeutic development.

The research into VEEV frameshifting in Chapter 2 elucidated the genetic and structural characteristics of encephalitic alphavirus -1 PRF signals. Key silent coding mutations to the VEEV slippery site resulted in attenuated pathogenesis in BALB/c mice infected through the aerosol route. In a follow-up study, it was also determined that VEEV_{PRFm} has immunogenic capabilities. Following subcutaneous infection, significant levels of neutralizing antibodies were detected in mice 28 days post infection (Figure 30A, 30B). Critically, all of these mice survived successive challenge with a lethal dose of VEEV_{TRD} (Figure 30C). In conjunction with the results from Chapter 2, these experiments demonstrated that mutations that ablate efficient -1 PRF constituted a promising foundation for the development of live-attenuated vaccines against VEEV and potentially other alphaviruses.

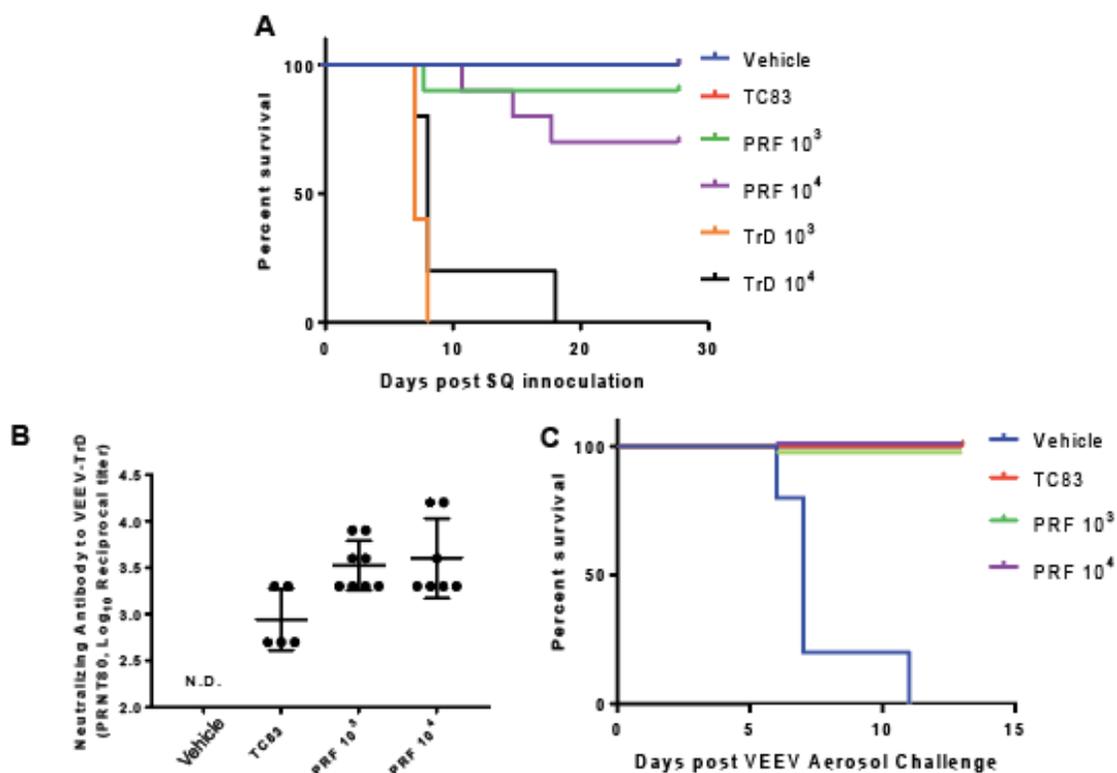


Figure 30. VEEV TrD_{PRFm} induces high levels of neutralizing antibodies and protects mice from VEEV TrD infection. (A) Kaplan Meyer survival plot of mice subcutaneously (SQ) vaccinated/infected with vehicle (N=5), TC83 10⁵ (N=5), PRF (N=10) or TrD (N=5) viruses. Mice were monitored up to 28 days post-infection. (B) Serum samples collected at Day 28 were analyzed to determine neutralizing antibody levels via PRNT₈₀ assay. (C) Kaplan Meyer survival plot of mice challenged with VEEV TrD on Day 29. Vehicle (N=5), TC83 10⁵ (N=5), PRF 10³ (N=9), PRF 10⁴ (N=7).

Further refinements of the current VEEV_{PRFm} model are warranted. An immediate criticism that can be leveraged is the observation that only 70% of infected mice survived aerosol infection of VEEV_{PRFm} (Figure 25A). Additionally, the incidence of reversion to WT VEEV TRD has not been investigated in the fatal cases of VEEV_{PRFm} inoculation. However, as VEEV_{PRFm} is the product of three point mutations to the VEEV slippery site, it is evident that a more robust

attenuated model that covers the entirety of the VEEV -1 PRF signal should be developed. Exploration of the downstream stimulatory structure in both a secondary (Figure 23B) and tertiary (Figure 23E) context will be instrumental in the elucidation of mutations of key regions that interface with ribosomal stalling.

This work was expanded in Chapter 3, with the investigation of CHIKV recoding signals. This work recapitulated the validated -1 PRF described in the new world encephalitic alphaviruses (Figure 27) and also provided the first structural characterization of the CHIKV TCR signal (Figure 29). It was further demonstrated that the CHIKV signals are remarkably conserved across disparate geographic lineages and that the higher order structure of these signals are resilient against incidental polymorphisms between geographic isolates. Taken together with the work conducted in Chapter 2, these results suggest that attenuating mutations to the CHIKV recoding signals will potentially have broad-range effectiveness. While direct experimentation remains needed, these results also raise the possibility that ablative mutations of both the TCR and -1 PRF signal will cumulatively impair the attenuation of CHIKV pathogenesis. If successful, this dual ablation strategy of recoding signals could serve as additional protection against the hazard of wild type reversion.

While beyond the scope of this work, it is also worth mentioning that several key aspects of alphavirus frameshifting remain unresolved. The precise function of the TF -1 PRF product are poorly understood, despite reports of unique post-translational modification compared to 6K¹⁴⁴. Additionally, the negative products of alphavirus -1 PRF have not been sufficiently explored. It is

known that expression of TF necessarily occurs at the expense of structural protein E1, meaning that the latter protein has three to ten percent less abundance than other structural alphavirus proteins. The importance of this discrepancy towards the viability of alphaviruses is unknown, but could potentially be explored in the future through recombinatory experiments. Finally, the previously described studies on alphavirus TCR signals necessarily require that only a small fragment of the sgRNA transcript is represented in a reporter vector. While this approach allows for the definitive identification of minimum structures required for promoting efficient recoding, by definition it cannot identify longer range RNA:RNA interactions along the transcript^{134,135}. At the time of writing, a new technology –SHAPE-MaP- is a promising tool for future whole genome secondary structure analysis²²⁵.

Another consideration worth investigating is the fact that alphavirus transmission predominately occurs through mosquito vectors²⁷. The replication cycle of alphaviruses in mosquito cell tissue is poorly explored in the literature. As a consequence, the degree to which different cell lines exert an influence on the dynamics of programmed translational recoding mechanisms has not been properly explored. Differences in recoding efficiency were observed in both CHIKV and the encephalitic alphaviruses across different cell lines (Figures 23, 27, 28), which lend credence to the hypothesis that viral recoding mechanisms may be optimized for specific host tissues. If specialized alphaviral recoding dynamics can be identified in mosquito cells, directed mutagenesis strategies

can be employed to specifically attenuate replication –and therefore transmission- in primary vectors.

It is important to note that research into the nuances of viral recoding will only be as effective as the methods of detection. As programmed translational recoding efficiencies are primarily quantified in a reporter vector system, the fidelity of these vectors must be continually scrutinized. Between the research conducted in Chapter 2 and Chapter 3, a publication surfaced that identified distortions that can arise from inserted test sequence interference with the flanking luciferase reporter proteins²⁰³. Traditional recoding efficiency is measured as a ratio between downstream firefly luciferase activity to upstream *Renilla* luciferase expression, further normalized against the ratio of an unmodified control plasmid¹⁷². The position of firefly luciferase is adjusted such that it is only expressed consequent to the predicted recoding event of the inserted experimental sequence. In contrast, *Renilla* luciferase will always be expressed regardless of the downstream recoding activity. In the previous pLuci reporter model from which pJD175f was derived, this means that the *Renilla* peptide can be followed by two distinct peptide sequences, either of which capable of exerting an influence on protein folding and subsequent luminescent activity (figure)²⁰³. Thus, firefly expression is potentially normalized against two variants of *Renilla*, leading to potential distortions of data analysis if not outright false positives. An upgraded reporter vector, pSGDluc, has been developed to address these issues, containing self-cleaving intein sequences that separate the luciferase proteins from the experimental sequence (Figure 31). A modification of

this plasmid was used for the CHIKV experiments in chapter 3 (modifications summarized in the Chapter 3 methods section). Critically, the -1 PRF efficiency of VEEV from the modified pSGDluc vector recapitulated the results in chapter 2, implying that distortions arising from the older pLuci model were minimal.

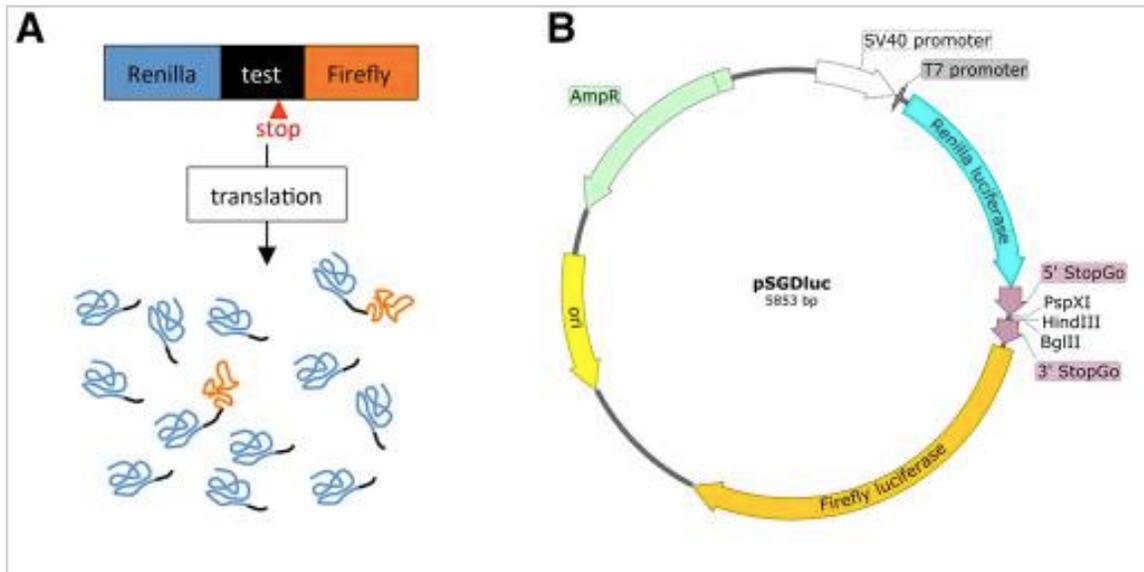


Figure 31. pSGDluc corrects potential sources of confounding distortions from the pLuci reporter vector. (A) General schematic of a dual luciferase assay conducted through the traditional pLuci reporter vector. A test sequence is cloned in between luciferase genes *Renilla* (blue) and Firefly (orange), such that Firefly is only expressed consequent to the respective recoding event. The products of standard and recoded translation are expressed a single polypeptide, linked by the test sequence. The recoding capabilities of the test sequence means that *Renilla* can potentially be followed by two distinct polypeptides, exerting differential influences on protein folding. (B) Plasmid map of the updated pSGDluc reporter vector constructed to address this issue. The cloning site for test sequences are now flanked 5' and 3' by StopGo intein sequences. The peptide sequences of these regions are post-translationally cleaved, allowing for standardized folding of *Renilla* and Firefly independent of test sequence recoding activity. Figure adapted from Louhgran *et al*, 2017²⁰³.

Orthogonal methods of quantification can also be employed to verify viral recoding activities. For example, all recoding experiments in Chapter 3 were supplemented with western blot quantification that recapitulated the dual

luciferase results (Figures 27, 28). Another modality in development is a reporter plasmid that substitutes the *Renilla* and firefly luciferase genes with acGFP and mCherry fluorescent genes, respectively. Preliminary experiments using VEEV and CHIKV -1 PRF signals as test sequences have elicited efficient frameshifting in HEK293T cells, albeit at a lower level than what is achieved with the dual luciferase counterparts (Figure 32). This discrepancy in -1 PRF activity may be the product of undiagnosed issues with this new vector. Currently, HEK293T is the only cell line where these dual fluorescent reporters produce enough protein to be detectable with a luminometer. Further optimization should explore the transfection efficiency of these constructs as well as the possibility of cell-specific effects on the expression of the selected fluorescent proteins. Lastly, it should be noted that future inquiries into attenuating mutations of viral recoding signals will be heavily predicated on the presence of comprehensive structural analysis. Previous computational predictions of encephalitic alphavirus TCR signals, for example, were minimal in scope, only modeling the first helical turn of the stem before abstracting into circular diagrams (Figure 20)¹³⁹. When the same signals are resolved through chemical probing assays, it quickly becomes apparent that the initial predictions omitted considerable detail of the minimum required structures (Figure 33). Direct experimentation, then, will always be the superior method of detecting critical structures for destabilizing mutations.

