

Copyright

by

Brian Richard Mann

2014

The Dissertation Committee for Brian Richard Mann

Certifies that this is the approved version of the following dissertation:

**GENOTYPIC AND MOLECULAR DETERMINANTS
IN THE CONTINUED EVOLUTION OF WEST NILE VIRUS
IN THE UNITED STATES**

Committee:

Alan D.T. Barrett, D.V.M., Ph.D., Mentor

David W.C. Beasley, Ph.D.

Kristen A. Bernard, M.S., D.V.M., Ph.D.

Gregg N. Milligan, M.S., Ph.D.

Tian Wang, Ph.D.

Dean, Graduate School

**GENOTYPIC AND MOLECULAR DETERMINANTS
IN THE CONTINUED EVOLUTION OF WEST NILE VIRUS
IN THE UNITED STATES**

by

Brian Richard Mann, B.S.

Dissertation

Presented to the Faculty of the Graduate School of
The University of Texas Medical Branch
in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

**The University of Texas Medical Branch
December, 2014**

Dedication

To my Mom and Dad, Cindy and William Mann
...and all who came before, and all who are starting their academic journey.

Acknowledgements

First and foremost, I would like to thank my mentor, Dr. Alan D.T. Barrett, for his guidance, advice, patience, and support throughout my tenure at the University of Texas Medical Branch (UTMB). Despite his hectic schedule and exhaustive institutional commitments and responsibilities, he made time to discuss the results of my research, answer emails, review manuscripts, and listen to presentations. His guidance both inspired and nurtured my growth as an academic, and I am thankful for his encouragement and criticisms that have shaped and refined a scientific curiosity, rigor, and integrity befitting an honored recipient of a Doctorate of Philosophy (PhD).

I am also grateful to the members of my dissertation committee: Drs. David Beasley, Kristen Bernard, Gregg Milligan, and Tian Wang for taking time out of their individual schedules to oversee and direct the course of this research.

In addition, several members of the Barrett lab (both past and present) deserve particular acknowledgement, including Andrew Beck, Natalie Collins, Li Li, Dr. Allison McMullen, Dr. Trevor Pitcher, Dr. Vanessa Sarathy, Dr. Amy Schuh, Daniele Swetnam, and Dr. Melissa Whiteman. These individuals offered helpful advice regarding this research, and provided an enjoyable and collegiate working environment.

I would also like to extend a special thanks to several individuals who have been essential in the completion of this dissertation research. Allison and Li trained me in tissue culture, molecular virology, general lab protocols, and BSL-3 lab procedures and techniques. Melissa, Dr. Jessica Plante, and Dr. Beasley trained me in all animal work; in

addition, Melissa helped me complete several of the mouse studies described in this dissertation research. Furthermore, Drs. Steve Widen and Thomas Wood (UTMB Next Generation Sequencing Core) and Andrew Beck offered critical advice in the rational design and interpretation of all *in silico* experiments and data management. Amy provided additional helpful insight in the development of the statistical skills needed to complete this dissertation research. Melissa, Dr. C.T. Davis, and Dr. Jason Wicker also generated the infectious clones applied in this dissertation. Dr. Robert Tesh, Hilda Guzman, and Dr. David Romero provided me with the numerous West Nile virus isolates utilized throughout my research projects; furthermore, Drs. Rudy Bueno, Jr. and James Dennett with the help of Martin Reyna and Vence Salvato from the Harris County Public Health and Environmental Services: Mosquito Control Division and the Arbovirus Team at the Texas Department of State Health Services conducted the collection of all field isolates applied throughout this dissertation. Dr. Wang and Guorui Xie were also instrumental in the design, execution, and interpretation of all described *in vitro* gene and protein expression studies.

I would also like to thank my UTMB peers and professional colleagues—David Briley, Matthew Huante, and Sergio Rodriguez in particular—for their personal support on this long academic road. Finally, I am forever grateful and thankful for my parents, Bill and Cindy Mann, for their limitless support, love, and encouragement in pursuit of an academic career that never seems to end.

All funding for this research was provided in part by National Institutes of Health (NIH) grant AI 067847 to A.D.T.B, contract HHSN272201000040I/HHSN27200004/D04 to both A.D.T.B and R.B.T and the NIH T32 training grant AI 007526 from the National Institute of Allergy and Infectious Diseases.

**GENOTYPIC AND MOLECULAR DETERMINANTS
IN THE CONTINUED EVOLUTION OF WEST NILE VIRUS
IN THE UNITED STATES**

Publication No. _____

Brian Richard Mann, Ph.D.

The University of Texas Medical Branch, 2014

Supervisor: Alan D.T. Barrett

Following the introduction of West Nile virus (WNV) into New York in 1999, the virus continues to evolve in the U.S. with the emergence, extinction, and co-circulation of multiple, distinct U.S. genotypes. Despite annual local, regional, or national epidemics (varying in size by year), limited progress has been made dissecting how and when viral outbreaks occur; therefore, the objective of this dissertation was to investigate both genotypic and molecular determinants involved in continued evolution of WNV by relating *in silico* and *in vitro* differences in the molecular epidemiology, population dynamics, and phenotypic properties of both natural and infectious clone-derived viruses. Phylogenetic analysis revealed bi-directional virus transmission (*i.e.*, gene-flow) on the U.S.-Mexican border despite limited prior isolation or evidence of WNV circulation in northern Mexico; in addition, paired geospatial and phylogenetic comparisons identified the emergence of four distinct genetic clusters (Groups 7-10) of novel 2010-2012 WNV isolates in both Houston (Harris County) and Dallas/Fort Worth, Texas concurrent with the regional 2012 Texas WNV epidemic—implicating recent emergence and circulation of virus isolates with novel genotypic signatures. *In silico* application of 2002-2012

Houston, Texas isolates implicated the Harris Co. paradigm as a novel, surrogate model for national WNV surveillance and evolution in the U.S. Incorporation of nineteen 2002-2012 Harris Co. isolates from this paradigm in the Illumina HiSeq1000 next generation sequencing (also known as “deep” sequencing) platform identified both host- and genotypic-dependent trends in the diversity and evolution of regional WNV populations; additional, paired phenotypic studies revealed a relationship between an attenuated neuroinvasive mouse phenotype and low genetic diversity in natural WNV quasispecies structure. These *in silico* results and prior applications of the Harris Co. paradigm identified several natural mutations encoded in the NS4B protein—a protein that has been linked to interferon (IFN)- α/β antagonism. *In vitro* implementation of engineered NY99 infectious clone (NY99ic)-derived viruses—encoding natural or *in vitro*-selected NS4B substitutions—in the IFN-competent human A549 *in vitro* cell line revealed both residue- and domain-dependent dysregulation of host innate antiviral responses. Overall, the studies in this dissertation support continued evolution of WNV with evidence of genetic and phenotypic changes over time.

TABLE OF CONTENTS

List of Tables	xiv
List of Figures	xvii
List of Abbreviations	23
CHAPTER 1.....	28
Introduction.....	28
1.1 Molecular epidemiology and phylogenetics	28
1.1.1 West Nile virus classification	28
1.1.2 Phylogenetics: defining a lineage	29
1.1.3 Introduction of an “Old World” virus into the “New World”	30
1.1.4 North American (U.S.) WNV genotypes	31
1.1.4.1 NY99 genotype	32
1.1.4.2 Southeast (SE) coastal Texas genotype	32
1.1.4.3 North American/WN 2002 (NA/WN02) genotype.....	32
1.1.4.4 Southwestern/SW 2003 (SW/WN03) genotype.....	33
1.1.4.5 Rise and fall of U.S. genotypes: underlying factors	33
1.2 Ecology and Natural Transmission.....	34
1.2.1 Principal vector: mosquitoes.....	34
1.2.2 Principal host: birds	35
1.2.3 Trade-Off Hypothesis: WNV host/vector fitness.....	36
1.2.4 Implications in human health: epidemiology in the U.S.....	37
1.3 Molecular virology	38
1.3.1 Structure and genomic organization	38
1.3.1.1 Structural proteins.....	38
1.3.1.2 Nonstructural proteins.....	39
1.3.1.3 Untranslated regions: 5’/3’ terminal sequences	42
1.3.2 Viral life cycle.....	42
1.3.3 Viral replication	43
1.4 Pathogenesis.....	44
1.4.1 Human infection.....	44

1.4.2 Clinical diagnoses	45
1.4.3 Innate immune responses	47
1.5 Gaps in knowledge.....	51
1.6 Specific Aims.....	52
1.6.1 Specific aim 1(a): Chapter 3	53
1.6.2 Specific aim 1(b): Chapter 4	54
1.6.3 Specific aim 2: Chapter 5.....	54
1.6.4 Specific aim 3: Chapter 6.....	55
CHAPTER 2.....	65
Materials & Methods	65
2.1 Buffers, media, and reagents.....	65
2.1.1 Buffers.....	65
2.1.1.1 Buffer RLT (Qiagen RNeasy Mini kit).....	65
2.1.2 Media	65
2.1.2.1 A549 and Vero cell culture growth media.....	65
2.1.2.2 C6/36 cell culture growth media.....	65
2.1.2.3 2x Minimal essential media (MEM)	66
2.1.2.4 1x Minimal essential media (MEM) agar	66
2.1.3 Reagents	66
2.1.3.1 Tryptose phosphate buffer	66
2.2 Cell culture methods	66
2.2.1 Cell-lines	66
2.2.2 Cell culture protocols	67
2.3 West Nile virus inoculation and <i>in vitro</i> cell culture passage.....	67
2.3.1 General protocol.....	67
2.3.2 <i>In vitro</i> A549 passage: qRT-PCR and Bio-Plex® studies	67
2.4 Determination of viral titers and <i>in vitro</i> WNV Phenotype.....	68
2.4.1 Plaque titration: Vero cell culture	68
2.4.2 Temperature-sensitive (<i>ts</i>) <i>in vitro</i> phenotype	68
2.4.3 <i>In vitro</i> plaque morphologies	68
2.5 Sanger sequencing	68