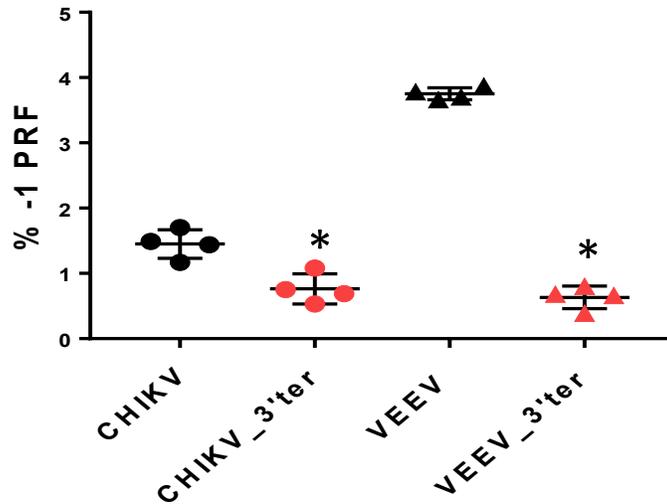


Figure 32. Efficient -1 PRF promoted by CHIKV and VEEV signals in bi-fluorescent (BiFL) reporter vectors. A bi-fluorescent (BiFL) vector (pJD2262) that uses AcGFP and mCherry as reporter sequences was developed as an orthogonal approach for assaying translational recoding efficiencies. HEK293T cells seeded at 1.0×10^5 per well in 6 well plates were transfected with 2.5 ug of either the control BiFL plasmid, or with an inserted CHIKV -1 PRF signal, VEEV -1 PRF signal, or associated mutants containing a termination codon 3' of the insert sequence in the -1 frame (3'ter). Cells were lysed with NP-40 lysis buffer mixed with HALT protease/phosphatase inhibitor and centrifuged 48 hours post-transfection. The supernatant was transferred into a black, clear-bottomed 96 well plate and fluorescence was quantified using the "Fluorescence" setting of the GloMax 96 microplate luminometer. The green (Ex: 525 nm, Em: 580-640 nm) and blue (Ex: 490, Em; 510-650) filters were used to quantify the fluorescent excitation of mCherry and AcGFP, respectively. In order to account for spectral overlap, lysates from cells transfected with a single color AcGFP expression vector were quantified for both AcGFP as well as mCherry fluorescence. A standard curve was generated in order to subtract AcGFP signal bleed over into the mCherry channel. *, $P > 0.05$.

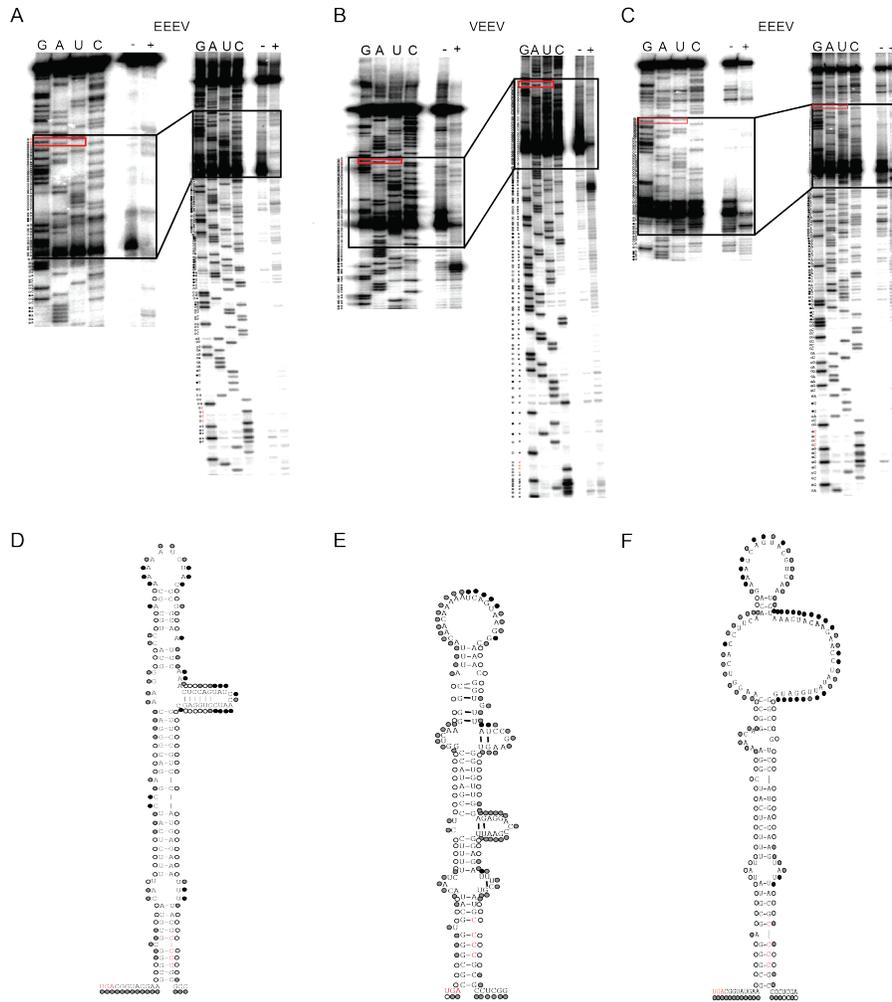


Figure 33. Structural analyses of the EEEV, VEEV and WEEV TCR signals. The TCR signals for (A) EEEV (B) VEEV and (C) WEEV were chemically resolved through selective 2' hydroxyl acylation and primer extension (SHAPE). RNA templates were transcribed from the corresponding dual-luciferase reporter structures and modified with NMIA. $\gamma^{32}\text{P}$ radiolabeled cDNA products were separated through 8% urea-PAGE and visualized via a Fujifilm phosphorimager. Gels have been annotated to indicate the sequencing lanes (G, A, U, C), an untreated control (-) and the NMIA modified experimental lane (+). Circles denote the relative reactivity of bases (white = unreactive; grey = partially reactive; black = highly reactive). Inset images of a longer run are provided for better separation of the 5' region of the sequence. The position of the UGA stop codon is indicated with a red box. (D, E, F) Corresponding structures of EEEV, VEEV and WEEV TCR structures derived from the above SHAPE gels. Circles correspond to the previously described nucleotide reactivities. The position of the 5' UGA sequence is highlighted in red. A 3' region of four Cs that appear on the gels as a single band are also highlighted in red text.

A final point of consideration from this research is the broad range application of its results. If ablation of -1 PRF attenuates the pathogenesis of VEEV, then the possibility is raised that this vulnerability extends to all viruses that rely on translational recoding to regulate gene expression. Pilot experiments pertaining to the identification, characterization and validation of recoding signals of other RNA viruses in the Togavirus, Flavivirus and Retrovirus families have been conducted (Figure 34). While beyond the scope of this thesis, next steps will focus on the identification of destabilizing mutations to these recoding signals as well as investigating potential effects on the replication and pathology of the respective viruses.

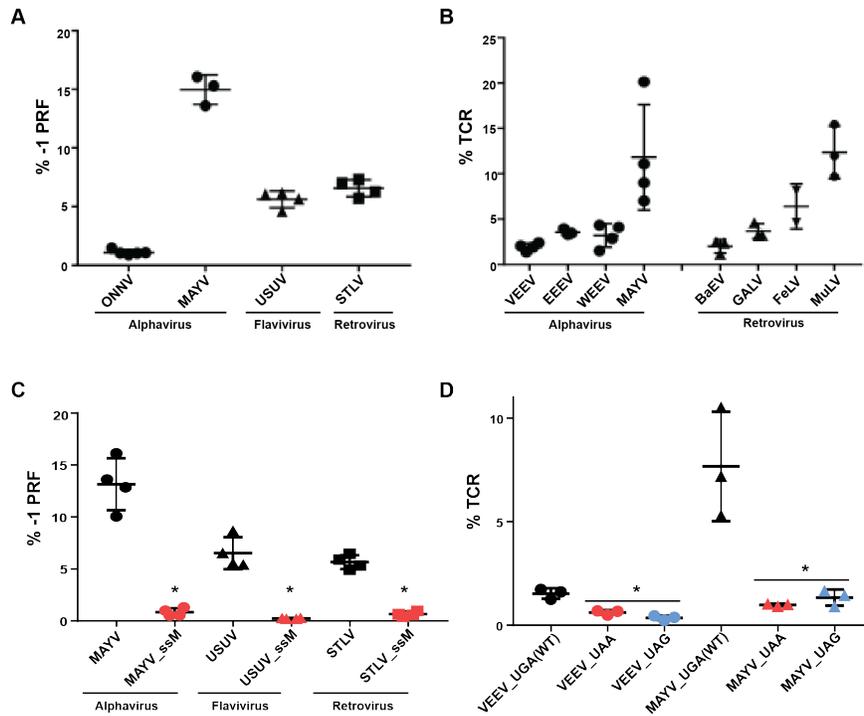


Figure 34. Efficient translational recoding documented in other RNA viruses. (A, B) Predicted -1 PRF and TCR structures from viruses in the Alphavirus (EEEV, VEEV and WEEV; O’Nyong’Nyong virus, ONNV; Mayaro virus, MAYV), Flavivirus (Usutu virus, USUV) and Retrovirus (simian T-cell leukemia virus, STLV; baboon endogenous virus, BaEV; gibbon ape leukemia virus, GaLV; feline leukemia virus, FeLV; murine leukemia virus, MuLV) families were cloned into dual-luciferase reporter vectors and their ability to promote efficient recoding was monitored in HEK293T cells. (C, D) Site-directed mutagenesis was used to introduce silent coding mutations into the slippery site (ssM) of -1 PRF viruses, or substitute the wild type UGA termination codon (WT) with alternative UAA and UAG codons. As this figure summarizes preliminary experiments, the full set of mutants for each virus has not yet been generated. Data is presented as means with standard deviations, where each data point denotes a biological replicate assayed as technical triplicates. * $P < 0.05$.

Appendices

Appendix A. Experimental insert sequences.

(Table 1) 5' and 3' positions of the oligonucleotide sequence are indicated. Lowercase text denotes sequence overlap with plasmid backbone sequence. Uppercase text denotes the respective experimental sequence. Underlined regions indicate the position of the -1 PRF or TCR recoding site. Bold text denotes mutations from the wild type sequence.

Gene Fragment	Sequence	Description
HIV_PRF_WT	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTAGGGAAGA TCTGGCCTTCCCACAAGGGGAGGCCAGGGAATTTT CTTCCGAGCTCGAAGACGCCAAAAACATAAAGAAA GGCCCGGCGCCATTCTATCCTCTAGAGGATGGAAC CGCTGGAGAGCAACTGCATAAGGCTgagctcatggaaga cgccaaaaacataa3'	HIV-1 -1 PRF sequence
CCR5_PRF_WT	5'gcgagttctcaaaaatgaacaaatgtcgacTTTAAAAGCCAGG ACGGTCACCTTTGGGGTGGTGACAAGTGTGATCAC TTGGGTGGTGGCTGTGTTTGCCTCTCTCCAGGAA TCATCTTACCAGATCTCAAAAAGAAGGTCTTCATTA CACCTGCAGCTCTCATTTTCCATACAGTCAGTATCA ATTCTGGAAGAATTTCCAGACATTAAGATAGTCAT CTTGGGgagctcatggaagacgccaaaaacataa3'	CCR5 -1 PRF sequence
EEEEV_PRF_WT	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTACTTGTCT GCGGCGCCTTGGGCGCCGCAGCGTACGAACACAC AGCAGTGATGCCGAACAAGGTGGGGATCCCGTACA AAGCTTTAGTCGAACGCCCAGGTTATGCACCCGTTT ACCTACAgagctcatggaagacgccaaaaacataa3'	EEEEV -1 PRF sequence
EEEEV_PRF_ssM	5'gcgagttctcaaaaatgaacaaatgtcgac GTTCTTG CTTGTCT GCGGCGCCTTGGGCGCCGCAGCGTACGAACACAC AGCAGTGATGCCGAACAAGGTGGGGATCCCGTACA AAGCTTTAGTCGAACGCCCAGGTTATGCACCCGTTT ACCTACAgagctcatggaagacgccaaaaacataa3'	EEEEV -1 PRF sequence with U UUU UUA -> G UUC UUG
EEEEV_PRF_5'ter	5'gcgagttctcaaaaatgaacaaatgtcgac ATAA TTTTTTACT TGTCTGCGGCGCCTTGGGCGCCGCAGCGTACGAA CACACAGCAGTGATGCCGAACAAGGTGGGGATCCC GTACAAAGCTTTAGTCGAACGCCCAGGTTATGCACC CGTTCACCTACAgagctcatggaagacgccaaaaacataa3'	EEEEV -1 PRF sequence with 5' UAA codon
EEEEV_PRF_3'ter	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTACTTGTCT GCGGCGCCTTGGGCGCCGCAGCGTACGAACACAC AGCAGTGATGCCGAACAAGGTGGGGATCCCGTACA AAGCTTTAGTCGAACGCCCAGGTTATGCACCCGTTT ACCTACAgagctcatggaagacgcttaaaacataa3'	EEEEV -1 PRF sequence with 3' UAA codon in -1 frame
tEEEEV_PRF	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTACTTGTCT	Truncation of

	GCGGCGCCTTGGGCGCCGCAGCGTACGAACACAC AGCAGTGATGCCGAACAAGGTGGGGATCCCCTACA AAGCgagctcatggaagacgcaaaaacataa3'	EEEV_PRF_W T
t2EEEV_PRF	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTACTTGTCT GCGGCGCCTTGGGCGCCGCAGCGagctcatggaagacg caaaaacataa3'	Truncation of tEEEV_PRF
VEEV_PRF_WT	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTAGTCGTGG CCGGCGCCGCAGGCGCCGGCGCCTACGAGCACGC GACCACGATGCCGAGCCAAGCGGGAATCTCGTACA ACACCATAGTCAACAGAGCAGGCTACGCGCCACTga gctcatggaagacgcaaaaacataa3'	VEEV -1 PRF sequence
VEEV_PRF_ssM	5'gcgagttctcaaaaatgaacaaatgtcgacGTTCTTGTCGTG GCCGGCGCCGCAGGCGCCGGCGCCTACGAGCACG CGACCACGATGCCGAGCCAAGCGGGAATCTCGTAC AACACCATAGTCAACAGAGCAGGCTACGCGCCACT gagctcatggaagacgcaaaaacataa3'	VEEV -1 PRF sequence with U UUU UUA -> G UUC UUG
VEEV_PRF_5'ter	5'gcgagttctcaaaaatgaacaaatgtcgacATAATTTTTAGTC GTGGCCGGCGCCGCAGGCGCCGGCGCCTACGAGC ACGCGACCACGATGCCGAGCCAAGCGGGAATCTC GTACAACACCATAGTCAACAGAGCAGGCTACGCGC CACTgagctcatggaagacgcaaaaacataa3'	VEEV -1 PRF sequence with 5' UAA codon
VEEV_PRF_3'ter	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTAGTCGTGG CCGGCGCCGCAGGCGCCGGCGCCTACGAGCACGC GACCACGATGCCGAGCCAAGCGGGAATCTCGTACA ACACCATAGTCAACAGAGCAGGCTACGCGCCACTga gctcatggaagacgcctaaaacataa3'	VEEV -1 PRF sequence with 3' UAA codon in -1 frame
tVEEV_PRF	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTAGTCGTGG CCGGCGCCGCAGGCGCCGGCGCCTACGAGCACGC GACCAC gagctcatggaagacgcaaaaacataa3'	Truncation of VEEV_PRF_W T
WEEV_PRF_WT	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTATTGGTTG CAGGCGTCTGCCTGGGGAAGGTAGACGCCTTCGAA CATGCGACCACTGTGCCAAATGTTCCGGGGATCCC GTAAGagctcatggaagacgcaaaaacataa3'	WEEV -1 PRF sequence
WEEV_PRF_ssM	5'gcgagttctcaaaaatgaacaaatgtcgacGTTCTTGTTGGTTG CAGGCGTCTGCCTGGGGAAGGTAGACGCCTTCGAA CATGCGACCACTGTGCCAAATGTTCCGGGGATCCC GTAAGagctcatggaagacgcaaaaacataa3'	WEEV -1 PRF sequence with U UUU UUA -> G UUC UUG
WEEV_PRF_5'ter	5'gcgagttctcaaaaatgaacaaatgtcgacATAATTTTTATT GGTTGCAGGCGTCTGCCTGGGGAAGGTAGACGCC TTCGAACATGCGACCACTGTGCCAAATGTTCCGGG GATCCCGTAAgagctcatggaagacgcaaaaacataa3'	WEEV -1 PRF sequence with 5' UAA codon
WEEV_PRF_3'ter	5'gcgagttctcaaaaatgaacaaatgtcgacTTTTTTATTGGTTG CAGGCGTCTGCCTGGGGAAGGTAGACGCCTTCGAA CATGCGACCACTGTGCCAAATGTTCCGGGGATCCC GTAAGagctcatggaagacgcctaaaacataa3'	WEEV -1 PRF sequence with 3' UAA codon in the -1 frame