2.5.1 Preparation of cDNA for sequencing	68
2.5.2 Sanger sequence analysis	70
2.6 Phylogenetic and sequence-based comparisons	70
2.6.1 Sequence Alignments (general)	70
2.6.2 Neighbor-joining (NJ) phylogenetic comparisons	70
2.6.3 Maximum Likelihood (ML) phylogenetic comparisons	71
2.6.4 Bayesian coalescent phylogenetic comparisons	71
2.6.5 Phylogenetic rendering	71
2.6.6 Selection studies	71
2.7 Next Generation Sequencing (NGS) pipe-line	72
2.7.1 Preparation of vRNA for Illumina HiSeq1000 sequencing	72
2.7.2 NGS pipe-line: downstream data processing	72
2.7.3 Source codes	73
2.7.3.1 FastQ contig alignment to NY99 reference sequence	73
2.7.3.2 Conversion from SAM to BAM file format	73
2.7.3.3 Removal of PCR duplicates	73
2.7.3.4 Down-sampling to normalize depth-of-coverage	74
2.7.4 Bayesian phylogenetic comparisons	74
2.7.5 Shannon entropy calculations	74
2.7.6 Single nucleotide variant (SNV) detection	75
2.8 Determination of <i>in vivo</i> mouse neuroinvasive phenotype	76
2.8.1 50% Lethal dose (LD ₅₀) studies	76
2.8.2 Ethics statement	76
2.9 Quantitative reverse-transcriptase PCR	76
2.9.1 RNA extraction	76
2.9.2 cDNA synthesis	77
2.9.3 SYBR® Green real-time PCR	78
2.9.4 Real-time PCR data management	78
2.9.5 NY99 infectious clone (NY99ic) studies: NS4B protein	79
2.10 Bio-Plex protein quantification	79
2.10.1 Preparation of cytokine/chemokine samples	79
2.10.2 BioRad® Bio-Plex Pro™ Human Cytokine 27-plex panel	80

CHAPTER 3.....	89
Dynamic Transmission of West Nile virus Across the United States-Mexican Border	89
3.1 Introduction.....	89
3.2 Results.....	90
3.2.1 Genomic sequence analysis: El Paso, Texas and Ciudad Juarez WNV isolates.....	90
3.2.2 Evidence of regional WNV evolution on the US-Mexican border.....	91
3.3 Discussion.....	93
CHAPTER 4.....	103
Continued Evolution of WNV, Houston, Texas, USA, 2002-2012.....	103
4.1 Introduction.....	103
4.2 Results.....	104
4.2.1 West Nile virus collection in Harris Co., Texas	104
4.2.2 Divergent WNV evolution in Harris Co., Texas, 2010-2012	105
4.2.2.1 Nucleotide changes	106
4.2.2.2 Amino acid substitutions.....	106
4.2.2.3 Phylogenetic analyses	107
4.2.2.4 Geospatial comparisons	108
4.2.3 Texas 2012 WNV outbreak: Novel WNV introduction event.....	109
4.2.4 Harris Co. paradigm—Model for WNV evolution in the U.S.	110
4.3 Discussion.....	110
CHAPTER 5.....	130
Natural Determinants of “Deep” WNV Population Structure	130
5.1 Introduction.....	130
5.2 Results.....	131
5.2.1 Bridging traditional Sanger and NGS methodologies.....	132
5.2.1.1 Consensus sequence comparisons.....	132
5.2.1.2 Phylogenetic comparisons	133
5.2.2 “Deep” NGS populations retain consensus-based trends	134

5.2.3 Host-dependent factors in natural WNV diversity.....	135
5.2.4 Determinants driving WNV population diversity: SNV detection .	136
5.2.5 Phenotypic link to WNV population diversity.....	138
5.3 Discussion.....	139
CHAPTER 6.....	163
Role of the West Nile virus NS4B protein in modulation of the innate immune response	163
6.1 Introduction.....	163
6.2 Results.....	164
6.2.1 <i>In vitro</i> platform: A549 cell culture	164
6.2.1.1 Multiplication kinetics and in vitro phenotype	164
6.2.1.2 Markers of innate immune modulation: Gene and protein expression	165
6.2.2 Role of the NS4B protein in the WNV innate immune response ...	167
6.2.2.1 Multiplication kinetics and in vitro phenotype	168
6.2.2.2 Differential gene/protein expression is both residue- and domain-dependent.....	168
6.2.2.3 NS4B protein-linked attenuation and innate immune expression	170
6.2.3 Role of the NS4B protein in natural WNV isolates	171
6.3 Discussion.....	172
CHAPTER 7.....	196
Discussion.....	196
7.1 Harris County (Co.) paradigm	197
7.2 Determinants of WNV population structure	200
7.3 Molecular determinants of innate immunity and attenuation	202
7.4 Conclusions.....	205

Appendix A: PrimePCR™ Flavivirus Infections H96 Panel.....	206
Appendix B: Protein Expression in the <i>in vitro</i> A549 Platform	210
Appendix C: NY99ic NS4B Mutant Gene Expression.....	211
Appendix D: NY99ic NS4B Mutant Protein Expression	214
Appendix E: Natural NS4B Mutant Protein Expression.....	216
References	219
Vitae	254
Curriculum Vitae	256

List of Tables

Table 1-1: Evolution of WNV lineage 1a—conserved amino acid substitutions	56
Table 2-1: WNV-specific reverse transcriptase (RT)-PCR primer sets.....	81
Table 2-2: Touch-down Titan One-Tube RT-PCR reaction protocol.....	82
Table 2-3: WNV-specific Sanger sequencing primers, related to Table 2-1	83
Table 2-4: BioRad® PrimePCR™ real-time PCR protocol	85
Table 3-1: West Nile virus isolates described in the conducted genetic and phylogenetic analyses, 1998-2010.....	97
Table 3-2: Nucleotide and amino acid divergence between El Paso and Ciudad Juarez WNV isolates	98
Table 3-3: Detected amino acid substitutions in applied El Paso and Ciudad Juarez WNV consensus sequences, 2005-2010	99
Table 4-1: West Nile virus isolates described in the conducted Harris Co., Texas and U.S. sequence and phylogenetic analyses, 1998-2012	115
Table 4-2: Detected amino acid substitutions in novel Harris Co., TX isolates, 2010- 2012	116
Table 4-3: Identified amino acid substitutions in Harris Co., TX isolates, 2002-2012	119

Table 5-1: WNV isolates in the described Next Generation Sequence (NGS) analyses.	144
Table 5-2: Depth-of-coverage and detected single nucleotide variants (SNVs).....	145
Table 5-3: Consensus differences between the NGS-indexed and published NY99- flamingo382-99 (NY99) sequences	146
Table 5-4: NGS and published consensus sequence discrepancies	147
Table 5-5: Encoded amino acid substitutions in the novel 2010 and 2011 Harris Co., TX WNV isolates.....	148
Table 5-6: Single nucleotide variants (SNVs) detection in natural WNV populations, <i>p</i> -value < 0.05 (Bonferroni-corrected).....	149
Table 5-7: Conserved SNVs in natural WNV populations, related to Table 5-6	155
Table 6-1: Natural and NY99ic-derived isolates applied in the <i>in vitro</i> A549 platform with documented <i>in vitro</i> and <i>in vivo</i> phenotypes	177
Table 6-2: NY99ic-derived NS4B mutant <i>in vitro</i> multiplication kinetics and temperature-sensitive (<i>ts</i>) phenotype in Vero cell culture	178
Table 6-3: Natural Texas WNV isolates encoding NS4B substitutions, related to Table 6-4	179
Table 6-4: Encoded non-NS4B amino acid substitutions in applied, natural Texas WNV isolates, related to Table 6-3	180

Table 6-5: Natural NS4B mutant <i>in vitro</i> multiplication kinetics and temperature-sensitive (<i>ts</i>) phenotype	182
---	-----

List of Figures

Figure 1-1: Simplified, condensed distribution of global WNV isolates in a Neighbor-joining tree clustered based on phylogenetic lineage (1-6), clade in lineage 1 (1a or 1b), or cluster in lineage 1a (1-6). U.S. isolates are indicated in red boldface. Scale bar, branch length representing % nucleotide divergence.	57
Figure 1-2: Bayesian phylogenetic tree depicting the simplified distribution of the North American genotypes within lineage 1a cluster 4. Posterior probabilities indicated ≥ 0.90 . Scale bar, divergence time in years.....	58
Figure 1-3: Natural WNV enzootic transmission cycle.....	59
Figure 1-4: National Incidence of clinical WNV disease in the U.S., 1999-2013. Number of death indicated per year (boldface red). WNF, West Nile fever; WNND, West Nile neuroinvasive disease.....	60
Figure 1-5: West Nile virus genome organization and processing.	61
Figure 1-6: 5' and 3' untranslated region (UTR) structure and organization in the WNV genome, related to Fig. 1-5 . Both A) conserved secondary RNA structures and B) long-distance base-pairings of RNA elements are indicated.....	62
Figure 1-7: West Nile virus life cycle.....	63
Figure 1-8: Innate immune responses to WNV infection.....	64
Figure 2-1: Illumina HiSeq1000 Next Generation Sequencing (NGS) pipe-line.....	86

Figure 2-2: BioRad® PrimePCR™ Flavivirus Infections H96 panel	87
Figure 2-3: Custom PrimePCR™ plate design.....	88
Figure 3-1: Bayesian-inferred 70% majority-rule phylogenetic trees of A) all published, full-length North American WNV isolates, 1999-2010; B) Group 1 isolates in the NA/WN02 genotype; and C) SW/WN03 genotype isolates. Boldface red, isolates sequenced in this chapter.....	101
Figure 3-2: West Nile virus reintroduction model on the U.S.-Mexican border. Dashed arrows, transmission of the SW/WN03 genotype from the southwestern U.S. to northern Mexico between 2003-2008 with U.S. reintroduction in 2008.....	102
Figure 4-1: Bayesian-inferred, 50% majority-rule, coalescent phylogenetic tree of published, full-length Harris Co., TX isolates, 2002-2012. Boldface, isolates sequenced in this chapter. Strain names link geographic map code (e.g. B1, B2, M1, M2)—refer to Fig. 4-2 . Year-of-collection annotated in parentheses. Scale bar indicates divergence time (in years).	126
Figure 4-2: Incidence of vector-borne WNV in Harris Co., Texas, 2002-2012, indicating the cumulative distribution of confirmed avian and mosquito (<i>Culex</i> and <i>Aedes</i> spp.) isolates. Small numbers, reference codes for each of the 268 mosquito control operational areas. Colored symbols indicate 2002-2012 Harris Co. isolates designated in Table 4-1 with open-symbols representing isolates clustering in Group 7 (red), Group 8 (blue), Group 9 (green), Group 10 (gold), or non-clustering isolates (black)..	127

Figure 4-3: Evolution of WNV in North America, 1999-2012. Bayesian coalescent tree of A) all published North American isolates clustering in the NY99 (blue), NA/WN02 (orange), or SW/WN03 (red) U.S. genotypes containing inferred monophyletic lineages B-G of the novel 2010-2012 Harris Co. isolate (indicated in boldface) color-coded (Panels B-G) based on **Fig. 4-1** designations. Posterior probabilities (range 0.00-1.00) indicated along branches. Scale bar, divergence time (in years).128

Figure 4-4: Bayesian phylogenetic support for expanded application of the Harris Co., Texas paradigm as a model for WNV evolution during the Dallas/Fort Worth, Texas outbreak, 2012. Boldface red, novel 2012 Collin and Denton Co. isolates sequenced in this chapter. Posterior probabilities ≥ 0.90 indicated. Scale bar, divergence time (in years).129

Figure 5-1: Bayesian-inferred coalescent phylogenetic tree using the NGS-indexed FASTA sequences of all applied, regional Houston, Texas isolates. Phylogenetic relationships indicated in the NY99 (blue), NA/WN02 (orange), and SW/WN03 (red) U.S. genotypes and published Harris Co. paradigm groupings (Groups 1-10)^{63,351}. Scale bar indicates divergence time (in years).157

Figure 5-2: West Nile virus incidence in Harris Co., Texas, 2002-2012. Shannon entropies (S_n ; scale $\times 10,000$) indicated for all sequenced bird- (open circles) and mosquito pool-derived (filled circles) WNV isolates per year⁴⁷.158

Figure 5-3: Host-specific genetic variation across the WNV genome for both A) bird- and B) mosquito pool-derived isolates per genomic region. Superimposed values, nucleotide positions exhibiting >0.05 mean S_n (red dashed line) per host-type. Circles represent detected, conserved synonymous (open, black) and non-synonymous (filled red) SNVs in bird-derived isolates relative to their genomic position.....159

Figure 5-4: Detection of conserved SNVs in natural WNV populations. A) SNVs identified in >1 WNV isolate with both synonymous (black) and non-synonymous (red) variants indicated relative to their genomic. B) Proportion of both bird- and mosquito-derived isolates encoding 4 of the 7 highlighted SNVs conserved in the 5'/3'UTRs at 1.0 to $>20\%$ mutational frequencies or in the encoded consensus sequence (black). *Six of the eight detected SNVs at position 10888 were significant ($p < 0.05$) in SNVer alone.160

Figure 5-5: West Nile virus virulence is linked to viral RNA population diversity, related to **Fig. 5-6**. Mean calculated Shannon entropy (scale $\times 10,000$) is indicated per region of the WNV genome for all applied isolates from *low* (dark blue; 0) to *high* (dark red; ≥ 600) genetic diversity.161

Figure 5-6: Mouse neuroinvasive phenotype is linked to increased WNV genetic diversity, related to **Fig. 5-5**. Mean Shannon entropy (\pm SEM) is indicated for both attenuated (grey) and virulent (red) isolates per region in the WNV genome. Dots, discrete mean values per isolate. Not significant, n.s.162

Figure 6-1: Flavivirus NS4B Protein Alignment.....183

Figure 6-2: <i>In vitro</i> A) multiplication kinetics and B) temperature-sensitive (<i>ts</i>) phenotype of applied isolates in the <i>in vitro</i> A549 expression platform	185
Figure 6-3: Expression profiles of targeted genes (clustered by function) in the <i>in vitro</i> human A549 expression platform at both 24 and 36 hpi. Fold up/down-regulation indicated relative to mock infection	186
Figure 6-4: Fold-expression of targeted genes at 24 hpi in the PrimePCR™ Flavivirus Infections H96 panel (relative to NY99). Bold red, significant up/down-regulated expression ($p < 0.05$; $\alpha = 0.05$)	187
Figure 6-5: Fold-expression of targeted genes at 36 hpi in the PrimePCR™ Flavivirus Infections H96 panel (relative to NY99). Bold red, significant up/down-regulated expression ($p < 0.05$; $\alpha = 0.05$)	188
Figure 6-6: Human cytokine/chemokine protein expression in the <i>in vitro</i> A549 platform at both 24 and 36 hpi (relative to NY99 infection)	189
Figure 6-7: Predicted WNV NS4B protein membrane topology <i>pre-cleavage</i> based on the DENV-2 NS4B topology. Putative transmembrane domains (pTMD), transmembrane domains (TMD), and loops indicated.....	190
Figure 6-8: Predicted WNV NS4B protein membrane topology <i>post-cleavage</i> based on the DENV-2 NS4B topology, related to Fig. 6-7	191
Figure 6-9: Residue- and domain-dependent gene expression (relative to NY99ic) for WNV NS4B mutants in the <i>in vitro</i> A549 platform	192

Figure 6-10: Fold-protein induction in NS4B mutant infection (relative to NY99ic)	193
.....	
Figure 6-11: Differential gene expression between NS4B mutants: A) P38A vs. P38G (attenuated) and B) C102A vs. C102S (attenuated). C) Fold-expression ≥ 2 -fold is indicated per comparison with significant fold-changes (p - value < 0.05 ; $\alpha = 0.000417$) designated in boldface red.....	194
Figure 6-12: Fold-protein induction in natural WNV isolates encoding NS4B substitutions (relative to NY99-flamingo382-99).....	195

List of Abbreviations

°C	Degrees Celsius
A	Absorbance
A	Adenine
ABSL3	Animal Biosafety Level 3
<i>Ae</i>	<i>Aedes</i> spp.
AFP	Acute flaccid paralysis
Ala (A)	Alanine
ANOVA	Analysis of variance
Arg (R)	Arginine
Asn (N)	Asparagine
Asp (D)	Aspartate
BAM	Binary alignment/map
bp	Base-pair
C protein	Capsid protein
C	Cytosine
CA	California
CDC	Centers for Disease Control and Prevention
cDNA	complimentary DNA
ClustalW	Clustal Omega
cm	Centimeter
Co.	County
CO ₂	Carbon dioxide
CPE	Cytopathic effects
CS	Conserved sequence
CSF	Cerebrospinal fluid
CT	Connecticut
C _T	Cycle threshold value
C-terminus	Carboxy (COO ⁻) terminus or terminal domain
<i>Cx.</i>	<i>Culex</i> spp.
Cys (C)	Cysteine
DB	Dumbell
DC	Dendritic cell
DC-SIGN	Dendritic cell-specific ICAM-grabbing non-integrin
DENV	Dengue virus
DMSO	Dimethyl sulfoxide
d _N	Rate of non-synonymous substitution
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide
dpi	Days post-infection
d _s	Rate of synonymous substitution
dsRNA	double-stranded RNA
DTT	Dithiothreitol

E protein	Envelope protein
EDIII	E protein domain III
ER	Endoplasmic reticulum
EtBr	Ethidium bromide
FBS	Fetal bovine serum
FGF Basic	Basic fibroblast growth factor
G	Guanine
GAG	Glycosaminoglycan
GB	Gigabyte
G-CSF	Granulocyte colony-stimulating factor
GHz	Gigahertz
Gln (Q)	Glutamine
Glu (E)	Glutamate
Gly (G)	Glycine
GM-CSF	Granulocyte macrophage colony-stimulating factor
GTR	General time reversible
His (H)	Histidine
HKY-85	Hasegawa-Kishino-Yano 85
HP	Hairpin
hpi	Hours post infection/inoculation
I	Invariant
i.c.	Intracranial
i.p.	Intraperitoneal
ID	Idaho
IFIT	IFN-inducible tetratricopeptide
IFITM	IFN-induced transmembrane
IFN-($\alpha/\beta/\gamma$)	Interferon- $\alpha/\beta/\gamma$
IFNAR	IFN- α and IFN- β receptor complex
Ig	Immunoglobulin
IL	Illinois
IL	Interleukin
IL-1R α	IL-1 receptor alpha
Ile (I)	Isoleucine
IP-10	Interferon gamma-induced protein 10 (aka CXCL10)
IPS-1	Interferon beta promoter stimulator 1
IRF	INF regulatory factor-3/4/5/7/9
ISGF3	ISG factor 3
ISGs	IFN-stimulated genes
Jak	Janus kinase
JEV	Japanese encephalitis virus
kb	kilobase
kDa	Kilodalton
KOUV	Koutango virus
KUNV	Kunjin virus
LD ₅₀	50% lethal median dose
Leu (L)	Leucine

<i>lp</i>	Large plaque morphology
Lys (K)	Lysine
M protein	Membrane protein
MAVS	Mitochondrial antiviral signaling protein
MCP-1	Monocyte chemotactic protein 1 (aka CCL2)
MD	Maryland
MDA-5	Melanoma differentiation-associated gene 5
MEM	Minimal essential media
Met (M)	Methionine
min	minute(s)
MIP-1 α	Macrophage inflammatory protein 1 alpha (aka CCL3)
MIP-1 β	Macrophage inflammatory protein 1 beta (aka CCL4)
ML tree	Maximum likelihood tree
mL	milliliter
mm	Millimeter
mM	Millimolar
MOI	Multiplicity of infection
<i>mp</i>	Medium plaque morphology
MUSCLE	Multiple sequence comparison by log-expectation
MVEV	Murray Valley encephalitis virus
MW/WN06	Midwest genotype (WN06)
Myd88	Myeloid differentiation factor 88
N	Ambiguous base
NA/WN02	North American genotype (WN02)
NC	Nucleocapsid
ND	North Dakota
NEAA	Non-essential amino acids
ng	nanogram
NGS	Next generation sequencing
NJ tree	Neighbor-joining tree
NJ	New Jersey
NLR	NOD-like receptor
nm	nanometer
nM	Nanomolar
NOD	Nuclear organization domain
NS1 protein	Non-structural protein 1
NS2A protein	Nonstructural protein 2A
NS2B protein	Nonstructural protein 2B
NS3 protein	Nonstructural protein 3
NS4A protein	Nonstructural protein 4A
NS4B protein	Nonstructural protein 4B
NS5 protein	Nonstructural protein 5
nt	Nucleotide(s)
<i>N</i> -terminus	Amino (NH ₃ ⁺)-terminus or terminal domain
NY	New York
NY99	NY99-flamingo382-99 prototype U.S. strain