tWEEV_PRF	5'ggagattctcaaaaatgaacaaatgtcgacTTTTTTATTGGTTG CAGGCGTCTGCCTGGGGAAGGTAGACGCCTTCGAA CATGCGACCACTGTGCCAAAagagctcatggaagacgcaaaa aacataa3'	Truncation of WEEV_PRF_W T
MuLV_TCR_WT	5'ggagacgtcgagtccaacccgggcccctcgacGACCCTAGA T <u>G</u> ACTAGGGAGGTCAGGGTCAGGAGCCCCCCT GAACCCAGGATAACCCTCAAAGTCGGGGGGCAACC CGTCACCTTCTGGTAGATACTGGGGCCCAAggatcc gaggcacggcataagcaaaagatcg3'	MuLV TCR sequence
CHIKV_Af/As_TCR _WT	5'ggagacgtcgagtccaacccgggcccctcgacCGACGAGTT ATGACTAGACAGGGCAGGTGGGTATATATTCTCGTC GGACACCGGTCCAGGTCATTTACAACAGAAGTCAG TACGCCAGTCAGTGCTGCCGGTGAACACCCTGGAG GAAGTCCACGAGGAGAAGTGTTACCCACCTAAGCT Gggatccgaggcacggcataagcaaaagatcg3'	CHIKV Africa/Asia TCR consensus sequence
CHIKV_Af/As_TCR _UAA	5'ggagacgtcgagtccaacccgggcccctcgacCGACGAGTT ATAA CTAGACAGGGCAGGTGGGTATATATTCTCGTC GGACACCGGTCCAGGTCATTTACAACAGAAGTCAG TACGCCAGTCAGTGCTGCCGGTGAACACCCTGGAG GAAGTCCACGAGGAGAAGTGTTACCCACCTAAGCT Gggatccgaggcacggcataagcaaaagatcg3'	CHIKV Africa/Asia TCR consensus sequence with UGA -> UAA
CHIKV_Af/As_TCR _UAG	5'ggagacgtcgagtccaacccgggcccctcgacCGACGAGTT ATAG CTAGACAGGGCAGGTGGGTATATATTCTCGT CGGACACCGGTCCAGGTCATTTACAACAGAAGTCA GTACGCCAGTCAGTGCTGCCGGTGAACACCCTGGA GGAAGTCCACGAGGAGAAGTGTTACCCACCTAAGC TGggatccgaggcacggcataagcaaaagatcg3'	CHIKV Africa/Asia TCR consensus sequence with UGA -> UAG
CHIKV_Af/As_TCR _AGA	5'ggagacgtcgagtccaacccgggcccctcgacCGACGAGTT AAGA CTAGACAGGGCAGGTGGGTATATATTCTCGT CGGACACCGGTCCAGGTCATTTACAACAGAAGTCA GTACGCCAGTCAGTGCTGCCGGTGAACACCCTGGA GGAAGTCCACGAGGAGAAGTGTTACCCACCTAAGC TGggatccgaggcacggcataagcaaaagatcg3'	CHIKV Africa/Asia TCR consensus sequence with UGA -> AGA
CHIKV_Carib_TCR _WT	5'ggagacgtcgagtccaacccgggcccctcgacCGACGAGTT ATGACTAGACAGGGCAGGTGGGTATATATTCTCGTC GGACACTGGTCCAGGCCATTTACAACAGAAGTCGG TACGCCAGTCAGTGCTGCCGGTAAACACCCTGGAG GAAGTTCACGAGGAGAAGTGTTACCCACCTAAGCT Gggatccgaggcacggcataagcaaaagatcg3'	CHIKV Caribbean TCR consensus sequence
CHIKV_Carib_TCR _UAA	5'ggagacgtcgagtccaacccgggcccctcgacCGACGAGTT ATAA CTAGACAGGGCAGGTGGGTATATATTCTCGTC GGACACTGGTCCAGGCCATTTACAACAGAAGTCGG TACGCCAGTCAGTGCTGCCGGTAAACACCCTGGAG GAAGTTCACGAGGAGAAGTGTTACCCACCTAAGCT Gggatccgaggcacggcataagcaaaagatcg3'	CHIKV Caribbean TCR consensus sequence with UGA -> UAA
CHIKV_Carib_TCR _UAG	5'ggagacgtcgagtccaacccgggcccctcgacCGACGAGTT ATAG CTAGACAGGGCAGGTGGGTATATATTCTCGT CGGACACTGGTCCAGGCCATTTACAACAGAAGTCG GTACGCCAGTCAGTGCTGCCGGTAAACACCCTGGA	CHIKV Caribbean TCR consensus sequence with

	GGAAGTTCACGAGGAGAAGTGTTACCCACCTAAGC TGgatccgaggcacggcataagcaaaagatcg3'	UGA -> UAG
CHIKV_Carib_TCR _AGA	5'ggagacgtcgagtccaacccgggccctcgacCGACGAGTT AAG ACTAGACAGGGCAGGTGGGTATATATTCTCGT CGGACACTGGTCCAGGCCATTTACAACAGAAGTCG GTACGCCAGTCAGTGCTGCCGGTAAACACCTGGA GGAAGTTCACGAGGAGAAGTGTTACCCACCTAAGC TGgatccgaggcacggcataagcaaaagatcg3'	CHIKV Caribbean TCR consensus sequence with UGA -> AGA
VEEV_PRF_WT	5'ggagacgtcgagtccaacccgggccctcgacTTTTTAGTC GTGGCCGGCCGCGCAGGCCGCGCCTACGAGC ACGCGACCACGATGCCGAGCCAAGCGGGAATCTC GTACAACACCATAGTCAACAGAGCAGGCTACGCGC CACTggatccgaggcacggcataagcaaaagatcg3'	VEEV -1 PRF sequence
CHIKV_PRF_WT	5'ggagacgtcgagtccaacccgggccctcgacTTTTTAGCC GTAATGAGCGTCGGTGCCCACTGTGAGCGCGTA CGAACACGTAACAGTGATCCCGAACACGGTGGGgg atccgaggcacggcataagcaaaagatcg3'	CHIKV -1 PRF sequence
CHIKV_PRF_ssM	5'ggagacgtcgagtccaacccgggccctcgac GTTCTTG GC CGTAATGAGCGTCGGTGCCCACTGTGAGCGCGT ACGAACACGTAACAGTGATCCCGAACACGGTGGGg gatccgaggcacggcataagcaaaagatcg3'	CHIKV -1 PRF sequence with U UUU UUA -> G UUC UUG
ONNV_PRF_WT	5'ggagacgtcgagtccaacccgggccctcgacTTTTTAGCC GTCATGAGCATCGGTGCCCGCACTGTGACCGCGTA CGAGCACGCAACAGTGATCCCGAACACGGTGGGA GTACCGTGgatccgaggcacggcataagcaaaagatcg3'	ONNV -1 PRF sequence
MAYV_PRF_WT	5'ggagacgtcgagtccaacccgggccctcgacTTTTTAGTC GCAATGAGCATCGGGAGTGCCGTTGCCAGTGCTTA CGAGCACACGGCAATCATTCCGAACCAAGTGGGAT TCCCGTATGGgatccgaggcacggcataagcaaaagatcg3'	MAYV -1 PRF sequence
MAYV_PRF_ssM	5'ggagacgtcgagtccaacccgggccctcgac GTTCTTG GT CGCAATGAGCATCGGGAGTGCCGTTGCCAGTGCTT ACGAGCACACGGCAATCATTCCGAACCAAGTGGGA TTCCCGTATGGgatccgaggcacggcataagcaaaagatcg3'	MAYV -1 PRF sequence with U UUU UUA -> G UUC UUG
USUV_PRF_WT	5'ggagacgtcgagtccaacccgggccctcgacTCCTTTTCA GTTGGGCCTTCTGGTGATGTTTCTGGCCACCCAGG AGGTCCTGAGGAAGAGGTGGACGGCCAGATTGACT GTTCCGGCTATTGTGGGAGCTCTACTCGTggatccgag gcacggcataagcaaaagatcg3'	USUV -1 PRF sequence
USUV_PRF_ssM	5'ggagacgtcgagtccaacccgggccctcgac GCCCTTG CA GTTGGGCCTTCTGGTGATGTTTCTGGCCACCCAGG AGGTCCTGAGGAAGAGGTGGACGGCCAGATTGACT GTTCCGGCTATTGTGGGAGCTCTACTCGTggatccgag gcacggcataagcaaaagatcg3'	USUV -1 PRF sequence where U CCU UUC -> G CCC UUG
STLV_PRF_WT	5'ggagacgtcgagtccaacccgggccctcgacAAAAAACTC CATAGGGGGGGAGGTCTAACCTCCCCCCCACATT	STLV -1 PRF sequence

	ACGGCAAGTCCTTCCCAACCAAGGCCCGGCATCTA TTCTGCCAGTTATACCGTTAGATCCCGCCCGCCGG CCCATAATTggatccgaggcacggcataagcaaaagatcg3'	
STLV_PRF_ssM	5'ggagacgtcgagtccaacccgggcccctcgac GAACAAGTC CATAGGGGGGGAGGTCTAACCTCCCCCCCACATT ACGGCAAGTCCTTCCCAACCAAGGCCCGGCATCTA TTCTGCCAGTTATACCGTTAGATCCCGCCCGCCGG CCCATAATTggatccgaggcacggcataagcaaaagatcg3'	STLV -1 PRF sequence where A AAA AAC -> G AAC AAG
VEEV_TCR_WT	5'ggagacgtcgagtccaacccgggcccctcgacACAACAACA ATGACGGTTTGACGCGGGTGCATACATCTTTTCCTC CGATACCGGTCAAGGGCATTTACAACAAAAATCAGT AAGGCAAACGGTGTATCCGAAGTGGTGTGGAGA GGACCGAATTGGAGATTCGTATGCCCGCGCCTC ggatccgaggcacggcataagcaaaagatcg3'	VEEV TCR sequence
VEEV_TCR_UAA	5'ggagacgtcgagtccaacccgggcccctcgacACAACAACA ATGACGGTT TAA CGCGGGTGCATACATCTTTTCCTC CGATACCGGTCAAGGGCATTTACAACAAAAATCAGT AAGGCAAACGGTGTATCCGAAGTGGTGTGGAGA GGACCGAATTGGAGATTCGTATGCCCGCGCCTC ggatccgaggcacggcataagcaaaagatcg3'	VEEV TCR sequence where UGA -> UAA
VEEV_TCR_UAG	5'ggagacgtcgagtccaacccgggcccctcgacACAACAACA ATGACGGTT TAG CGCGGGTGCATACATCTTTTCCTC CGATACCGGTCAAGGGCATTTACAACAAAAATCAGT AAGGCAAACGGTGTATCCGAAGTGGTGTGGAGA GGACCGAATTGGAGATTCGTATGCCCGCGCCTC ggatccgaggcacggcataagcaaaagatcg3'	VEEV TCR sequence where UGA
EEEV_TCR_WT	5'ggagacgtcgagtccaacccgggcccctcgacGAGGCACTC GAAT TG ACGGTACGAAGCGGGCGGTACATTTTCT CATCCGAGACGGGACAAGGGCACCTGCAGCAAAAA TCTACGCGCAATGCAAACCTCCAGTATCCAATCCTG GAGCGTTCCGTCCATGAGAAATTTACGCCCGCGG CCTCGATggatccgaggcacggcataagcaaaagatcg3'	EEEV TCR sequence
WEEV_TCR_WT	5'ggagacgtcgagtccaacccgggcccctcgacTCAACACTC CAACT GA CGGTATGAAGCGGGAGCGTATATTTTCTC ATCGGAAACAGGCCAAGGTCACCTTCAACAGAAAT CAGTACGTCAATGTAACTACAAGAACCTATATTGG ATCGGGCCGTCCATGAGAAGTATTACGCCCGCGC CTCGATggatccgaggcacggcataagcaaaagatcg3'	WEEV TCR sequence
MAYV_TCR_WT	5'ggagacgtcgagtccaacccgggcccctcgacTAGTAATTCA TCT TG ACTAGCCGTGCGGGGGCCTATATTTTCTCA TCCGACGTCCGTCCAGGGCACCTGCAACAGAAATC AGTGAGGCAGCATGACTTAGAGGTGCCGATTATGG ATCGTGTGATTGAGGAAAAGGTCTACCCGCCTAAAT TAGATggatccgaggcacggcataagcaaaagatcg3'	MAYV TCR sequence
MAYV_TCR_UAA	5'ggagacgtcgagtccaacccgggcccctcgacTAGTAATTCA TCT TAA CTAGCCGTGCGGGGGCCTATATTTTCTCA TCCGACGTCCGTCCAGGGCACCTGCAACAGAAATC	MAYV TCR sequence with UGA -> UAA

	AGTGAGGCAGCATGACTTAGAGGTGCCGATTATGG ATCGTGTGATTGAGGAAAAGGTCTACCCGCCTAAAT TAGATggatccgaggcagggcataagcaaaagatcg3'	
MAYV_TCR_UAG	5'ggagacgtcgagtccaacccccgggccctcgacTAGTAATTCA TCT TAG CTAGGCCGTGCGGGGCCTATATTTTCTCA TCCGACGTCGGTCCAGGGCACCTGCAACAGAAATC AGTGAGGCAGCATGACTTAGAGGTGCCGATTATGG ATCGTGTGATTGAGGAAAAGGTCTACCCGCCTAAAT TAGATggatccgaggcagggcataagcaaaagatcg3'	MAYV TCR sequence with UGA -> UAG
BaEV_TCR_WT	5'ggagacgtcgagtccaacccccgggccctcgacTGAGGACAG CGAAT TAG GGGTGTCAGGGCTCTGGAGCCCCCCC GAGCCCCGGCTAACTCTATCTGTAGGGGGGCATCC CACCACCTTCTTGGTGGACACAGGCGCCCAAggatcc gaggcagggcataagcaaaagatcg3'	BaEV TCR sequence
GALV_TCR_WT	5'ggagacgtcgagtccaacccccgggccctcgacAGCCCTAGA TAACT TAG GGGAGTCAGGGTTCGGACCCCCTCCCCG AACCTAGGGTAACACTGACTGTGGAGGGGACCCCC ATTGAGTTCCTGGTGCACACCGGAGCTGAAggatccg aggcagggcataagcaaaagatcg3'	GALV TCR sequence
FeLV_TCR_WT	5'ggagacgtcgagtccaacccccgggccctcgacCAACTTAGG AGATT TAG GAGAGTCAGGGCCAGGACCCCCCCCCT GAGCCCAGGATAACCTTAAAAATAGGGGGGCAACC GGTACTTTCCTGGTGGACACGGGAGCCCAGggatc cgaggcagggcataagcaaaagatcg3'	FeLV TCR sequence

Appendix B. Plasmids used in this work

(Table 2) List of plasmid names and descriptions that were used in this dissertation.