NY99ic	NY99 infectious clone
OAS1	2'-5'-oligoadenylate synthetase 1
OHFV	Omsk hemorrhagic fever virus
ORF	Open reading frame
<i>p</i>	<i>p</i> -value
PAMPs	Pattern-associated molecular patterns
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PDGF	Platelet-derived growth factor
PFU	Plaque forming units
pg	picogram
Phe (F)	Phenylalanine
PKR	Protein kinase R
PMN	Peripheral mononuclear cell
POWV	Powassan virus
prM protein	Pre-membrane protein
Pro (P)	Proline
PRR	Pattern recognition receptor
pTMD	Putative TMD
Q	Phred-like quality score
qRT-PCR	Quantitative reverse-transcriptase PCR
RabV	Rabensburg virus
RANTES	Regulated on activation, normal T cell expressed and secreted
RIG-I	Retinoic-acid-inducible gene I
RLR	RIG-I-like receptor
RNA	Ribonucleic Acid
RNaseL	Ribonuclease L
SAMtools	Sequence alignment map toolbox
SEM	Standard error of the mean
Ser (S)	Serine
SL	Stem loop
SLEV	St. Louis encephalitis virus
S_n	Shannon entropy
SNV	Single nucleotide variant
<i>sp</i>	Small plaque morphology
ssRNA	Single-stranded RNA
STAT1/2	Signal transducers and activators of transcription 1 and 2
SW/WN03	Southwest genotype (WN03)
T cell	T lymphocyte
T	Thymine
TBEV	Tick-borne encephalitis virus
TGN	<i>Trans</i> -golgi network
Thr (T)	Threonine
TIR	Toll/IL-1 receptor
TLR	Toll-like receptor
TMD	Transmembrane domain

TNF- α	Tumor necrosis factor alpha
TPB	Tryptose phosphate broth/buffer
TRE	Directory tree file
TRIF	TIR domain-containing adaptor-inducing IFN- β
TRIM25	Tripartite motif-containing protein 25
Trp (W)	Tryptophan
<i>ts</i>	Temperature-sensitive
TX	Texas
Tyr (Y)	Tyrosine
U	Uracil
U.S.	United States
UTMB	University of Texas Medical Branch
UTR	Untranslated region (either 5' and 3')
UV	Ultraviolet
Val (V)	Valine
VCF	vCard file
VEGF	Vascular endothelial growth factor
vRNA	viral RNA
WNF	West Nile fever
WNND	West Nile neuroinvasive disease
WNV	West Nile virus
WRCEVA	World Reference Center for Emerging Viruses and Arboviruses
YFV	Yellow fever virus
Γ	Gamma distribution
μg	Microgram
μL	Microliter

CHAPTER 1

Introduction

1.1 MOLECULAR EPIDEMIOLOGY AND PHYLOGENETICS¹

1.1.1 West Nile virus classification

West Nile virus (WNV) is an “Old World” virus which re-emerged in sporadic epizootic outbreaks in the mid-1990s to international concern as a rising, biological threat to public health. Initial isolation of the virus occurred during routine YFV surveillance in sub-Saharan Africa from a febrile woman in the West Nile district of Uganda¹; paired host-range, electron micrograph (EM)-based, and serological (*e.g.* antisera and hemagglutination inhibition (HI) tests) studies confirmed common characteristics and cross-neutralization with 14 other Group B *Togaviridae* mosquito-borne arboviruses (flaviviruses) and nine members of the “Japanese encephalitis (JE) serocomplex”—including JE, Kunjin (KUN), Koutango (KOU), St. Louis encephalitis (SLE), and Murray Valley encephalitis (MVE) viruses²⁻⁵. Further studies applied the distinct antigenic and structural flavivirus characteristics in establishing the taxonomic designation *Flaviviridae* genus *Flavivirus*⁶⁻⁸. Integration of traditional genetic and phylogenetic techniques in the mid-1990s, related these original *Flavivirus* designations in the JE serocomplex with shared genomic structure launching further, international inquiries into the genotypic origins of both “Old World” and “New World” WNV epizootic outbreaks⁹.

In this regard, sporadic outbreaks of febrile disease in Africa, Australia, and the Middle East and larger, self-limited outbreaks in Israel, France, and South Africa

¹ Text, figures, and graphs in this section are reproduced and/or modified with permission from **Mann, B.R.** *Int J Environ Res Public Health*; published by MDPI, 2013—based on the terms and conditions of the Creative Common Attribution license. The citation is: **Mann, B.R., McMullen, A.R., Swetnam, D.M., & Barrett, A.D.T. (2013).** Molecular epidemiology and evolution of West Nile virus in North America. *Int J Environ Res Public Health* **10**(10), 5111-5129.

throughout the 1950s-1980s linked clinical WNV infection with overt, epidemic febrile disease^{10,11}. Incidence of severe WNV-related human and equine (horses) disease gained momentum in the mid-1990s with wide-spread, regional outbreaks in northern Africa, the Middle East, and central Europe culminating in the 1996 Romanian epidemic marking the first significant epidemic incidence of human encephalitic disease^{10,11}. Paired serological and host/vector surveillance during these epidemics since the mid/late-1950s has provided an expanding collection of both prototype and regional WNV strains applied to in-depth phylogenetic studies. Using this approach, identification of conserved and divergent mutational trends in applied isolates has aided in deciphering underlying genotypic determinants involved in the global evolution and expansion of WNV to the western hemisphere.

1.1.2 Phylogenetics: defining a lineage

Based on these phylogenetic studies of the genome, WNV is classified into five distinct lineages—designated 1-5 with KOUV as a potential lineage 6—based on nucleotide (nt) divergence $\geq 20\%$ between isolates in different lineages (**Fig. 1-1**)¹²⁻¹⁴. Lineages 1 and 2 constitute the largest global distribution of circulating WNV isolates including endemic and epidemic strains collected from humans, horses, birds, and mosquitoes in Europe, Africa, the Middle East, Australia (KUNV sub-type), and the Americas^{13,15,16}. Historical circulation of lineage 2 strains was restricted to local, endemic transmission cycles in sub-Saharan/southern Africa and Madagascar with recent introductions into both Europe (Cyprus, Greece, Hungary, Italy, and Romania) and Russia¹⁷⁻²⁷. In-depth genetic comparisons group lineage 2 isolates into four clades (2a-2d) based on 0.2-16.9% nt divergence and conserved clade-specific substitutions^{13,24,28,29}. Lineage 3 consists of a single 1997 isolate collected from a WNV-positive *Culex pipiens* mosquito pool in the Czech Republic—designated Rabensburg virus (RabV)—exhibiting limited *in vitro* replication in mammalian cells and *in vivo* virulence in either mice or

birds^{30,31}. Lineage 4 includes several Russian isolates identified in circulation since 1988³². Lineage 5 consists of 13 1955-1982 Indian isolates^{13,33}.

Lineage 1 is classified into two distinct clades: 1a and 1b (Kunjin virus sub-type circulating in Australasia) constituting the largest global WNV distribution. In particular, Clade 1a is further sub-divided into six clusters (1-6) demonstrating $\leq 5.4\%$ nt divergence despite broad geographic coverage (in some clusters) and conserved genotypic signatures (**Fig. 1-1** and **Table 1-1**)^{13,34,35}. Clusters 1 and 2 include isolates from northern Africa, Europe, the Middle East, Russia, and India; in particular, isolates in cluster 2 encode a conserved NS1-A70S substitution with further sub-division into either the Eastern European or Mediterranean sub-types based on conserved NS1-L206F, NS2B-A103V, and NS5-T898I substitutions in all Eastern European sub-type isolates^{13,36-40}. Cluster 3 includes 1995-2005 isolates collected from the Russian Astrakhan region encoding two unique substitutions: NS1-S99P and NS2A-A224T. Both clusters 5 and 6 are composed of 1965-1979 central African isolates^{13,16}. Relevant to this dissertation, cluster 4 constitutes four “Old World” isolates collected from Tunisia (1997), Israel (1998 and 2000), and Hungary (2003) and an increasing repertoire of “New World” isolates from the Americas (1999-present) exhibiting conserved expression of both the E-T126I and NS4A-V85A substitutions—indicated in **Fig. 1-1** (red) and **Table 1-1**¹³. In line with these phylogenetic designations, divergent evolution of a proline (residue 249) in the NS3 protein is lineage-specific: lineage 2 (H), lineage 3 (N), lineage 4 (T), lineage 5 (T), and lineages 1a (T/P) and 1b (A) linked to host-selected virulence in birds (**Table 1-1**)^{13,41}.

1.1.3 Introduction of an “Old World” virus into the “New World”

Initial isolation of this “Old World” virus in the western hemisphere—the “New World”—followed a self-limited, epizootic outbreak in the New York metropolitan area in 1999 that resulted in 62 human encephalitic cases concurrent with 25 equine cases (9 fatal) and wide-spread mortality in local wild bird populations^{42,43}. From 1999-2001,

WNV surveillance efforts detected WNV-positive dead birds (*Corvus* spp. in particular) and *Culex* spp. mosquito pools in CT, MD, NJ, and NY confirming the *in situ* establishment of WNV in endemic and enzootic cycles throughout the northeastern U.S.^{42,44-46}. Expansion continued down the eastern coast and westward across the continental U.S. with detection in all 48 contiguous states by 2004^{47,48}. National and international surveillance campaigns have since confirmed WNV circulation in regions of Canada, the Caribbean, Central America, and South America⁴⁹⁻⁵³.

How did WNV accomplish its trans-Atlantic immigration in the U.S.? Numerous theories exist; however, few rate more than speculation. In this regard, initial, full-length sequencing of the prototype NY99-flamingo382-99 U.S. isolate (NY99; AF196835)—cultured from a dead Chilean flamingo (*Phoenicopterus chilensis*) in the Bronx zoo—in parallel with E-glycoprotein antigenic mapping revealed 99.8% nt sequence identity with a 1998 Israeli domestic goose isolate (IS-98 STD; AF481864) in lineage 1a Cluster 4^{43,54}. Limited sequence similarity ($\leq 96.9\%$) to other published lineage 1a cluster 2 (European) or Clade 1b (Kunjin) strains relative to cluster 4 sub-Saharan Africa isolates supported the independent initiation of both the Israeli and New York outbreaks in the late-1990s upon introduction of a common progenitor strain from an unknown location in northern Africa^{13,54}.

1.1.4 North American (U.S.) WNV genotypes

However it began, the establishment of WNV in local transmission cycles fueled its rapid expansion from the northeastern U.S. into naïve, domestic bird and mosquito populations throughout North America with the resulting emergence, extinction, and co-circulation of multiple U.S. genotypes (**Fig. 1-2**):

1.1.4.1 NY99 GENOTYPE

Initial phylogenetic studies focusing on the WNV pre-membrane/membrane (prM/M) and envelope (E) protein genes in 1999-2001 CT and NY surveillance cohorts confirmed the *in situ* evolution of a homogeneous WNV population (termed the NY99 genotype) in local, naïve *Culex* spp. mosquito and bird populations in the northeastern U.S.—named for the included NY99 prototype isolate. Limited genetic change (<0.18% nt divergence) is observed between isolates in this original genotype with detection of synonymous mutations not fixed within the population⁴²⁻⁴⁶.

1.1.4.2 SOUTHEAST (SE) COASTAL TEXAS GENOTYPE

Following the introduction of WNV into Texas in 2002, surveillance efforts in the coastal region of southeast Texas confirmed the emergence of a divergent cohort of seven uniform isolates encoding five amino acid substitutions: E-T76A, NS1-E94G, NS2A-V138I, NS4B-V173I, and NS5-T526I supporting its designation as a novel U.S. genotype—termed the SE coastal Texas genotype^{55,56}. No additional isolates (to date) have been detected belonging to this regional U.S. genotype, suggesting that it has become extinct.

1.1.4.3 NORTH AMERICAN/WN 2002 (NA/WN02) GENOTYPE

Rapid expansion of WNV transmission in the central/Midwestern U.S. coincided with peak clinical incidence between 2002-2003 with a combined total of 5,812 WNND cases and 548 fatalities^{47,48}. Phylogenetic analyses of paired 2001-2004 North American isolates confirmed the displacement of the original NY99 genotype with a heterogeneous pool of isolates (up to 0.58% nt divergence)—termed the North American or West Nile 2002 (NA/WN02) genotype—in 2002 as defined by 13 conserved nt transitions and a single encoded substitution, V159A, in the E protein (**Table 1-1** and **Fig. 1-2**)⁵⁷⁻⁵⁹.

1.1.4.4 SOUTHWESTERN/SW 2003 (SW/WN03) GENOTYPE

Progressive declines in the annual incidence of clinical WNV disease from 2004-2011 prompted sustained surveillance efforts to elucidate novel genotypic determinants driving the selection of regional homeostatic virus populations and a decreasing national growth rate consistent with now endemic circulation in the U.S.^{47,48,60-62}. McMullen *et al.* (2011) investigated a heterogeneous pool of 2003-2009 isolates collected in the southwestern U.S. and northern Mexico confirming the circulation of a novel U.S. genotype—termed the southwestern (SW/WN03) genotype first isolated in 2003 (**Fig. 1-2**). This genotype in 2003 is defined by 13 unique nt changes with positive selection for both the NS4A-A85T and NS5-K314R substitutions in multiple isolates⁶³. Recent studies using 2006-2011 isolates identified a novel monophyletic cluster of 2006-2007 isolates derived from human blood donors and birds in ID and ND—termed the Midwestern (MW/WN06) cluster, **Fig. 1-2**—within the SW/WN03 genotype encoding the signature NS4A and NS5 substitutions^{64,65}. Both the NA/WN02 and SW/WN03 genotypes circulate in the U.S. to date^{63,64}.

1.1.4.5 RISE AND FALL OF U.S. GENOTYPES: UNDERLYING FACTORS

Despite the confirmed emergence, co-circulation, and apparent extinction of multiple local or national U.S. genotypes from 1999-2011, current advances in identifying specific genotypic or molecular determinants driving the continued evolution of WNV in the U.S. still remain limited. In particular, selection of the E-V159A substitution in all published U.S. isolates (to date) has been linked to a reduced extrinsic incubation period in *Culex* spp. mosquitoes; in effect, extinction of the original NY99 and regional SE coastal Texas genotypes is attributed to more efficient dissemination of NA/WN02 genotype isolates (encoding this substitution) in the mosquito vector^{59,66}. However, this hypothesis remains controversial due to identified species- and region-

specific differences in *Culex* spp. mosquito populations⁶⁷. Other proposed factors include annual reintroduction, local over-wintering, and/or long-distance migration events in bird or mosquito hosts as alternative, potential drivers in the evolution of WNV in the U.S.^{49,68,69}. Furthermore, the biological and epidemiological roles of the conserved NS4A-A85T and NS5-K314R substitutions in the SW/WN03 genotype remain undefined.

1.2 ECOLOGY AND NATURAL TRANSMISSION

West Nile virus is maintained in a natural, enzootic transmission cycle with birds as the principal amplifying host and mosquitoes (*Culex* spp. in particular) serving as the principal vector (**Fig. 1-3**). However, not all hosts are equally important with evidence of both host- and vector-specific viral competencies⁷⁰⁻⁷⁶. In particular, horses, humans, and other mammals—with limited exceptions (*e.g.*, chipmunks, eastern cottontail rabbits, and fox squirrels)—serve as “dead-end” hosts with insufficient peak viremia (virus titer in the serum/blood) needed to infect a mosquito upon ingestion of an infected bloodmeal^{48,70,77-82}.

1.2.1 Principal vector: mosquitoes

Local, regional, and national vector-surveillance campaigns, from 1999-2012, in the U.S. detected WNV in 65 distinct mosquito species including *Aedes* spp., *Anopheles* spp., *Culex* spp., and *Culiseta* spp. (among others)⁴⁸; however, ornithophilic (bird-seeking) *Culex* (*Cx.*) spp. mosquitoes appear to drive both epizootic and epidemic transmission with spill-over feeding on humans, horses, and other mammals as opportunistic hosts^{83,84}. In effect, wide-spread distribution of several dominant, regional species: *Cx. pipiens pipiens* (north)^{45,85}, *Cx. salinarius* (north-east)^{86,87}, *Cx. pipiens quinquesfasciatus* (south)⁸⁵, and *Cx. tarsalis* (west) throughout the U.S. further

facilitated the establishment of WNV into local, enzootic transmission cycles during its westward expansion^{71,88-90}. Other *Culex* species: *restuans*, *nigripalpus*, and *stigmatosoma* have been implicated as potential bridge vectors in the enzootic maintenance and amplification of WNV early in the annual transmission season⁷⁵; however, the overall impact of these vector populations is less well-defined.

From an ecological perspective, vector-competence is a critical part of the viral transmission cycle and consists of multiple steps. Upon ingestion of an infected bloodmeal, WNV must overcome several barriers to transmission including **(1)** infection of the midgut⁹¹; **(2)** escape from the midgut⁹²; **(3)** salivary gland infection⁹³; and **(4)** salivary gland escape^{71,94}. How well a virus can overcome these barriers is termed vector-competence, and how long it takes is the extrinsic incubation period (EIP)⁷¹. Based on these principles, a number of intrinsic and extrinsic factors affect *Culex* spp. vector-competence for WNV infection including population density, longevity, immature survival, infected viral genotype, and temperature (among others)^{71,95-98}. In this regard, displacement of the original NY99 genotype with the dominant NA/WN02 genotype coincided with a reduced EIP in *Cx. pipiens* mosquitoes attributing vector-adaptation to selection of the E-V159A substitution^{59,66}.

1.2.2 Principal host: birds

“Vertebrate-competence” (similar to vector-competence) is the extent to which a virus can overcome host-specific barriers in various vertebrates and replicate to peak viremias sufficient to infect a mosquito upon ingestion of a bloodmeal⁹⁹. For WNV, this includes birds; however, as would be expected, not all birds are competent hosts. In fact, (until 2012) WNV has been isolated from 326 bird species in the U.S. alone, but Passeriformes—corvids in particular—serve as the principal WNV-competent avian hosts in the U.S.^{47,48,70,99-102}. Extensive *in vivo* studies confirmed a 10⁴-10⁵ pfu/mL minimum threshold in peak WNV viremia needed to infect feeding *Cx. pipiens* and

quinquefasciatus mosquitoes^{76,100}. In this regard, certain Passeriformes—including the American crow (*Corvus brachyrhynchos*) and blue jay (*Cyanocitta cristata*)—have >10⁹ pfu/mL mean peak viremias¹⁰⁰. However, *Corvidae* species are also more susceptible to WNV infection and succumb to West Nile disease^{70,99-102}. In effect, several alternative Passeriforme species: house sparrows, house finches, and American robins¹⁰³⁻¹⁰⁵ (among others) have been implicated as critical amplification hosts in both the local spread and long-distance migration (*e.g.*, American robin) of WNV in the U.S.^{101,106,107}.

Despite extensive U.S. surveillance efforts in wild bird populations, elucidation of critical molecular determinants driving WNV host-competence and/or adaptation in U.S. avian species remains limited. In this regard, applied *in vivo* studies identified a single positively selected locus (249) in the NS3 protein linked to both a virulent (P residue) and non-virulent (T residue) phenotype in American crows^{41,108}. Furthermore, envelope (E) protein-linked glycosylation—(N-Y-T/S) motif at residues 154-156—produced higher peak viremias and reduced survival times in both American crows and house sparrows compared to variants lacking the E-glycosylation motif consistent with prior *in vivo* mouse studies¹⁰⁹⁻¹¹³. However, additional research is still needed.

1.2.3 Trade-Off Hypothesis: WNV host/vector fitness

Vertebrate (bird)- and vector (mosquito)-competence sustain an inverse relationship in the selection and adaptation of natural WNV isolates; unique biological barriers in both hosts restrict or expand infected WNV populations resulting in the selection of constituent members able to adapt to the disparate, compounding pressures (bottlenecks) in natural, enzootic WNV transmission—albeit at the cost of enhanced viral fitness in either host. This theoretical model in competing host fitness as a natural constraint in viral evolution and adaptation is termed the Trade-Off Hypothesis¹¹⁴⁻¹¹⁷. In terms of natural WNV evolution, an expansion in viral populations is observed in the *Culex* spp. mosquito vector with subsequent purifying selection in the bird host despite

no apparent fitness costs^{116,118-122}. How virus RNA population structure (referred to as *viral quasispecies*) is defined through host selection and its impact on viral adaptation and *in vivo* host phenotype are critical questions in defining the past, present, and future of WNV evolution in the U.S.

1.2.4 Implications in human health: epidemiology in the U.S.

Historical records highlight the local, sporadic impact of self-limited WNV epizootic outbreaks in the “Old World”; in this regard, outbreaks in northern/sub-Saharan Africa, the Middle East, Australia, and Europe from 1951-1984 resulted in annual infections numbering in the hundreds of cases—few developing encephalitic complications^{14,123}. Incidence not only gained momentum in the mid/late-1990s but heralded a paradigmatic shift in clinical disease with unprecedented levels of clinical and symptomatic West Nile neuroinvasive disease (WNND) in humans, horses, and birds including naïve host populations in North America^{10,11,14,47,48,123}. In the U.S. alone, >39,000 confirmed WNV infections (with >17,000 WNND cases) and 1,668 deaths have been reported to the Centers for Disease Control and Prevention (CDC) through 2014 (**Fig. 1-4**)^{47,48}. Clinical WNV incidence follows the classic iceberg effect with an estimated 80% of infections being asymptomatic, 20% developing acute West Nile fever (WNF), and <1% culminating in WNND with encephalitis, meningoencephalitis, and severe neurological sequelae that prove fatal in ~10% of WNND cases⁴⁸. Based on these epidemiological data, an estimated 1:150 to 1:250 natural infections progress to WNND (dependent on a range of host determinants) corresponding to 2.6-4.4 million predicted human infections in the U.S. from 1999-2012¹²⁴⁻¹²⁶. However, passive reporting of confirmed, clinical diagnoses to the CDC ArboNET database coupled with discontinued WNF reporting mandates in 2004 limit the real-time detection of active disease incidence in local, regional, or national U.S. populations⁴⁷. Furthermore, viremic blood donors were not reported until 2003 following reformation of U.S. blood collection organizations and

national WNV reporting guidelines with implementation of nucleic acid amplification testing (NAT)^{124,127,128}. Despite these deficits, paired on-going host/vector surveillance campaigns in several U.S. regions: CA, NY, and TX (among others) continue to provide a predictive model of human risk allowing regional and state-wide programs to better facilitate public education and clinical preparedness.