Plasmid	Description
pJD175f	pLuci derived dual luciferase readthrough control
pJD187	Dual luciferase reporter with inserted HIV-1 -1 PRF signal, pJD175f backbone
pJD827	Dual luciferase reporter with inserted CCR5 -1 PRF signal, pJD175f backbone
pJD1911	Dual luciferase reporter with inserted EEEV -1 PRF signal, pJD175f backbone
pJD1622	Dual luciferase reporter with inserted EEEV -1 PRF signal and GUUCUUG slip site mutation, pJD175f backbone
pJD1623	Dual luciferase reporter with inserted EEEV -1 PRF signal and 5' termination codon, pJD175f backbone
pJD1624	Dual luciferase reporter with inserted EEEV -1 PRF signal and 3' termination codon, pJD175f backbone
pJD1639	Dual luciferase reporter with inserted tEEEV -1 PRF signal, pJD175f backbone
pJD1640	Dual luciferase reporter with inserted t2EEEV -1 PRF signal, pJD175f backbone
pJD1910	Dual luciferase reporter with inserted VEEV -1 PRF signal, pJD175f backbone
pJD1625	Dual luciferase reporter with inserted VEEV -1 PRF signal and GUUCUUG slip site mutation, pJD175f backbone
pJD1626	Dual luciferase reporter with inserted VEEV -1 PRF signal and 5' termination codon, pJD175f backbone
pJD1627	Dual luciferase reporter with inserted VEEV -1 PRF signal and 3' termination codon, pJD175f backbone
pJD1641	Dual luciferase reporter with inserted tVEEV -1 PRF signal, pJD175f backbone
pJD1928	Dual luciferase reporter with inserted WEEV -1 PRF signal, pJD175f backbone
pJD1628	Dual luciferase reporter with inserted WEEV -1 PRF signal and GUUCUUG slip site mutation, pJD175f backbone
pJD1629	Dual luciferase reporter with inserted WEEV -1 PRF signal and 5' termination codon, pJD175f backbone
pJD1630	Dual luciferase reporter with inserted WEEV -1 PRF signal 3' termination codon, pJD175f backbone
pJD1642	Dual luciferase reporter with inserted tWEEV -1 PRF signal, pJD175f backbone
pJD2257	Modified pSGDluc (dual luciferase with inteins) readthrough control
pJD2267	Modified pSGDluc (dual luciferase with inteins) with 5' termination codon upstream of multiple cloning site,
pJD2269	Modified pSGDluc (dual luciferase with inteins) with 3' -1 termination codon downstream of multiple cloning site,
pJD2231	Modified pSGDluc (dual luciferase with inteins) with MuLV TCR insert, pJD2257 backbone
pJD1843	Modified pSGDluc (dual luciferase with inteins) with CHIKV Africa/Asia TCR consensus insert, pJD2257 backbone
pJD2245	Modified pSGDluc (dual luciferase with inteins) with CHIKV Africa/Asia TCR consensus insert and 5' termination codon, pJD2257 backbone
pJD2246	Modified pSGDluc (dual luciferase with inteins) with CHIKV Africa/Asia TCR consensus insert and 3' termination codon, pJD2257 backbone
pJD1829	Modified pSGDluc (dual luciferase with inteins) with CHIKV Africa/Asia TCR consensus insert and stop codon mutated to UAA, pJD2257 backbone
pJD1830	Modified pSGDluc (dual luciferase with inteins) with CHIKV Africa/Asia TCR consensus insert and stop codon mutated to UAG, pJD2257 backbone

pJD1831	Modified pSGDluc (dual luciferase with inteins) with CHIKV Africa/Asia TCR consensus insert with stop codon mutated to AGA (Arginine), pJD2257 backbone
pJD1844	Modified pSGDluc (dual luciferase with inteins) with CHIKV Caribbean TCR consensus TCR insert, pJD2257 backbone
pJD2247	Modified pSGDluc (dual luciferase with inteins) with CHIKV Caribbean TCR consensus TCR insert and 5' termination codon, pJD2257 backbone
pJD2248	Modified pSGDluc (dual luciferase with inteins) with CHIKV Caribbean TCR consensus TCR insert and 3' -1 termination codon, pJD2257 backbone
pJD1832	Modified pSGDluc (dual luciferase with inteins) with CHIKV Caribbean TCR consensus TCR insert and stop codon mutated to UAA, pJD2257 backbone
pJD1833	Modified pSGDluc (dual luciferase with inteins) with CHIKV Caribbean TCR consensus TCR insert and insert stop codon mutated to UAG, pJD2257 backbone
pJD1834	Modified pSGDluc (dual luciferase with inteins) with CHIKV Caribbean TCR consensus TCR insert and stop codon mutated to AGA (Arginine), pJD2257 backbone
pJD2360	Modified pSGDluc (dual luciferase with inteins) with VEEV -1 PRF insert, pJD2257 backbone
pJD2288	Modified pSGDluc (dual luciferase with inteins) with CHIKV -1 PRF insert, pJD2257 backbone
pJD2289	Modified pSGDluc (dual luciferase with inteins) with CHIKV -1 PRF insert and upstream 5' premature terminator, pJD2267 backbone
pJD2378	Modified pSGDluc (dual luciferase with inteins) with CHIKV -1 PRF insert and 3' -1 termination codon, pJD2269
pJD2368	Modified pSGDluc (dual luciferase with inteins) with CHIKV -1 PRF insert and GUUCUUG slip-site mutation, pJD2257 backbone
pJD2261	Bifluorescent readthrough vector
pJD2266	Bifluorescent readthrough vector with 3' termination codon
pJD2361	Bifluorescent readthrough vector with CHIKV -1 PRF insert, pJD2261 backbone
pJD2362	Bifluorescent readthrough vector with CHIKV -1 PRF insert and 3' termination codon, pJD2266 backbone
pJD2365	Bifluorescent readthrough vector with VEEV -1 PRF insert, pJD2261 backbone
pJD2366	Bifluorescent readthrough vector with VEEV -1 PRF insert and 3' termination codon, pJD2266 backbone
pJD2215	Modified pSGDluc (dual luciferase with inteins) with ONNV -1 PRF insert, pJD2257 backbone
pJD2229	Modified pSGDluc (dual luciferase with inteins) with MAYV -1 PRF insert, pJD2257 backbone
pJD2238	Modified pSGDluc (dual luciferase with inteins) with MAYV -1 PRF insert and GUUCUUG slip-site mutation, pJD2257 backbone
pJD2220	Modified pSGDluc (dual luciferase with inteins) with USUV -1 PRF insert, pJD2257 backbone
pJD2239	Modified pSGDluc (dual luciferase with inteins) with USUV -1 PRF insert and GCCCTTG slip-site mutation, pJD2257 backbone
pJD2218	Modified pSGDluc (dual luciferase with inteins) with STLV -1 PRF insert, pJD2257 backbone
pJD2240	Modified pSGDluc (dual luciferase with inteins) with STLV -1 PRF insert GAACAAG, pJD2257 backbone
pJD2226	Modified pSGDluc (dual luciferase with inteins) with VEEV TCR insert, pJD2257 backbone
pJD2243	Modified pSGDluc (dual luciferase with inteins) with VEEV TCR insert and insert stop codon mutated to UAA, pJD2257 backbone
pJD2244	Modified pSGDluc (dual luciferase with inteins) with VEEV TCR insert and insert stop codon mutated to UAG, pJD2257 backbone
pJD2227	Modified pSGDluc (dual luciferase with inteins) with EEEV TCR insert, pJD2257 backbone

pJD2228	Modified pSGDluc (dual luciferase with inteins) with WEEV TCR insert, pJD2257 backbone
pJD2229	Modified pSGDluc (dual luciferase with inteins) with MAYV TCR insert, pJD2257 backbone
pJD2241	Modified pSGDluc (dual luciferase with inteins) with MAYV TCR insert and insert stop codon mutated to UAA, pJD2257 backbone
pJD2242	Modified pSGDluc (dual luciferase with inteins) with MAYV TCR insert and insert stop codon mutated to UAG, pJD2257 backbone
pJD2232	Modified pSGDluc (dual luciferase with inteins) with BaEV TCR insert, pJD2257 backbone
pJD2234	Modified pSGDluc (dual luciferase with inteins) with GALV TCR insert, pJD2257 backbone
pJD2234	Modified pSGDluc (dual luciferase with inteins) with FeLV TCR insert, pJD2257 backbone

Appendix C. Supplementary Figures for Chapter 3

Figure 35. Multiple sequence alignments of CHIKV TCR and -1 PRF sequences from 14 African/Asian (Af/As) strains (Top), and 5 Caribbean (Carib) strains.

CHIKV Af/As

TCR

```
EU564335 5645 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
EF027137 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
EF027134 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
HM045801 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
GU013528 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
FJ513628 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
FJ445433 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
FJ445502 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
FJ807896 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
HM045823 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
EF012359 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
HM045812 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
KJ941050 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
HM045811 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACCGGUCCAGGUCAUUUACAACAGAA
*****

EU564335 5785 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
EF027137 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
EF027134 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
HM045801 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
GU013528 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
FJ513628 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
FJ445433 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
FJ445502 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
FJ807896 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
HM045823 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
EF012359 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
HM045812 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
KJ941050 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
HM045811 GUCAGUACGCCAGUCAGUCGUCGGUGAACACCCUGGAGGAAGUCCACGAGGAGAAGUUAUACCCUUAAGCUG
*****
```

-1 PRF

```
EU564335 9951 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
EF027137 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
EF027134 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
HM045801 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
GU013528 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
FJ513628 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
FJ445433 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
FJ445502 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
FJ807896 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
HM045823 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
EF012359 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
HM045812 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
KJ941050 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
HM045811 UUUUUUAGCCGUAUAGAGCAUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
*****
```

CHIKV Carib

TCR

```
KX702401 5645 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACUGGUCCAGGCCAUUUACAACAGAA
KY415978 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACUGGUCCAGGCCAUUUACAACAGAA
KY415985 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACUGGUCCAGGCCAUUUACAACAGAA
KR046232 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACUGGUCCAGGCCAUUUACAACAGAA
KR046228 CGACGAGUUUAGCAUAGACAGGGCAAGGUGGGUUAUUAUUCUCGUCGGACACUGGUCCAGGCCAUUUACAACAGAA
*****

KX702401 5785 GUCGGUACGCCAGUCAGUCGUCGGUUAACACCCUGGAGGAAGUUCACGAGGAGAAGUUAUACCCUUAAGCUG
KY415978 GUCGGUACGCCAGUCAGUCGUCGGUUAACACCCUGGAGGAAGUUCACGAGGAGAAGUUAUACCCUUAAGCUG
KY415985 GUCGGUACGCCAGUCAGUCGUCGGUUAACACCCUGGAGGAAGUUCACGAGGAGAAGUUAUACCCUUAAGCUG
KR046232 GUCGGUACGCCAGUCAGUCGUCGGUUAACACCCUGGAGGAAGUUCACGAGGAGAAGUUAUACCCUUAAGCUG
KR046228 GUCGGUACGCCAGUCAGUCGUCGGUUAACACCCUGGAGGAAGUUCACGAGGAGAAGUUAUACCCUUAAGCUG
*****
```

-1 PRF

```
KX702401 9951 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
KY415978 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
KY415985 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
KR046232 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
KR046228 UUUUUUAGCCGUAUAGAGCGUCGGUGCCACACUGUGAGCGCGUACGAACACGUAACAGUGAUCCCGAACCGUAAACAGUGAUCCCGAACCGGUGGG
*****
```