1.3 MOLECULAR VIROLOGY

1.3.1 Structure and genomic organization

West Nile virus is a 50 nm enveloped virus encoding a single-stranded, positive-sense RNA genome and single 10-kb open reading frame (ORF) flanked with both a 5' and 3' untranslated region (UTR). Packaged RNA molecules are capped at the 5'-end with a Type-I m⁷GpppAmp cap structure terminating at the 3'-end with the 5'-CU_{OH}-3' element^{129,130}. Translation of this capped RNA molecule produces a single polypeptide sequence that is both co- and post-translationally cleaved into three structural: capsid (C), pre-membrane (prM), and envelope (E) and seven nonstructural (NS) proteins—NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 (**Fig. 1-5**). Host signalase, furin, and still undefined peptidases cleave pr-M, prM-E, E-NS1, NS1-NS2A, and 3'-end of the NS4A protein; in addition, the virus-encoded NS2B/NS3 serine protease cleaves C-prM and all remaining junctions in the NS proteins¹³¹⁻¹³³. Included below is a brief description of the structure and function of all encoded structural and NS protein and UTR elements.

1.3.1.1 STRUCTURAL PROTEINS

Post-translational processing of the C, prM, and E structural proteins in the host endoplasmic reticulum (ER)-lumen and Golgi apparatus provide the protein scaffold in formation of immature, non-infectious virus particles¹³⁴. Dimerization of the C protein

forms the nucleocapsid core binding genomic RNA; C protein dimers (12 kDa) exhibit an asymmetric high (positive) net charge distribution that forms both an apolar surface (interacts with the virion envelope) and basic surface (binds genomic RNA)¹³⁵. Furin-mediated cleavage of the precursor 22 kDa prM protein is required to produce infectious virus particles with retention of the 8 kDa C-terminal portion of the mature, membrane-associated M protein in mature, extracellular virions^{136,137}.

In addition, the *N*-linked glycosylated 50-55 kDa E protein forms a head-to-tail dimer with three β -barrel domains (I-III) per monomeric subunit. Domain II encodes the dimerization domain with the fusion peptide at its distal end; domain III (EDIII) encodes immunoglobulin (Ig)-like regions and putative receptor binding sites linked to WNV type-specific epitopes—residues 307, 330, and 332 (among others) in the WNV EDIII—and host cell receptor interactions^{134,138-142}. Domain I is the central domain that connects both domain II and III via short, flexible loops. In addition, *N*-linked glycosylation at residue 154 in the 154-156 (N-Y-T/S) motif has been identified as an *in vivo* virulence determinant^{28,109,112,113,143}. In immature virions, prM-E complex in 90 heterodimers that form 90 antiparallel E protein homo-dimers arranged in pseudo-icosahedral $T = 3$ symmetry post-prM cleavage^{132,144}; additional prM protein chaperonin-like activity has been implicated in maintaining E protein folding patterns in immature virus particles¹⁴⁵.

1.3.1.2 NONSTRUCTURAL PROTEINS

In-depth structural and phenotypic studies in several flaviviral systems: WNV, DENV, YFV, and JEV (among others) have implicated co-localization of the NS proteins in the formation/function of the flaviviral replication complex, virion assembly, and post-translational protein modifications^{131,133,146}.

NS1 protein

NS1 is a multi-functional *N*-linked glycosylated, oligomeric (46-55 kDa) protein that dimerizes following post-translational modification in the ER-lumen and is secreted as a soluble hexamer with a lipid core¹⁴⁷⁻¹⁵². In particular, *trans*-complementation of the NS1 dimer regulates negative-strand RNA synthesis^{149,152-154}. *N*-linked glycosylation of the WNV NS1 protein at three residues: 130, 175, and 207 has been implicated as a determinant of mouse *in vivo* neuroinvasiveness^{155,156}. Extracellular secretion of soluble NS1 is also correlated with host complement responses¹⁵⁷⁻¹⁶¹.

NS3 protein

NS3 is a 70 kDa monomeric protein encoding an *N*-terminal 180 aa serine protease function facilitating both *cis*- and *trans*-cleavage of the encoded polyprotein complemented in *trans* with the NS2B protein as an essential cofactor¹⁶²⁻¹⁶⁴. Post-translational protease function results in the cleavage of several encoded junctions—including NS2A-NS2B (*cis*), NS2B-NS3 (*cis*), NS3-NS4A, and NS4B-NS5—in addition to auto-cleavage (**Fig. 1-5**)¹³². In addition, the *C*-terminal two-thirds of the NS3 protein encodes RNA helicase¹⁶⁵, nucleotide triphosphatase¹⁶³, and RNA triphosphatase¹⁶⁶ activities with putative regulation of virus particle production¹⁶⁷.

NS4B protein

NS4B is a 27 kDa hydrophobic, dimeric protein embedded in the host perinuclear ER membrane with no known enzymatic function^{131,168,169}. However, NS4B initiates remodeling of the ER-membrane implicated in its potential *cis*-acting structural role in the flaviviral replication complex^{170,171}. Putative interactions have also been identified with both the NS1 and NS3 proteins; in particular, interaction of NS4B with NS3 dissociates NS3 from single-stranded RNA enhancing its helicase activity^{165,172,173}. Further attempts to define the phenotypic role of the NS4B protein have implicated

several residues and regions (*N*-terminus) as critical determinants in interferon (IFN)- α/β antagonism, RNAi suppression, and *in vivo* attenuation in mammalian cells¹⁷⁴⁻¹⁸⁰.

NS5 protein

NS5 is a multi-functional protein encoding integral methyltransferase and RNA-dependent RNA polymerase (RdRp) functions in flaviviral replication. In particular, the 33 kDa *N*-terminal portion (residues 1-296) encodes the *N*-7 and 2'-*O* methyltransferase domains required in formation of 5' Type-I m⁷GpppAmp capping of viral RNA¹⁸¹⁻¹⁸³; NS5-mediated 2'-*O* methylation has also been linked to antagonism of IFIT-mediated antiviral IFN- α/β induction¹⁸⁴. RdRp activity is encoded in the *C*-terminal portion of the NS5 protein in three canonical domains: palm (aa 499-541 and 610-717), fingers (aa 274-498 and 542-609), and thumb (aa 718-899) demonstrating *trans*-complementation with the NS3 protein in the formation and function of the flavivirus replication complex^{133,185-189}. Further studies have identified the role of the NS5 protein in STAT1/2-linked IFN- β antagonism for some flaviviruses^{190,191}.

“Other” small, hydrophobic NS proteins: NS2A, NS2B, and NS4A

Limited information is available on the functional roles of the encoded NS2B and NS4A proteins. NS2B serves as an essential cofactor for NS3 serine protease function¹⁶²⁻¹⁶⁴; and the NS4A protein has been implicated in vesicle formation via rearrangement of cytoplasmic membranes in concert with its *C*-terminal 2K peptide domain^{192,193}. In contrast, the 22 kDa NS2A protein is a membrane-bound protein with putative roles in virus assembly/release and formation of virus-induced membrane structures¹⁹⁴⁻¹⁹⁶; furthermore, the NS2A protein has also been implicated in Type-I IFN α/β antagonism¹⁹⁷⁻¹⁹⁹. Co-localization of the encoded NS2A, NS2B, NS4A, and NS4B proteins with dsRNA provide membrane-spanning regions facilitating putative assembly and anchoring of the flaviviral replication complex on the ER membrane^{133,146}.

1.3.1.3 UNTRANSLATED REGIONS: 5'/3' TERMINAL SEQUENCES

Both a 96 nt 5'UTR and variable 337-649 nt 3'UTR flank the encoded WNV polyprotein sequence^{133,200}; in addition, despite lineage-specific variation in the initial 200 nt of the 3'UTR, the 5'UTR sequence is strongly conserved between strains^{133,201}. Conserved secondary RNA structures: stemloops (SL), pseudoknots, dumbbells (DB), and hairpins (HP) in both 5' and 3'UTRs provide critical RNA structural elements in flaviviral replication and translation (**Fig. 1-6A**). In particular, conserved base-pairings (*i.e.*, conserved sequences; CS) and stemloops in both the 5'UTR and 3'UTR: SLA and 3'SL (among others) facilitate long-distance, *cis*-complementation of RNA elements resulting in cyclization of the positive-sense RNA strand prior to the initiation of minus-strand RNA synthesis—though not required for translation (**Fig. 1-6B**)^{200,202-206}.

1.3.2 Viral life cycle

Members of the *Flaviviridae: Flavivirus* genus exhibit a viral life cycle conserved among positive-sense (ssRNA), enveloped viruses—WNV is no exception (**Fig. 1-7**)¹³⁴.

Attachment, entry, and fusion

Infectious virions attach to the surface of the host cell and enter via receptor-mediated endocytosis of clathrin-coated pits^{133,207,208}; however, the associated receptor(s) remain disputed for *all* flaviviruses. Possible receptors include the glycosaminoglycan (GAG) heparin-sulfate²⁰⁹ and DC-SIGN (dendritic cell-specific ICAM-grabbing non-integrin)²¹⁰ (among others) with putative binding to the E protein Ig-like EDIII or *N*-linked glycosylation of the M or E proteins^{133,211,212}. However, there is no conclusive evidence or agreement for any flavivirus cell receptor, with GAGs thought to be co-receptors (sticky molecules on the surface of cells) rather than real receptors. Transition from early to intermediate to mature-late stages of virus uptake coincides with the progressive acidification (pH < 6.5) of the delivered endosomes²¹³; E protein dimers on

the virion surface dissociate at low pH with rotation of the domain II “hinge” region away from domain I forming trimeric spikes projected away from the virion-surface^{214,215}. Subsequent exposure of the fusion peptide (loop)—conserved among all flaviviruses—in domain II precedes its insertion into the outer lipid leaflet of the host cell membrane^{216,217}. Domain III then folds back against the exterior surface of the core trimer pulling the membrane-anchored M protein transmembrane domains (TMDs) towards the target endosomal membrane²¹⁵; this conformational change brings the virus envelope close to the host membrane—facilitating fusion^{215,217-219}.

Assembly, packaging, and exocytosis

Fusion of the late endosome with the virion envelope releases the nucleocapsid (NC; composed of C protein) into the host cell cytoplasm resulting in the dissociation of the C protein and vRNA. Translation is initiated using the delivered, capped vRNA producing a single polyprotein cleaved into the structural and nonstructural proteins. Remodeling of the host ER-membrane by virus-encoded NS proteins and formation of the replication complex initiates viral replication and particle assembly¹³⁴. Immature virions accumulate in the ER containing E, prM, nucleocapsid (with bound RNA), and host lipid membrane with export to the Golgi apparatus and *trans*-Golgi network (TGN); prM is cleaved in the TGN via host furin-like proteases generating mature, infectious virus particles²²⁰. Packaged infectious particles are transported to the host cell surface and released via exocytosis¹³⁴.

1.3.3 Viral replication

Following the release of NC-bound viral RNA from mature-late endosomes—as described above (**Fig. 1-7**)—*cis*-complementation of encoded RNA structural elements facilitates cyclization of the flaviviral genome in a “pan-handle” structure required for the initiation of minus-strand RNA synthesis (but not translation)^{202,205,206,216}. In this regard, a

stable hairpin (cHP) in the 5'UTR directs codon selection to the AUG start site in the adjacent SLB stem-loop structure in translation initiation (**Fig. 1-6B**)²²¹. cHP is present in the “linear” form of the RNA genome; in effect, cyclization disrupts this RNA structure allowing *cis*-complementation of the 5'UTR SLB and 3'UTR3'SL elements and NS5 protein binding to initiate minus-strand RNA synthesis^{133,222,223}. Formation of the replication complex occurs on the ER-membrane consisting of cyclized vRNA, NS5 (RdRp), and other virus-encoded and host cellular proteins (*e.g.*, eEF1 α)^{133,224,225}. Indirect roles have been proposed for the NS1, NS2A, NS4A, and NS4B proteins including *in situ* restructuring of the host cell ER-membrane and anchoring of the replication complex^{133,200,226,227}. During transcription, the *N*-terminal region (methyltransferase) of the NS5 protein both caps and modifies the 5'-terminus of the nascent RNA strand protecting it from targeted host cellular protein degradation^{182,183,185,228}. Overall, positive-sense RNA synthesis is >10-fold more efficient than minus-strand synthesis from positive strand templates²²⁹. Furthermore, studies suggest that nascent genomic RNA can serve as a template for translation, transcription, and virion assembly; however, translation must precede transcription, and encapsidation of the NC does not occur until transcription has initiated^{131,133}.

1.4 PATHOGENESIS

1.4.1 Human infection

From mosquito bite to host viremia

Humans are “dead-end” hosts to natural WNV infection with multiple physical and immunological barriers restricting symptomatic disease to $\leq 20\%$ of opportunistic infections. Natural infection is initiated upon the bite of an infected mosquito with the subsequent infection of resident antigen-presenting cells including dendritic cell (DC)

subsets (*i.e.*, Langerhans cells) in the skin^{230,231}. In particular, *Culex* spp. mosquitoes can deliver 10^1 to 10^7 pfu of virus to the bite site²³²⁻²³⁴. Infected DCs migrate to the draining lymph nodes facilitating secondary infection of neutrophils, macrophages, and DC subsets with subsequent systemic viral spread^{235,236}. Serum viremia disseminates to the peripheral organs including both permissive (*e.g.*, spleen) and non-permissive (*e.g.*, liver) tissues resulting in systemic WNV infection²³⁷.

Infiltrating the brain: WNV neuroinvasion

Principal routes-of-infection into the central nervous system (CNS) are undefined; however, four main mechanisms have been proposed²³⁷: **(1)** infection of the endothelium in the blood-brain-barrier (BBB); **(2)** direct viral crossing of the BBB due to vasoactive cytokine-induced (*e.g.*, TNF- α) vascular permeability and/or matrix metalloproteinase (MMP)-mediated breakdown of endothelial tight junctions²³⁸⁻²⁴⁰; **(3)** the “Trojan horse” mechanism with trafficking of infected leukocytes and T cells across the BBB²⁴¹⁻²⁴³; and **(4)** retrograde axonal transport of the virus to the CNS via infection of the peripheral neurons²⁴⁴. Regardless of *how* WNV infiltrates the brain, peripheral mononuclear (PMN) cells (neutrophils) predominate in the cerebrospinal fluid (CSF) of WNND patients with meningitis and encephalitis^{245,246}; furthermore, numerous cell-types support WNV infection in the brain: neurons, cortical astrocytes, and microvascular endothelial cells (among others) with detection of WNV in the cerebral cortex, thalamus, brainstem, basal ganglia, cerebellum, and anterior horn of the spinal cord²⁴⁷⁻²⁴⁹.

1.4.2 Clinical diagnoses

West Nile fever (WNF)

West Nile fever (WNF) is an acute, mild febrile illness with onset 2-14 days following infection which may occur via mosquito bite, blood transfusion, or organ transplantation. Symptoms mimic “flu-like” illness including headache, weakness,

myalgia (muscle pain), and arthralgia (joint pain) with gastrointestinal complications and transient maculopapular rash also reported. Infection is self-limited with <1% of diagnosed cases progressing to severe central nervous system (CNS) disease. No FDA-licensed vaccine exists; thus, treatment is supportive^{48,250}.

West Nile neuroinvasive disease (WNND)

Initial reports of West Nile neuroinvasive disease (WNND) were described during the 1957 Israel outbreak^{250,251}. Patients develop acute aseptic meningitis, encephalitis, or meningoencephalitis resulting in death in ~10% of diagnoses. Based on historical data, 40% of WNND diagnosed patients presented meningitis, 16% encephalitis, and 44% meningoencephalitis in the 1996 Romanian outbreak; similar statistics were observed in the 1999 New York outbreak—meningitis (40%) and encephalitis (60%)^{250,252,253}. Clinical presentation of WNND-related meningitis includes (but is not limited to) fever, stiff neck, headache, photophobia, and cerebrospinal fluid (CSF) pleocytosis (increased cell counts). Furthermore, encephalitic patients tend to exhibit additional motor dysfunctions: tremors, Guillan-Barre-like syndrome, and/or acute flaccid paralysis (AFP) with evidence of parenchymal involvement affecting diverse regions of the brain including the cerebral cortex, basal ganglia, brainstem, and anterior horn of the spinal cord^{250,253,254}. Similar to WNF diagnoses, treatment is limited to supportive care. Long-term sequelae include fatigue, cognitive dysfunctions, and AFP—with one-third of patients with AFP recovering to near base-line while one-third exhibit little to no improvement up to 6 months post-onset^{255,256}. Risk factors for WNND include diabetes, alcohol abuse, high blood pressure, and cardiovascular disease with increased propensities in the aged (≥ 50 years) and immunocompromised^{253,257,258}.

Persistent WNV infection

Persistent renal infection in WNV survivors has been implicated in the prolonged sequelae and chronic symptoms of WNV patients²⁵⁹. Recent *in vivo* studies in hamster and murine models have also exhibited persistent CNS and peripheral (*e.g.*, kidney or spleen) infection with prolonged viremia in the hamster model^{253,260-262}. Further research is needed to confirm the epidemiological significance of these studies.

1.4.3 Innate immune responses

Extensive *in vitro* and *in vivo* studies have implicated the roles of both innate and adaptive (*i.e.*, B and T cell-mediated) immune responses in the clearance, control, and progression of WNV infection in humans. In the specific aims of this dissertation, critical components of host innate antiviral responses are reviewed below:

Interferon (IFN) expression

Identification of an “interfering” component in virus-containing cell culture—termed interferon (IFN)—facilitated induction of an antiviral state that exhibited a protective effect against virus infection²⁶³⁻²⁶⁶. IFN is part of the host innate immune response to infection; however, viruses encode proteins that act to antagonize and overcome the IFN response. Induction of Type-I IFN (IFN- α/β) has been linked to peripheral and CNS protection in acute WNV infection^{267,268}. In-depth *in vitro* and *in vivo* studies have investigated mechanisms driving viral evasion and suppression of Type-I IFN antiviral responses. In brief, autocrine and paracrine signaling of IFN- β stimulates the extracellular IFN- α/β receptor (IFNAR) complex resulting in Tyk2-Jak1 phosphorylation of STAT1 and STAT2^{199,269,270}. Subsequent formation of the trimeric IFN-stimulated gene factor 3 (ISGF3) with IFN regulatory factor (IRF)-9 translocates to the nucleus resulting in up-regulated transcription of ISGs (**Fig 1-8**): protein kinase R (PKR), IFN-inducible tetratricopeptides (IFITs), IFN-induced transmembrane (IFITM)

genes, and 2'-5'-oligoadenylate synthase (OAS) in addition to *IFNA* (IFN- α) (and others)²³⁷. IFN- β -stimulated expression of these genes further induces a cellular antiviral state with the suppression of viral replication/translation and degradation of “non-self” viral RNA and virus-encoded proteins^{184,271,272}; *how* WNV evades these responses is a topic of continued debate and extensive research including (but not limited to) initiation and/or modulation of RIG-I-, Toll-, and NOD-like receptor signaling^{237,273}.

RIG-I-like receptors (RLRs)

Retinoic-acid inducible gene-I (RIG-I)-like receptors (RLRs) constitute cytosolic RNA helicase proteins: RIG-I (DDX58), myeloma differentiation antigen 5 (MDA-5; IFIH1), and LGP2 that bind “non-self” RNA elements—including 5'-triphosphorylated (5'ppp) ends²⁷⁴. Stimulation triggers RIG-I and MDA-5 binding (in particular) of the mitochondrial antiviral signaling (MAVS) adaptor activating IRF-3, IRF-5, or IRF-7 transcription of target genes leading to up-regulated IFN- β expression and induction of an innate antiviral state (**Fig. 1-8**)²⁷⁵⁻²⁷⁷; recent studies also confirmed the critical, cooperative roles of RIG-I, MDA-5, and MAVS in innate control of WNV infection with viral RNA-stimulated RIG-I activation early in infection and MDA-5 at later time-points²⁷⁸. However, both RLRs also exhibit non-canonical IPS-1 signaling in acute infection controlling WNV-induced pro-inflammatory responses (*e.g.*, Type-I IFN production) linking innate and adaptive immunity²⁷⁹⁻²⁸¹. In addition, *in vitro* studies correlated RIG-I stimulation to the expression of PAMPs in both genomic and anti-genomic RNA fragments in the 5'UTR, NS2A, and NS4A; however, these responses were masked in the context of full-length viral RNAs warranting further investigation²⁸².