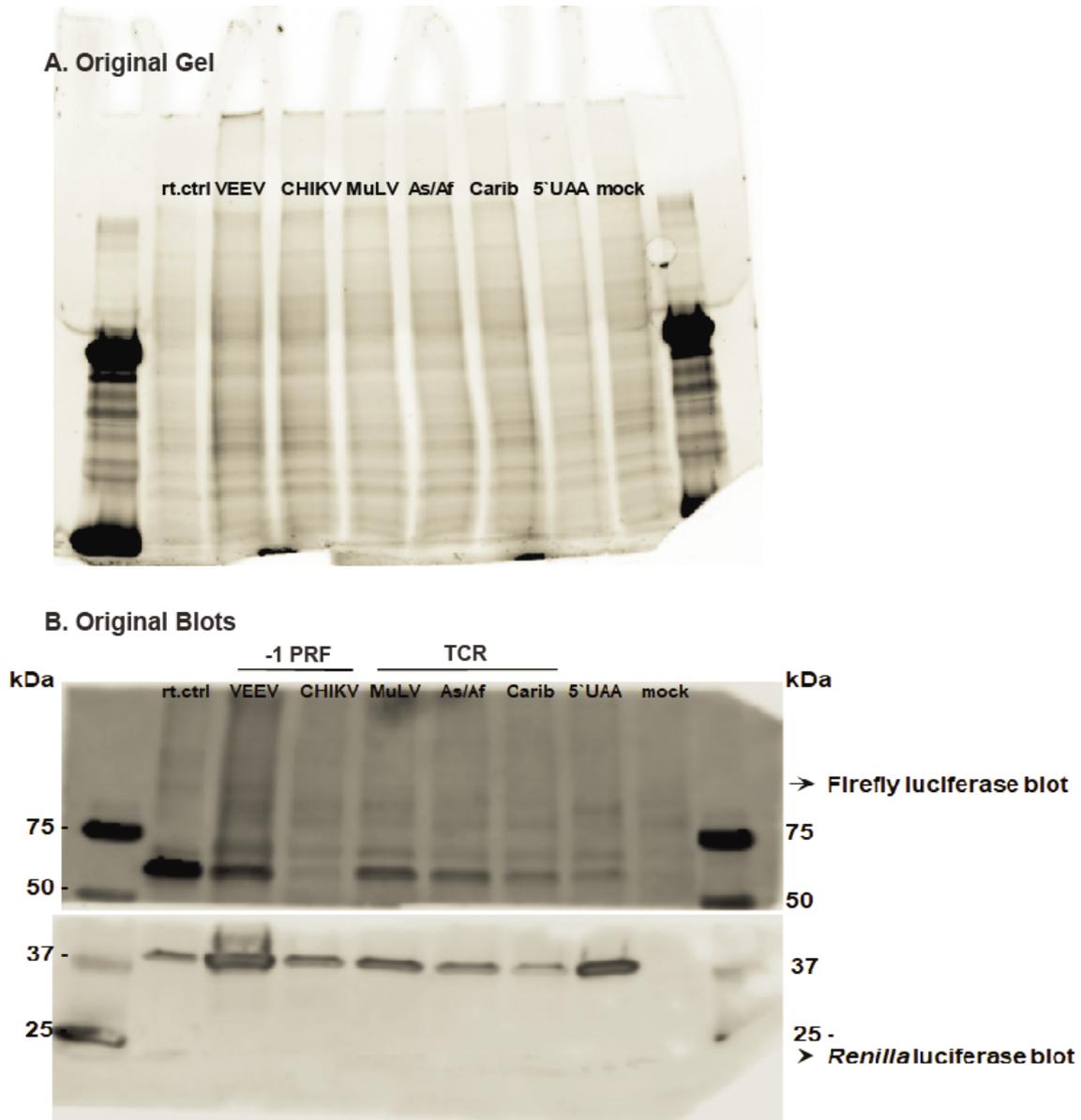
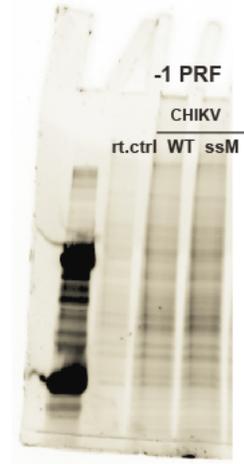
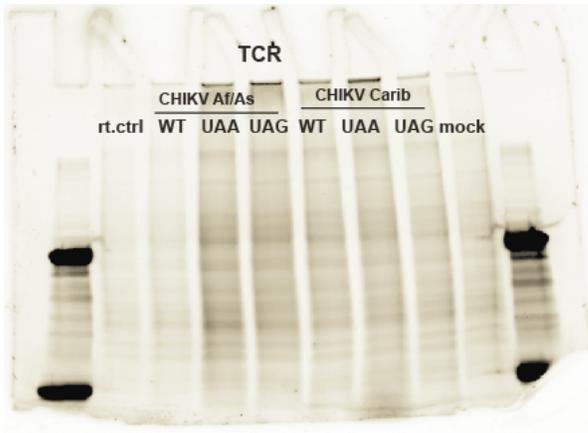
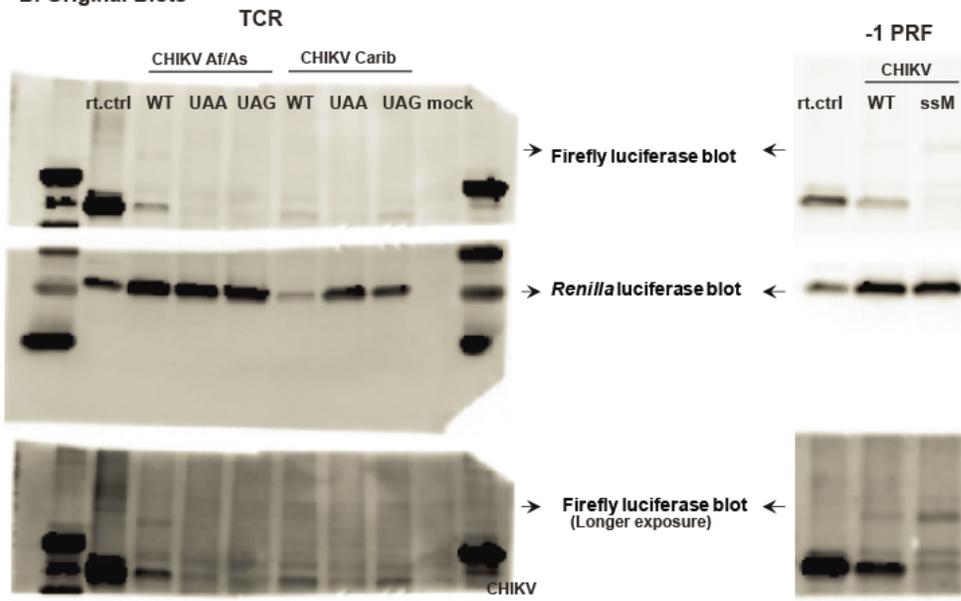


Figure 36. A. Total protein of samples separated through 4-20% SDS-PAGE gel. 10 μ g of readthrough control (rt.ctrl) sample and 15-20 μ g other samples were loaded for immunoblot analysis. **B.** Raw images of Firefly luciferase (short and long exposures) and *Renilla* luciferase immunoblots of protein lysates from HEK293T cells transiently transfected with the indicated reporter plasmids. The lanes on either end is molecular weight marker (Precision Plus Protein™ Kaleidoscope™ Bio-Rad #1610375).

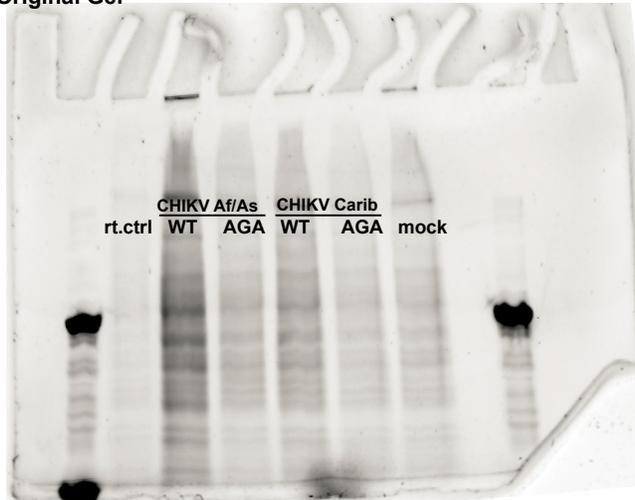
A. Original Gels



B. Original Blots



C. Original Gel



D. Original Blots

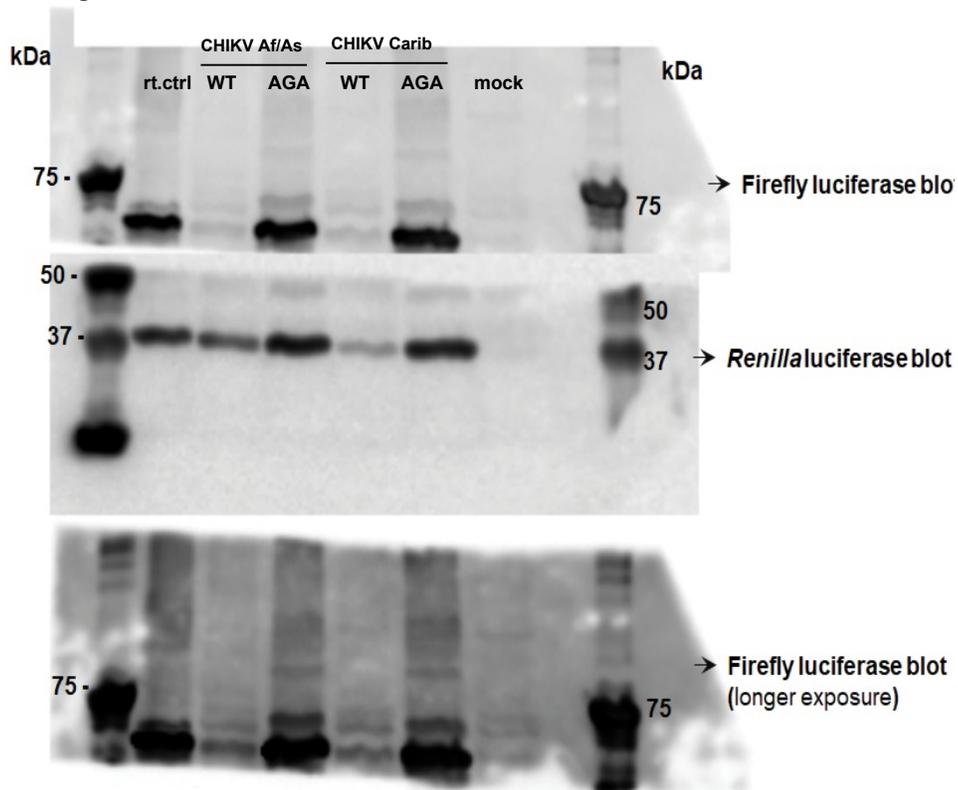


Figure 37. Panels A and B correspond to Figure 28C. Panels C and D correspond to Fig 28D. **A and C. Original Gels.** Total protein of samples separated through 4-20% SDS-PAGE gels. 10 μ g of readthrough control (rt.ctrl) sample and 15-20 μ g other samples were loaded for immunoblot analyses. **B and D: Original Blots.** Raw images of Firefly luciferase (short and long exposures) and *Renilla* luciferase immunoblots of protein lysates from HEK293T cells transiently transfected with the indicated reporter plasmids. The lanes on either end is molecular weight marker (Precision Plus Protein™ Kaleidoscope™ Bio-Rad #1610375).

Bibliography

1. Rappuoli, R. From Pasteur to genomics: progress and challenges in infectious diseases. *Nat. Med.* **10**, 1177–1185 (2004).
2. The top 10 causes of death. *World Health Organization* Available at: <http://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>. (Accessed: 14th August 2018)
3. WHO | Disease burden and mortality estimates. *WHO* Available at: http://www.who.int/healthinfo/global_burden_disease/estimates/en/. (Accessed: 14th August 2018)
4. Geoghegan, J. L., Duchêne, S. & Holmes, E. C. Comparative analysis estimates the relative frequencies of co-divergence and cross-species transmission within viral families. *PLOS Pathog.* **13**, e1006215 (2017).
5. Team, W. B. D. New country classifications by income level: 2017-2018. *The Data Blog* (2017). Available at: <https://blogs.worldbank.org/opendata/new-country-classifications-income-level-2017-2018>. (Accessed: 20th August 2018)
6. PringleJun. 4, H., 2015 & Pm, 4:00. How Europeans brought sickness to the New World. *Science / AAAS* (2015). Available at: <http://www.sciencemag.org/news/2015/06/how-europeans-brought-sickness-new-world>. (Accessed: 14th August 2018)
7. Patterson, K. B. & Runge, T. Smallpox and the Native American. *Am. J. Med. Sci.* **323**, 216–222 (2002).

8. Riley, J. C. Smallpox and American Indians revisited. *J. Hist. Med. Allied Sci.* **65**, 445–477 (2010).
9. Zappa, A., Amendola, A., Romanò, L. & Zanetti, A. Emerging and re-emerging viruses in the era of globalisation. *Blood Transfus.* **7**, 167–171 (2009).
10. Morens, D. M. & Fauci, A. S. The 1918 influenza pandemic: insights for the 21st century. *J. Infect. Dis.* **195**, 1018–1028 (2007).
11. WHO | Pandemic (H1N1) 2009. *WHO* Available at:
<http://www.who.int/csr/disease/swineflu/en/>. (Accessed: 15th August 2018)
12. Weiss, R. A. & McMichael, A. J. Social and environmental risk factors in the emergence of infectious diseases. *Nat. Med.* **10**, S70-76 (2004).
13. Gao, F. *et al.* Origin of HIV-1 in the chimpanzee *Pan troglodytes troglodytes*. *Nature* **397**, 436–441 (1999).
14. Sharp, P. M. *et al.* The origins of acquired immune deficiency syndrome viruses: where and when? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **356**, 867–876 (2001).
15. Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS Pandemic. *Cold Spring Harb. Perspect. Med.* **1**, (2011).
16. Interviews - Don Francis | The Age Of Aids | FRONTLINE | PBS. Available at:
<https://www.pbs.org/wgbh/pages/frontline/aids/interviews/francis.html>.
(Accessed: 15th August 2018)
17. Homepage. Available at: <http://www.unaids.org/en/Homepage>. (Accessed: 15th August 2018)

18. WHO | Data and statistics. *WHO* Available at: <http://www.who.int/hiv/data/en/>.
(Accessed: 15th August 2018)
19. Shi, Z. & Hu, Z. A review of studies on animal reservoirs of the SARS coronavirus.
Virus Res. **133**, 74–87 (2008).
20. Stadler, K. *et al.* SARS--beginning to understand a new virus. *Nat. Rev. Microbiol.* **1**,
209–218 (2003).
21. Centers for Disease Control and Prevention (CDC). Update: outbreak of severe acute
respiratory syndrome--worldwide, 2003. *MMWR Morb. Mortal. Wkly. Rep.* **52**,
269–272 (2003).
22. 2014-2016 Ebola Outbreak in West Africa | History | Ebola (Ebola Virus Disease) |
CDC. (2018). Available at: <https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/index.html>. (Accessed: 14th August 2018)
23. 2014 Ebola Outbreak in West Africa Epidemic Curves | 2014-2016 Outbreak West
Africa | History | Ebola (Ebola Virus Disease) | CDC. (2018). Available at:
<https://www.cdc.gov/vhf/ebola/history/2014-2016-outbreak/cumulative-cases-graphs.html>. (Accessed: 14th August 2018)
24. Shuman, E. K. Global climate change and infectious diseases. *N. Engl. J. Med.* **362**,
1061–1063 (2010).
25. Gould, E. A. & Higgs, S. Impact of climate change and other factors on emerging
arbovirus diseases. *Trans. R. Soc. Trop. Med. Hyg.* **103**, 109–121 (2009).
26. Patterson, J., Sammon, M. & Garg, M. Dengue, Zika and Chikungunya: Emerging
Arboviruses in the New World. *West. J. Emerg. Med.* **17**, 671–679 (2016).

27. Weaver, S. C. & Reisen, W. K. Present and Future Arboviral Threats. *Antiviral Res.* **85**, 328 (2010).
28. Khormi, H. M. & Kumar, L. Climate change and the potential global distribution of *Aedes aegypti*: spatial modelling using GIS and CLIMEX. *Geospatial Health* **8**, 405–415 (2014).
29. Caminade, C. *et al.* Global risk model for vector-borne transmission of Zika virus reveals the role of El Niño 2015. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 119–124 (2017).
30. Eisen, L. & Moore, C. G. *Aedes (Stegomyia) aegypti* in the continental United States: a vector at the cool margin of its geographic range. *J. Med. Entomol.* **50**, 467–478 (2013).
31. Pagès, F., Corbel, V. & Paupy, C. [Aedes albopictus: chronical of a spreading vector]. *Med. Trop. Rev. Corps Sante Colon.* **66**, 226–228 (2006).
32. Benedict, M. Q., Levine, R. S., Hawley, W. A. & Lounibos, L. P. Spread of the tiger: global risk of invasion by the mosquito *Aedes albopictus*. *Vector Borne Zoonotic Dis. Larchmt. N* **7**, 76–85 (2007).
33. Brugman, V. A. *et al.* The Role of *Culex pipiens* L. (Diptera: Culicidae) in Virus Transmission in Europe. *Int. J. Environ. Res. Public Health* **15**, (2018).
34. Baylis, M. Potential impact of climate change on emerging vector-borne and other infections in the UK. *Environ. Health* **16**, (2017).
35. Braack, L., Gouveia de Almeida, A. P., Cornel, A. J., Swanepoel, R. & de Jager, C. Mosquito-borne arboviruses of African origin: review of key viruses and vectors. *Parasit. Vectors* **11**, (2018).

36. Lee, S. H. *et al.* The Effects of Climate Change and Globalization on Mosquito Vectors: Evidence from Jeju Island, South Korea on the Potential for Asian Tiger Mosquito (*Aedes albopictus*) Influxes and Survival from Vietnam Rather Than Japan. *PLoS ONE* **8**, (2013).
37. Lee, H., Kim, J. E., Lee, S. & Lee, C. H. Potential effects of climate change on dengue transmission dynamics in Korea. *PLoS ONE* **13**, (2018).
38. Campbell, L. P. *et al.* Climate change influences on global distributions of dengue and chikungunya virus vectors. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).
39. Musso, D. & Gubler, D. J. Zika Virus. *Clin. Microbiol. Rev.* **29**, 487–524 (2016).
40. Iosifidis, S. *et al.* Current Zika virus epidemiology and recent epidemics. *Med. Mal. Infect.* **44**, 302–307 (2014).
41. Weaver, S. C. *et al.* Zika Virus: History, Emergence, Biology, and Prospects for Control. *Antiviral Res.* **130**, 69–80 (2016).
42. Dick, G. W. A., Kitchen, S. F. & Haddow, A. J. Zika virus. I. Isolations and serological specificity. *Trans. R. Soc. Trop. Med. Hyg.* **46**, 509–520 (1952).
43. White, M. K., Wollebo, H. S., David Beckham, J., Tyler, K. L. & Khalili, K. Zika virus: An emergent neuropathological agent. *Ann. Neurol.* **80**, 479–489 (2016).
44. Haddow, A. D. *et al.* Genetic characterization of Zika virus strains: geographic expansion of the Asian lineage. *PLoS Negl. Trop. Dis.* **6**, e1477 (2012).
45. WHO | One year into the Zika outbreak: how an obscure disease became a global health emergency. *WHO* Available at: <http://www.who.int/emergencies/zika-virus/articles/one-year-outbreak/en/>. (Accessed: 17th August 2018)

46. Zanoluca, C. *et al.* First report of autochthonous transmission of Zika virus in Brazil. *Mem. Inst. Oswaldo Cruz* **110**, 569–572 (2015).
47. Pinto Junior, V. L., Luz, K., Parreira, R. & Ferrinho, P. [Zika Virus: A Review to Clinicians]. *Acta Med. Port.* **28**, 760–765 (2015).
48. Krauer, F. *et al.* Zika Virus Infection as a Cause of Congenital Brain Abnormalities and Guillain–Barré Syndrome: Systematic Review. *PLoS Med.* **14**, (2017).
49. Lozano-Fuentes, S. *et al.* Susceptibility and Vectorial Capacity of American *Aedes albopictus* and *Aedes aegypti* (Diptera: Culicidae) to American Zika Virus Strains. *J. Med. Entomol.* (2018). doi:10.1093/jme/tjy114
50. Garoff, H., Sjöberg, M. & Cheng, R. H. Budding of alphaviruses. *Virus Res.* **106**, 103–116 (2004).
51. Suomalainen, M., Liljeström, P. & Garoff, H. Spike protein-nucleocapsid interactions drive the budding of alphaviruses. *J. Virol.* **66**, 4737–4747 (1992).
52. Sjöberg, M. & Garoff, H. Interactions between the transmembrane segments of the alphavirus E1 and E2 proteins play a role in virus budding and fusion. *J. Virol.* **77**, 3441–3450 (2003).
53. Li, L., Jose, J., Xiang, Y., Kuhn, R. J. & Rossmann, M. G. Structural changes of envelope proteins during alphavirus fusion. *Nature* **468**, 705–708 (2010).
54. Go, Y. Y., Balasuriya, U. B. R. & Lee, C. Zoonotic encephalitides caused by arboviruses: transmission and epidemiology of alphaviruses and flaviviruses. *Clin. Exp. Vaccine Res.* **3**, 58–77 (2014).

55. Weaver, S. C., Winegar, R., Manger, I. D. & Forrester, N. L. Alphaviruses: Population genetics and determinants of emergence. *Antiviral Res.* **94**, 242–257 (2012).
56. ViralZone. Available at: <https://viralzone.expasy.org/>. (Accessed: 20th August 2018)
57. Strauss, J. H. & Strauss, E. G. The alphaviruses: gene expression, replication, and evolution. *Microbiol. Rev.* **58**, 491–562 (1994).
58. Atkins, G. J. The Pathogenesis of Alphaviruses. *International Scholarly Research Notices* (2013). doi:10.5402/2013/861912
59. Meltzer, E. Arboviruses and viral hemorrhagic fevers (VHF). *Infect. Dis. Clin. North Am.* **26**, 479–496 (2012).
60. Rulli, N. E., Melton, J., Wilmes, A., Ewart, G. & Mahalingam, S. The molecular and cellular aspects of arthritis due to alphavirus infections: lesson learned from Ross River virus. *Ann. N. Y. Acad. Sci.* **1102**, 96–108 (2007).
61. Laine, M., Luukkainen, R. & Toivanen, A. Sindbis viruses and other alphaviruses as cause of human arthritic disease. *J. Intern. Med.* **256**, 457–471 (2004).
62. Toivanen, A. Alphaviruses: an emerging cause of arthritis? *Curr. Opin. Rheumatol.* **20**, 486–490 (2008).
63. Ross, R. W. The Newala epidemic. *J. Hyg. (Lond.)* **54**, 177–191 (1956).
64. Weaver, S. C. & Forrester, N. L. Chikungunya: Evolutionary history and recent epidemic spread. *Antiviral Res.* **120**, 32–39 (2015).
65. Amraoui, F. & Failloux, A.-B. Chikungunya: an unexpected emergence in Europe. *Curr. Opin. Virol.* **21**, 146–150 (2016).

66. Notes from the Field: Transmission of Chikungunya Virus in the Continental United States — Florida, 2014. Available at:
<https://www.cdc.gov/mmWr/preview/mmwrhtml/mm6348a4.htm>. (Accessed: 13th July 2018)
67. Geographic Distribution | Chikungunya virus | CDC. (2018). Available at:
<https://www.cdc.gov/chikungunya/geo/index.html>. (Accessed: 20th August 2018)
68. MacDonald, G. H. & Johnston, R. E. Role of Dendritic Cell Targeting in Venezuelan Equine Encephalitis Virus Pathogenesis. *J. Virol.* **74**, 914–922 (2000).
69. Gardner, C. L. *et al.* Eastern and Venezuelan Equine Encephalitis Viruses Differ in Their Ability To Infect Dendritic Cells and Macrophages: Impact of Altered Cell Tropism on Pathogenesis. *J. Virol.* **82**, 10634–10646 (2008).
70. Salimi, H., Cain, M. D. & Klein, R. S. Encephalitic Arboviruses: Emergence, Clinical Presentation, and Neuropathogenesis. *Neurotherapeutics* **13**, 514–534 (2016).
71. Zacks, M. A. & Paessler, S. Encephalitic alphaviruses. *Vet. Microbiol.* **140**, 281–286 (2010).
72. Steele, K. E. & Twenhafel, N. A. REVIEW PAPER: Pathology of Animal Models of Alphavirus Encephalitis. *Vet. Pathol.* **47**, 790–805 (2010).
73. Weaver, S. C., Powers, A. M., Brault, A. C. & Barrett, A. D. Molecular epidemiological studies of veterinary arboviral encephalitides. *Vet. J. Lond. Engl.* **1997** **157**, 123–138 (1999).

74. Deresiewicz, R. L., Thaler, S. J., Hsu, L. & Zamani, A. A. Clinical and neuroradiographic manifestations of eastern equine encephalitis. *N. Engl. J. Med.* **336**, 1867–1874 (1997).
75. Armstrong, P. M. & Andreadis, T. G. Eastern Equine Encephalitis Virus — Old Enemy, New Threat. <http://dx.doi.org/10.1056/NEJMp1213696> (2013).
doi:10.1056/NEJMp1213696
76. Confirmed and Probable Eastern Equine Encephalitis Cases, Human, United States, 1964-2003, By State. *N. Y.* 3
77. Mitchell, C. J. *et al.* Isolation of eastern equine encephalitis virus from *Aedes albopictus* in Florida. *Science* **257**, 526–527 (1992).
78. Reisen, W. K., Fang, Y. & Brault, A. C. Limited interdecadal variation in mosquito (Diptera: Culicidae) and avian host competence for Western equine encephalomyelitis virus (Togaviridae: Alphavirus). *Am. J. Trop. Med. Hyg.* **78**, 681–686 (2008).
79. Forrester, N. L., Kenney, J. L., Deardorff, E., Wang, E. & Weaver, S. C. Western equine encephalitis submergence: lack of evidence for a decline in virus virulence. *Virology* **380**, 170–172 (2008).
80. Reeves, W. C., Hammon, W. M., Longshore, W. A., McCLURE, H. E. & Geib, A. F. Epidemiology of the arthropod-borne viral encephalitides in Kern County, California, 1943-1952. *Publ. Public Health* **4**, 1–257 (1962).

81. Western Equine Encephalitis Fact Sheet - Minnesota Dept. of Health. Available at:
<http://www.health.state.mn.us/divs/idepc/diseases/weencephalitis/wee.html>.
(Accessed: 19th July 2018)
82. Neira, M. V., Mahmood, F., Reisen, W. K., James, C. B. L. & Romoser, W. S.
Comparative Study of the Pathological Effects of Western Equine
Encephalomyelitis Virus in Four Strains of *Culex tarsalis* Coquillett (Diptera:
Culicidae). *Front. Public Health* **2**, 184 (2014).
83. Reisen, W. K., Meyer, R. P., Presser, S. B. & Hardy, J. L. Effect of temperature on the
transmission of western equine encephalomyelitis and St. Louis encephalitis
viruses by *Culex tarsalis* (Diptera: Culicidae). *J. Med. Entomol.* **30**, 151–160 (1993).
84. Baer, A. *et al.* Venezuelan Equine Encephalitis Virus Induces Apoptosis through the
Unfolded Protein Response Activation of EGR1. *J. Virol.* **90**, 3558–3572 (2016).
85. Weaver, S. C., Ferro, C., Barrera, R., Boshell, J. & Navarro, J.-C. Venezuelan equine
encephalitis. *Annu. Rev. Entomol.* **49**, 141–174 (2004).
86. Weaver, S. C. *et al.* Re-emergence of epidemic Venezuelan equine encephalomyelitis
in South America. VEE Study Group. *Lancet Lond. Engl.* **348**, 436–440 (1996).
87. Venezuelan Equine Encephalitis Virus Activity in the Gulf Coast Region of Mexico,
2003–2010. Available at:
<http://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0001875>.
(Accessed: 20th July 2018)
88. Aguilar, P. V. *et al.* Endemic Venezuelan equine encephalitis in the Americas: hidden
under the dengue umbrella. *Future Virol.* **6**, 721–740 (2011).

89. Taylor, K. G. & Paessler, S. Pathogenesis of Venezuelan equine encephalitis. *Vet. Microbiol.* **167**, 145–150 (2013).
90. León, C. A. Sequelae of Venezuelan equine encephalitis in humans: a four year follow-up. *Int. J. Epidemiol.* **4**, 131–140 (1975).
91. Weaver, S. C. & Barrett, A. D. T. Transmission cycles, host range, evolution and emergence of arboviral disease. *Nat. Rev. Microbiol.* **2**, 789–801 (2004).
92. Reichert, E., Clase, A., Bacetty, A. & Larsen, J. Alphavirus antiviral drug development: scientific gap analysis and prospective research areas. *Biosecurity Bioterrorism Biodefense Strategy Pract. Sci.* **7**, 413–427 (2009).
93. Federal Select Agent Program - Select Agents and Toxins List. Available at: <https://www.selectagents.gov/selectagentsandtoxinslist.html>. (Accessed: 23rd July 2018)
94. Spurgers, K. B. & Glass, P. J. Vaccine Development for Biothreat Alpha viruses. *J. Bioterrorism Biodefense* **01**, (2011).
95. Bronze, M. S., Huycke, M. M., Machado, L. J., Voskuhl, G. W. & Greenfield, R. A. Viral agents as biological weapons and agents of bioterrorism. *Am. J. Med. Sci.* **323**, 316–325 (2002).
96. Hawley, R. J. & Eitzen, E. M. Biological weapons--a primer for microbiologists. *Annu. Rev. Microbiol.* **55**, 235–253 (2001).
97. Phelps, A. L. *et al.* Susceptibility and Lethality of Western Equine Encephalitis Virus in Balb/c Mice When Infected by the Aerosol Route. *Viruses* **9**, (2017).

98. Porter, A. I. *et al.* Characterization and pathogenesis of aerosolized eastern equine encephalitis in the common marmoset (*Callithrix jacchus*). *Virol. J.* **14**, 25 (2017).
99. Hunt, A. R. *et al.* Treatment of mice with human monoclonal antibody 24 hours after lethal aerosol challenge with virulent Venezuelan equine encephalitis virus prevents disease but not infection. *Virology* **414**, 146–152 (2011).
100. Ryzhikov, A. B., Ryabchikova, E. I., Sergeev, A. N. & Tkacheva, N. V. Spread of Venezuelan equine encephalitis virus in mice olfactory tract. *Arch. Virol.* **140**, 2243–2254 (1995).
101. Bernstein, B. J. The Birth of the U.S. Biological-Warfare Program. *Sci. Am.* **256**, 116–121 (1987).
102. Kortepeter, M. G., Cieslak, T. J. & Eitzen, E. M. Bioterrorism. *J. Environ. Health* **63**, 21–24 (2001).
103. Berge, T. O., Banks, I. S. & Tigertt, W. D. ATTENUATION OF VENEZUELAN EQUINE ENCEPHALOMYELITIS VIRUS BY IN VITRO CULTIVATION IN GUINEA-PIG HEART CELLS. *Am. J. Epidemiol.* **73**, 209–218 (1961).
104. Kulasegaran-Shylini, R., Thiviyathan, V., Gorenstein, D. G. & Frolov, I. The 5'UTR-specific mutation in VEEV TC-83 genome has a strong effect on RNA replication and subgenomic RNA synthesis, but not on translation of the encoded proteins. *Virology* **387**, 211–221 (2009).
105. 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. *J VIROL* **67**, 9 (1993).

106. Calisher, C. H. & Maness, K. S. Laboratory studies of Venezuelan equine encephalitis virus in equines, Texas, 1971. *J. Clin. Microbiol.* **2**, 198–205 (1975).
107. Ludwig, G. V. *et al.* Comparative neurovirulence of attenuated and non-attenuated strains of Venezuelan equine encephalitis virus in mice. *Am. J. Trop. Med. Hyg.* **64**, 49–55 (2001).
108. Jacks, T. & Varmus, H. E. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. *Science* **230**, 1237–1242 (1985).
109. Jacks, T., Madhani, H. D., Masiarz, F. R. & Varmus, H. E. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. *Cell* **55**, 447–458 (1988).
110. Jacks, T. *et al.* Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. *Nature* **331**, 280–283 (1988).
111. Nikolić, E. I. C., King, L. M., Vidakovic, M., Irigoyen, N. & Brierley, I. Modulation of Ribosomal Frameshifting Frequency and Its Effect on the Replication of Rous Sarcoma Virus. *J. Virol.* **86**, 11581–11594 (2012).
112. Kang, H. Direct structural evidence for formation of a stem-loop structure involved in ribosomal frameshifting in human immunodeficiency virus type 1. *Biochim. Biophys. Acta* **1397**, 73–78 (1998).
113. Giedroc, D. P., Theimer, C. A. & Nixon, P. L. Structure, stability and function of RNA pseudoknots involved in stimulating ribosomal frameshifting¹¹Edited by D. E. Draper. *J. Mol. Biol.* **298**, 167–185 (2000).
114. PLANT, E. P. *et al.* The 9-Å solution: How mRNA pseudoknots promote efficient programmed –1 ribosomal frameshifting. *RNA* **9**, 168–174 (2003).

115. Dinman, J. D. Mechanisms and implications of programmed translational frameshifting. *Wiley Interdiscip. Rev. RNA* **3**, 661–673 (2012).
116. Liao, P.-Y., Choi, Y. S., Dinman, J. D. & Lee, K. H. The many paths to frameshifting: kinetic modelling and analysis of the effects of different elongation steps on programmed –1 ribosomal frameshifting. *Nucleic Acids Res.* **39**, 300–312 (2011).
117. YOSHINAKA, Y., KATOH, I., COPELAND, T. D. & OROSZLAN, S. Translational Readthrough of an Amber Termination Codon During Synthesis of Feline Leukemia Virus Protease. 4
118. Yoshinaka, Y., Katoh, I., Copeland, T. D. & Oroszlan, S. Murine leukemia virus protease is encoded by the gag-pol gene and is synthesized through suppression of an amber termination codon. *Proc. Natl. Acad. Sci. U. S. A.* **82**, 1618–1622 (1985).
119. Panganiban, A. T. Retroviral gag gene amber codon suppression is caused by an intrinsic cis-acting component of the viral mRNA. *J. Virol.* **62**, 3574–3580 (1988).
120. Schueren, F. & Thoms, S. Functional Translational Readthrough: A Systems Biology Perspective. *PLoS Genet.* **12**, (2016).
121. Li, G. & Rice, C. M. The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon. *J. Virol.* **67**, 5062–5067 (1993).
122. Icho, T. & Wickner, R. B. The double-stranded RNA genome of yeast virus L-A encodes its own putative RNA polymerase by fusing two open reading frames. *J. Biol. Chem.* **264**, 6716–6723 (1989).

123. Fujimura, T., Ribas, J. C., Makhov, A. M. & Wickner, R. B. Pol of gag-pol fusion protein required for encapsidation of viral RNA of yeast L-A virus. *Nature* **359**, 746–749 (1992).
124. Dinman, J. D. & Wickner, R. B. Ribosomal frameshifting efficiency and gag/gag-pol ratio are critical for yeast M1 double-stranded RNA virus propagation. *J. Virol.* **66**, 3669–3676 (1992).
125. Hung, M., Patel, P., Davis, S. & Green, S. R. Importance of Ribosomal Frameshifting for Human Immunodeficiency Virus Type 1 Particle Assembly and Replication. *J. Virol.* **72**, 4819–4824 (1998).
126. Baranov, P. V. *et al.* Programmed ribosomal frameshifting in decoding the SARS-CoV genome. *Virology* **332**, 498–510 (2005).
127. Brierley, I. & Dos Ramos, F. J. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. *Virus Res.* **119**, 29–42 (2006).
128. Plant, E. P., Rakauskaitė, R., Taylor, D. R. & Dinman, J. D. Achieving a Golden Mean: Mechanisms by Which Coronaviruses Ensure Synthesis of the Correct Stoichiometric Ratios of Viral Proteins. *J. Virol.* **84**, 4330–4340 (2010).
129. Plant, E. P., Sims, A. C., Baric, R. S., Dinman, J. D. & Taylor, D. R. Altering SARS Coronavirus Frameshift Efficiency Affects Genomic and Subgenomic RNA Production. *Viruses* **5**, 279–294 (2013).
130. Plant, E. P. *et al.* A three-stemmed mRNA pseudoknot in the SARS coronavirus frameshift signal. *PLoS Biol.* **3**, e172 (2005).

131. Kim, H.-K. *et al.* A frameshifting stimulatory stem loop destabilizes the hybrid state and impedes ribosomal translocation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 5538–5543 (2014).
132. Ritchie, D. B. *et al.* Conformational dynamics of the frameshift stimulatory structure in HIV-1. *RNA N. Y. N* **23**, 1376–1384 (2017).
133. Su, M.-C., Chang, C.-T., Chu, C.-H., Tsai, C.-H. & Chang, K.-Y. An atypical RNA pseudoknot stimulator and an upstream attenuation signal for -1 ribosomal frameshifting of SARS coronavirus. *Nucleic Acids Res.* **33**, 4265–4275 (2005).
134. Gao, F. & Simon, A. E. Differential use of 3’CITEs by the subgenomic RNA of Pea enation mosaic virus 2. *Virology* **510**, 194–204 (2017).
135. Kuhlmann, M. M., Chattopadhyay, M., Stupina, V. A., Gao, F. & Simon, A. E. An RNA Element That Facilitates Programmed Ribosomal Readthrough in Turnip Crinkle Virus Adopts Multiple Conformations. *J. Virol.* **90**, 8575–8591 (2016).
136. Le, M.-T., Kasprzak, W. K., Shapiro, B. A. & Simon, A. E. Combined single molecule experimental and computational approaches for understanding the unfolding pathway of a viral translation enhancer that participates in a conformational switch. *RNA Biol.* **14**, 1466–1472 (2017).
137. Paul, C. P., Barry, J. K., Dinesh-Kumar, S. P., Brault, V. & Miller, W. A. A sequence required for -1 ribosomal frameshifting located four kilobases downstream of the frameshift site. *J. Mol. Biol.* **310**, 987–999 (2001).

138. Strauss, E. G., Rice, C. M. & Strauss, J. H. Sequence coding for the alphavirus nonstructural proteins is interrupted by an opal termination codon. *Proc. Natl. Acad. Sci. U. S. A.* **80**, 5271–5275 (1983).
139. Firth, A. E., Wills, N. M., Gesteland, R. F. & Atkins, J. F. Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. *Nucleic Acids Res.* **39**, 6679–6691 (2011).
140. Jones, J. E. *et al.* Disruption of the Opal Stop Codon Attenuates Chikungunya Virus-Induced Arthritis and Pathology. *mBio* **8**, (2017).
141. Myles, K. M., Kelly, C. L. H., Ledermann, J. P. & Powers, A. M. Effects of an Opal Termination Codon Preceding the nsP4 Gene Sequence in the O'Nyong-Nyong Virus Genome on *Anopheles gambiae* Infectivity. *J. Virol.* **80**, 4992–4997 (2006).
142. Melton, J. V. *et al.* Alphavirus 6K Proteins Form Ion Channels. *J. Biol. Chem.* **277**, 46923–46931 (2002).
143. Firth, A. E., Chung, B. Y., Fleeton, M. N. & Atkins, J. F. Discovery of frameshifting in Alphavirus 6K resolves a 20-year enigma. *Viol. J.* **5**, 108 (2008).
144. Ramsey, J., Renzi, E. C., Arnold, R. J., Trinidad, J. C. & Mukhopadhyay, S. Palmitoylation of Sindbis Virus TF Protein Regulates Its Plasma Membrane Localization and Subsequent Incorporation into Virions. *J. Virol.* **91**, (2017).
145. Gaedigk-Nitschko, K. & Schlesinger, M. J. Site-directed mutations in Sindbis virus E2 glycoprotein's cytoplasmic domain and the 6K protein lead to similar defects in virus assembly and budding. *Virology* **183**, 206–214 (1991).

146. Gaedigk-Nitschko, K., Ding, M. X., Levy, M. A. & Schlesinger, M. J. Site-directed mutations in the Sindbis virus 6K protein reveal sites for fatty acylation and the underacylated protein affects virus release and virion structure. *Virology* **175**, 282–291 (1990).
147. Snyder, J. E. *et al.* Functional Characterization of the Alphavirus TF Protein. *J. Virol.* **87**, 8511–8523 (2013).
148. Ramsey, J. & Mukhopadhyay, S. Disentangling the Frames, the State of Research on the Alphavirus 6K and TF Proteins. *Viruses* **9**, (2017).
149. Chung, B. Y.-W., Firth, A. E. & Atkins, J. F. Frameshifting in alphaviruses: a diversity of 3' stimulatory structures. *J. Mol. Biol.* **397**, 448–456 (2010).
150. Kendra, J. A. *et al.* Ablation of Programmed –1 Ribosomal Frameshifting in Venezuelan Equine Encephalitis Virus Results in Attenuated Neuropathogenicity. *J. Virol.* **91**, (2017).
151. Leblanc, J., Weil, J. & Beemon, K. Posttranscriptional regulation of retroviral gene expression: primary RNA transcripts play three roles as pre-mRNA, mRNA, and genomic RNA. *Wiley Interdiscip. Rev. RNA* **4**, 567–80 (2013).
152. Sanchez, A., Trappier, S. G., Mahy, B. W., Peters, C. J. & Nichol, S. T. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 3602–7 (1996).
153. Firth, A. E. & Brierley, I. Non-canonical translation in RNA viruses. *J. Gen. Virol.* **93**, 1385–409 (2012).

154. Dinman, J. D. Programmed Ribosomal Frameshifting Goes Beyond Viruses: Organisms from all three kingdoms use frameshifting to regulate gene expression, perhaps signaling a paradigm shift. *Microbe* **1**, 521–527 (2006).
155. Moomau, C., Musalgaonkar, S., Khan, Y. A., Jones, J. E. & Dinman, J. D. Structural and Functional Characterization of Programmed Ribosomal Frameshift Signals in West Nile Virus Strains Reveals High Structural Plasticity Among cis-Acting RNA Elements. *J. Biol. Chem.* **291**, 15788–15795 (2016).
156. Belew, A. T., Hepler, N. L., Jacobs, J. L. & Dinman, J. D. PRFdb: A database of computationally predicted eukaryotic programmed -1 ribosomal frameshift signals. *BMC Genomics* **9**, 339 (2008).
157. Zadeh, J. N. *et al.* NUPACK: Analysis and design of nucleic acid systems. *J. Comput. Chem.* **32**, 170–173 (2011).
158. Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms Mol. Biol. AMB* **6**, 26 (2011).
159. Grentzmann, G., Ingram, J. A., Kelly, P. J., Gesteland, R. F. & Atkins, J. F. A dual-luciferase reporter system for studying recoding signals. *RNA* **4**, 479–486 (1998).
160. Schultz, K. L. W., Vernon, P. S. & Griffin, D. E. Differentiation of neurons restricts Arbovirus replication and increases expression of the alpha isoform of IRF-7. *J. Virol.* **89**, 48–60 (2015).
161. Belew, A. T. *et al.* Ribosomal frameshifting in the CCR5 mRNA is regulated by miRNAs and the NMD pathway. *Nature* **512**, 265–269 (2014).

162. Harger, J. W. & Dinman, J. D. An in vivo dual-luciferase assay system for studying translational recoding in the yeast *Saccharomyces cerevisiae*. *RNA* **9**, 1019–1024 (2003).
163. Merino, E. J., Wilkinson, K. A., Coughlan, J. L. & Weeks, K. M. RNA Structure Analysis at Single Nucleotide Resolution by Selective 2'-Hydroxyl Acylation and Primer Extension (SHAPE). *J. Am. Chem. Soc.* **127**, 4223–4231 (2005).
164. Wilkinson, K. A., Merino, E. J. & Weeks, K. M. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.* **1**, 1610–1616 (2006).
165. Parisien, M. & Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **452**, 51–55 (2008).
166. Johnson, B. J. B., Kinney, R. M., Kost, C. L. & Trent, D. W. Molecular Determinants of Alphavirus Neurovirulence: Nucleotide and Deduced Protein Sequence Changes during Attenuation of Venezuelan Equine Encephalitis Virus. *J. Gen. Virol.* **67**, 1951–1960 (1986).
167. Bernard, K. A., Klimstra, W. B. & Johnston, R. E. Mutations in the E2 glycoprotein of Venezuelan equine encephalitis virus confer heparan sulfate interaction, low morbidity, and rapid clearance from blood of mice. *Virology* **276**, 93–103 (2000).
168. Davis, N. L., Willis, V. L., Smith, J. F. & Johnston, R. E. In vitro synthesis of infectious venezuelan equine encephalitis virus RNA from a cDNA clone: Analysis of a viable deletion mutant. *Virology* **171**, 189–204 (1989).

169. 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 Res.* **78**, 230–41 (2008).
170. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{(-Delta Delta C(T))} Method. *Methods* **25**, 402–408 (2001).
171. *Guide for the Care and Use of Laboratory Animals.*
172. Jacobs, J. L. & Dinman, J. D. Systematic analysis of bicistronic reporter assay data. *Nucleic Acids Res.* **32**, e160 (2004).
173. Dunkle, J. A. & Dunham, C. M. Mechanisms of mRNA frame maintenance and its subversion during translation of the genetic code. *Biochimie* **114**, 90–6 (2015).
174. Mortimer, S. A. & Weeks, K. M. A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. *J.Am.Chem.Soc.* **129**, 4144–4145 (2007).
175. Brierley, I. Ribosomal frameshifting on viral RNAs. *J.Gen.Virol.* **76**, 1885–1892 (1995).
176. Melian, E. B. *et al.* Programmed ribosomal frameshift alters expression of west Nile virus genes and facilitates virus replication in birds and mosquitoes. *PLoS Pathog.* **10**, e1004447 (2014).
177. Melian, E. B. *et al.* NS1' of flaviviruses in the Japanese encephalitis virus serogroup is a product of ribosomal frameshifting and plays a role in viral neuroinvasiveness. *J.Virol.* **84**, 1641–1647 (2010).

178. Taylor, A. *et al.* Effects of an In-Frame Deletion of the
Gene Locus from the Genome of
Ross River Virus. *J. Virol.* **90**, 4150–4159 (2016).
179. Atkins, J. F., Loughran, G., Bhatt, P. R., Firth, A. E. & Baranov, V. P. Ribosomal
frameshifting and transcriptional slippage: From genetic steganography and
cryptography to adventitious use. *Nucleic Acids Res.* **44**, 7007–7078 (2016).
180. Staple, D. W. & Butcher, S. E. Solution structure of the HIV-1 frameshift inducing
stem-loop RNA. *Nucleic Acids Res* **31**, 4326–4331 (2003).
181. Staple, D. W. & Butcher, S. E. Solution structure and thermodynamic investigation
of the HIV-1 frameshift inducing element. *J.Mol.Biol.* **349**, 1011–1023 (2005).
182. Garcia-Miranda, P. *et al.* Stability of HIV Frameshift Site RNA Correlates with
Frameshift Efficiency and Decreased Virus Infectivity. *J. Virol.* **90**, 6906–6917
(2016).
183. Jacks, T. Translational suppression in gene expression in retroviruses and
retrotransposons. *Curr.Top.Microbiol.Immunol.* **157**, 93–124 (1990).
184. Sheju-Shilaga, M., Crowe, S. M. & Mak, J. Maintenance of the Gag/Gag-Pol ratio is
important for human immunodeficiency virus type 1 RNA dimerization and viral
infectivity. *J.Virol.* **75**, 1834–1841 (2001).
185. Dinman, J. D., Ruiz-Echevarria, M. J. & Peltz, S. W. Translating old drugs into new
treatments: Identifying compounds that modulate programmed -1 ribosomal
frameshifting and function as potential antiviral agents. *Trends Biotech* **16**, 190–
196 (1998).

186. Barry, J. K. & Miller, W. A. A -1 ribosomal frameshift element that requires base pairing across four kilobases suggests a mechanism of regulating ribosome and replicase traffic on a viral RNA. *Proc.Natl.Acad.Sci.U.S.A* **99**, 11133–11138 (2002).
187. Plant, E. P. & Dinman, J. D. The role of programmed-1 ribosomal frameshifting in coronavirus propagation. *Front Biosci* **13**, 4873–4881 (2008).
188. Young, L. B., Melian, E. B. & Khromykh, A. A. NS1' colocalizes with NS1 and can substitute for NS1 in West Nile virus replication. *J. Virol.* **87**, 9384–90 (2013).
189. Winkelmann, E. R., Widman, D. G., Suzuki, R. & Mason, P. W. Analyses of mutations selected by passaging a chimeric flavivirus identify mutations that alter infectivity and reveal an interaction between the structural proteins and the nonstructural glycoprotein NS1. *Virology* **421**, 96–104 (2011).
190. Advani, V. M. & Dinman, J. D. Reprogramming the genetic code: The emerging role of ribosomal frameshifting in regulating cellular gene expression. *BioEssays* **38**, 21–26 (2016).
191. Mukhopadhyay, S. *et al.* Mapping the structure and function of the E1 and E2 glycoproteins in alphaviruses. *Struct. Lond. Engl. 1993* **14**, 63–73 (2006).
192. Berge, W. D., T. O. .. Banks, I. S. .. Tigertt. Attenuation of Venezuelan equine encephalomyelitis virus by in vitro cultivation in guinea pig heart cells. *Am J Hyg* **73**, 209–218 (1961).
193. Pederson C.E., F. E., Robinson, D. M. .. Cole. Isolation of the vaccine strain of Venezuelan equine encephalomyelitis virus from mosquitoes in Louisiana. *Am J Epidemiol* **95**, 490–496 (1972).

194. Pittman, P. R. *et al.* Long-term duration of detectable neutralizing antibodies after administration of live-attenuated VEE vaccine and following booster vaccination with inactivated VEE vaccine. *Vaccine* **14**, 337–343 (1996).
195. Paessler, S. & Weaver, S. C. Vaccines for Venezuelan equine encephalitis. *Vaccine* **27 Suppl 4**, D80–5 (2009).
196. Kariuki Njenga, M. *et al.* Tracking epidemic Chikungunya virus into the Indian Ocean from East Africa. *J. Gen. Virol.* **89**, 2754–2760 (2008).
197. Yactayo, S., Staples, J. E., Millot, V., Cibrelus, L. & Ramon-Pardo, P. Epidemiology of Chikungunya in the Americas. *J. Infect. Dis.* **214**, S441–S445 (2016).
198. Brizzi, K. Neurologic Manifestation of Chikungunya Virus. *Curr. Infect. Dis. Rep.* **19**, 6 (2017).
199. Chandak, N. H. *et al.* Neurological complications of Chikungunya virus infection. *Neurol. India* **57**, 177 (2009).
200. Bandeira, A. C., Campos, G. S., Sardi, S. I., Rocha, V. F. D. & Rocha, G. C. M. Neonatal encephalitis due to Chikungunya vertical transmission: First report in Brazil. *IDCases* **5**, 57–59 (2016).
201. Cardona-Correa, S. E., Castaño-Jaramillo, L. M. & Quevedo-Vélez, A. [Vertical transmission of chikungunya virus infection. Case Report]. *Rev. Chil. Pediatr.* **88**, 285–288 (2017).
202. Li, G. P. & Rice, C. M. Mutagenesis of the in-frame opal termination codon preceding nsP4 of Sindbis virus: studies of translational readthrough and its effect on virus replication. *J. Virol.* **63**, 1326–1337 (1989).

203. Loughran, G., Howard, M. T., Firth, A. E. & Atkins, J. F. Avoidance of reporter assay distortions from fused dual reporters. *RNA* **23**, 1285–1289 (2017).
204. Gibson, D. G. *et al.* Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–5 (2009).
205. Merino, E. J., Wilkinson, K. A., Coughlan, J. L. & Weeks, K. M. RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* **127**, 4223–4231 (2005).
206. Smola, M. J., Rice, G. M., Busan, S., Siegfried, N. A. & Weeks, K. M. Selective 2'-hydroxyl acylation analyzed by primer extension and mutational profiling (SHAPE-MaP) for direct, versatile, and accurate RNA structure analysis. *Nat. Protoc.* **10**, 1643–1669 (2015).
207. Li, W. *et al.* The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* **43**, W580–W584 (2015).
208. Gorchakov, R. *et al.* Attenuation of Chikungunya virus vaccine strain 181/clone 25 is determined by two amino acid substitutions in the E2 envelope glycoprotein. *J. Virol.* **86**, 6084–96 (2012).
209. Sahadeo, N. *et al.* Correction: Molecular Characterisation of Chikungunya Virus Infections in Trinidad and Comparison of Clinical and Laboratory Features with Dengue and Other Acute Febrile Cases. *PLoS Negl. Trop. Dis.* **9**, e0004305 (2015).
210. Somogyi, P., Jenner, A. J., Brierley, I. A. & Inglis, S. C. Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell. Biol.* **13**, 6931–6940 (1993).

211. Lopinski, J. D., Dinman, J. D. & Bruenn, J. A. Kinetics of Ribosomal Pausing during Programmed -1 Translational Frameshifting. *MolCell Biol* **20**, 1095–1103 (2000).
212. Doma, M. K. & Parker, R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* **440**, 561–564 (2006).
213. Naphine, S., Vidakovic, M., Girnary, R., Namy, O. & Brierley, I. Prokaryotic-style frameshifting in a plant translation system: conservation of an unusual single-tRNA slippage event. *EMBO J* **22**, 3941–3950 (2003).
214. Schuller, A. P. & Green, R. Roadblocks and resolutions in eukaryotic translation. *Nat. Rev. Mol. Cell Biol.* **19**, 526–541 (2018).
215. Plant, E. P. & Dinman, J. D. Torsional restraint: A new twist on frameshifting pseudoknots. *Nucleic Acids Res.* **33**, 1825–1833 (2005).
216. Tajima, Y., Iwakawa, H., Kaido, M., Mise, K. & Okuno, T. A long-distance RNA–RNA interaction plays an important role in programmed – 1 ribosomal frameshifting in the translation of p88 replicase protein of Red clover necrotic mosaic virus. *Virology* **417**, 169–178 (2011).
217. Gao, F. & Simon, A. E. Multiple Cis-acting elements modulate programmed -1 ribosomal frameshifting in Pea enation mosaic virus. *Nucleic Acids Res.* **44**, 878–895 (2016).
218. Newburn, L. R., Nicholson, B. L., Yosefi, M., Cimino, P. A. & White, K. A. Translational readthrough in Tobacco necrosis virus-D. *Virology* **450–451**, 258–265 (2014).

219. Nicholson, B. L. & White, K. A. Functional long-range RNA–RNA interactions in positive-strand RNA viruses. *Nat. Rev. Microbiol.* **12**, 493–504 (2014).
220. Newburn, L. R. & White, K. A. Cis-acting RNA elements in positive-strand RNA plant virus genomes. *Virology* **206**, 37–45 (2015).
221. Goff, S. P. & Green, L. Translational readthrough-promoting drugs enhance pseudoknot-mediated suppression of the stop codon at the Moloney murine leukemia virus gag–pol junction. *J. Gen. Virol.* **96**, 3411–3421 (2015).
222. Csibra, E., Brierley, I. & Irigoyen, N. Modulation of stop codon read-through efficiency and its effect on the replication of murine leukemia virus. *J. Virol.* **88**, 10364–76 (2014).
223. Adamson, C. S. *et al.* A block in virus-like particle maturation following assembly of murine leukaemia virus in insect cells. *Virology* **314**, 488–496 (2003).
224. Lemm, J. A., Durbin, R. K., Stollar, V. & Rice, C. M. Mutations which alter the level or structure of nsP4 can affect the efficiency of Sindbis virus replication in a host-dependent manner. *J. Virol.* **64**, 3001–11 (1990).
225. Siegfried, N. A., Busan, S., Rice, G. M., Nelson, J. A. E. & Weeks, K. M. RNA motif discovery by SHAPE and mutational profiling (SHAPE-MaP). *Nat. Methods* **11**, 959–965 (2014).



UNIVERSITY OF
MARYLAND

BIOLOGICAL SCIENCES GRADUATE PROGRAM

2101 Bioscience Research Building
College Park, Maryland 20742-4415
301.405.6905/6991 TEL, 301.314.9921 FAX

Steve Fetter, Ph.D.
Associate Dean of the Graduate School
University of Maryland

Dear Dr. Fetter,

This letter is written to signify that the dissertation committee, committee chair, and the graduate director have all approved the use of previously published co-authored work in the final dissertation of JOSEPH KENDRA, Biological Sciences Graduate Program, 112530433. In accordance with the Graduate School's policy the dissertation committee has determined that they made substantial contributions to the included work.

The citations for the published works are:

Ablation of Programmed -1 Ribosomal Frameshifting in Venezuelan Equine Encephalitis Virus Results in Attenuated Neuropathogenicity. Joseph A. Kendra, Cynthia de la Fuente, Ashwini Brahms, Caitlin Woodson, Todd M. Bell, Bin Chen, Yousuf A. Khan, Jonathan L. Jacobs, Kylene Kehn-Hall, Jonathan D. Dinman. *Journal of Virology* Jan 2017, 91 (3) e01766-16; **DOI:** 10.1128/JVI.01766-16

Functional and structural characterization of the Chikungunya virus translational recoding signals. Joseph A. Kendra, Vivek M. Advani. Bin Chen, Joseph W. Briggs, Jinyi Zhu, Hannah J. Bress, Sushrut M. Pathy, Jonathan D. Dinman. *Journal of Biological Chemistry* Sept 2018 (*In Press*)

Per Graduate School policy the dissertation forward will identify the scope and nature of the student's contributions to the jointly authored work included in the dissertation and a copy of this letter will be submitted with the dissertation.

Sincerely,

A handwritten signature in black ink, appearing to read "JD Dinman".

Dr. Jonathan D. Dinman, Dissertation Committee Chair, Professor and Chair, Cell Biology and Molecular Genetics

A handwritten signature in blue ink, appearing to read "Zakiya Whatley".

Dr. Zakiya Whatley, Program Manager, Biological Sciences Graduate Program

Joseph Kendra, Graduate Student, Biological Sciences

A handwritten signature in black ink, appearing to read "Joseph Kendra".