Toll-like receptors (TLRs)

Toll-like receptors (TLRs): TLR3, TLR7, and TLR8 (in particular) are located in intracellular endosomal vesicles and recognize distinct viral PAMPs including dsRNA

(TLR3) and “non-self” GU-rich ssRNA motifs (TLR7/8). Recognition of dsRNA triggers a conformational change activating the TIR domain-containing adaptor inducing IFN- β (TRIF) adaptor protein which binds IRF-3, IRF-7, or NF- κ B stimulating translocation to the nucleus and *IFNB* (IFN- β) and pro-inflammatory cytokine/chemokine transcription; in contrast, both TLR7/8 signal through the myeloid differentiation factor 88 (Myd88) adaptor protein and downstream IRF-7- and NF- κ B-mediated transcription (**Fig. 1-8**)^{237,273,283}. However, unlike RLRs, antiviral TLR signaling is both redundant and variable, functioning in a cell- and tissue-specific manner. For example, TLR3 signaling stimulates Type-I IFN production and controls WNV replication in cortical neurons but not macrophages or DCs; however, the role of TLR3 in mediating viral entry and infection of the CNS remains controversial^{242,284}. In particular, NS1 expression has been implicated in TLR3 and IFN- β antagonism²⁸⁵; however, additional NS1 studies failed to disrupt TLR3 signaling²⁸⁶. TLR7 and Myd88 signaling have been implicated in the recruitment of immune cells: macrophages, neutrophils, and T cells to the CNS—clearing WNV CNS infection—and Langerhans DC cell migration to the draining lymph nodes^{287,288}.

NOD-like receptors

Nuclear organization domain (NOD)-like receptors involved in antiviral responses include the cytosolic NLRP3 inflammasome complex comprised of multiple pattern recognition receptors (PRRs) regulating (1) cleavage of pro-IL-1 β and (2) programmed cell death responses (*i.e.*, pyroptosis)²⁷³. In brief, activation of NLRP3 stimulates caspase-1 cleavage of pro-IL-1 β into the mature IL-1 β form leading to both autocrine and paracrine activation of the IL-1 receptor and NF- κ B-regulated pro-inflammatory cytokine and ISG expression (**Fig. 1-8**)^{237,273}. In this regard, up-regulated IL-1 β expression has been detected in WNV-infected macrophages and neural tissue (*e.g.*, glial cells and cortical neurons); in particular, IL-1 β synergizes with Type-I IFN responses to suppress viral replication in cortical neurons promoting viral-clearance in the CNS²⁸⁹⁻²⁹¹. In

addition, NLR-induced IL-1 β directs peripheral WNV-infected Langerhans DC cell migration to the draining node²⁹².

Role of the NS4B protein: innate immune evasion

Based on these studies (and others), extensive research has been conducted to dissect *how* Flaviviruses interact with and modulate host cellular and immune responses in the establishment of peripheral, neuroinvasive, and lethal infection. In particular, WNV-induced Type-I IFN α/β antagonism has been linked to specific, virus-encoded molecular determinants in the NS2A, NS4B, and NS5 proteins (among others). In relation to the aims of this dissertation, the *N*-terminus of the NS4B protein has been implicated in flaviviral antagonism of Jak-STAT signaling in Type-I IFN- α/β expression^{175,177,199}; additional studies have also identified the role of the NS4B protein in innate RNAi suppression, unfolded protein responses (UPRs), and cell-specific pro-inflammatory cytokine/chemokine induction: IL-6, IL-8, and IP-10 (and others)^{176,293,294}. In-depth, *in vivo* studies of the attenuated NS4B-P38G/T116I (+NS3-N480H) WNV mutant have also confirmed NS4B-linked (cell-specific) induction of Type-I IFN- α/β and Myd88-dependent IL-1 β , IL-6, and TNF- α cytokine expression with protection against lethal infection in the TLR7^{-/-} and Myd88^{-/-} mouse models^{178,295}.

1.5 GAPS IN KNOWLEDGE

West Nile virus is a prime example for arboviral evolution in the U.S. Following its initial introduction into the western hemisphere in the 1999 New York epizootic outbreak, subsequent transmission within domestic bird and ornithophilic (bird-seeking; *Culex* spp.) mosquito populations fueled its rapid expansion westward across the continental U.S., north into Canada, and south into Central/South America^{43,49-53}. In its wake, the rising burden on the U.S. public health infrastructure cultivated the establishment of several regional and state-wide WNV surveillance programs in CA, CT, IL, NY, and TX tasked with identifying the underlying ecological determinants driving now endemic WNV circulation in the U.S. In collaboration with these efforts, paired *in silico* and phenotypic studies identified the emergence, extinction, and co-circulation of multiple U.S. genotypes linked to conserved molecular markers in the WNV genome^{42,55,58,59,63,66}; however, despite sustained efforts, limited progress has been made dissecting *how* and *when* viral outbreaks occur. WNV continues to evolve in the U.S.; it's not a question of "if" but "when" a significant outbreak will re-occur. Based on these scientific deficits, this dissertation is aimed at elucidating novel genotypic and molecular determinants driving the continued evolution of WNV in the U.S.

1.6 SPECIFIC AIMS

Prior applications of isolates collected in Harris Co., TX have provided a robust geographic model for both regional and national WNV surveillance. In this dissertation, applied *in silico* use of this model will enable elucidation of critical genotypic and genetic determinants driving the regional selection of WNV-encoded mutations and variants. In particular, despite the confirmed emergence, extinction, and co-circulation of multiple WNV genotypes in the southwestern U.S., limited evidence of WNV circulation has been reported in northern Mexico^{53,296-299}. Based on these trends, *how* does viral movement (gene-flow) in host/vector populations on the U.S.-Mexican border influence or constrain WNV evolution in the southwestern U.S.? In addition, national WNV incidence deviated from the 2006-2011 *status quo* in the recent 2012 WNV epizootic outbreak focused in the Dallas/Fort Worth, Texas metropolitan region^{47,48}. *In silico* application of novel 2005-2012 isolates from both northern Mexico and central Texas isolates will be used to map conserved genetic and genotypic determinants underlying the adaptation and evolution of natural, circulating WNV isolates in the southwestern U.S. and periodic epidemics.

Based on these and prior *in silico* studies, it remains unclear whether dominant or minor, allelic variants in natural WNV populations influence national epidemiological trends or infected host phenotype; in effect, further *in silico* application of the Harris Co. surveillance model using state-of-the-art next generation sequencing (NGS) technologies will enable the elucidation of viral quasispecies structure as a biological driver in host-selection and *in vivo* phenotype in natural WNV populations. From an applied perspective, identification of natural, genetic determinants linked to attenuated *in vitro* or *in vivo* viral phenotypes can be essential in the development of live-attenuated vaccine candidates or antiviral therapeutic targets—neither of which exist for WNV^{178,179,300,301}. In this regard, there is limited understanding of viral determinants involved in the phenotypic properties of WNV. Therefore, the NS4B protein will be used as a model to

investigate phenotypic properties due to several identified natural or *in vitro*-selected variants as the overall immunological role of the NS4B protein in mediating these responses remains unclear. Linking differential innate antiviral responses to encoded NS4B residues or domains using an established reverse-genetics approach—as proposed in this dissertation—will define critical molecular determinants influencing the development of observed host phenotypes in acute WNV infection. Periodic local, regional, and national WNV epidemics continue to persist in the U.S. despite endemic circulation with no signs of declining; thus, bridging *in silico* surveillance and natural, genetic flux in viral populations with concomitant phenotypic trends is imperative to better understand underlying determinants driving the continued evolution of WNV in the U.S. and the impact of this evolution on future human health. Therefore, the following specific aims are proposed:

1.6.1 Specific aim 1(a): Chapter 3

Elucidate WNV transmission dynamics on the U.S.-Mexican border using 2005-2010 isolates collected in El Paso, Texas and Ciudad Juarez, Mexico.

Hypothesis: *In silico* sequencing and phylogenetic analyses will confirm distinct genetic determinants and genotypic trends between viruses isolated in either location during different years consistent with published sequences originating from either northern Mexico or the southwestern U.S.

Rationale: Host/vector-surveillance efforts in northern Mexico from 2002-2010 identified wide-spread WNV seroprevalence in horses despite limited incidence of clinical human disease. Paradigmatic models for WNV in North America support the unidirectional transmission of WNV into northern Mexico but not vice-versa.

1.6.2 Specific aim 1(b): Chapter 4

Investigate genotypic and geospatial determinants driving the 2012 Texas WNV outbreak in 2010-2012 isolates collected in Harris Co. and Dallas/Fort Worth, Texas.

Hypothesis: *Implementation of in silico comparisons in novel, full-length WNV sequences collected during the 2012 WNV outbreak will reveal the genetic, genotypic, and spatio-temporal determinants underlying the emergence of periodic viral epidemics.*

Rationale: Extensive host/vector-surveillance efforts in Harris Co., Texas from 2002-2009 confirmed the emergence, extinction, and co-circulation of multiple U.S. genotypes: NA/WN02, coastal SE Texas, and SW/WN03 as an established geographic model for WNV evolution in the U.S.^{55,58,63}. Furthermore, prior *in silico* studies identified a relative genetic stasis in WNV evolution from 2006-2011 consistent with the detection of homogeneous viral populations in multiple, regional surveillance cohorts and a lull in national clinical incidence^{47,60-62,68,69}. In-depth *in silico* comparisons of full-length 2010-2012 Harris Co. and Dallas/Fort Worth, Texas sequences will test the hypothesis that the 2012 outbreak resulted from the emergence of novel, conserved genetic determinants.

1.6.3 Specific aim 2: Chapter 5

Determine the role of quasispecies structure in both temporal- and host-dependent selection of natural WNV populations conferring *in vivo* phenotype variation.

Hypothesis: *Paired application of Shannon entropic indices and SNV detection on “deep” sequenced 2002-2012 Harris Co. isolates will confirm distinct host- and phenotype-dependent determinants in natural WNV population structure underlying the selection of dominant viral sequences over time.*

Rationale: *In vivo* fitness studies with laboratory-derived, sub-clonal WNV populations exhibit host-dependent selection pressures with population expansion in

mosquitoes and subsequent purifying selection in the avian host despite no apparent fitness costs^{116,119,120,122,302}. Furthermore, host fitness is a proposed correlate of viral quasispecies structure; in effect, diverse populations are better equipped to adapt to host-selection pressures^{117,119,302-304}. Based on the results from Specific Aim 1b, inclusion of a diverse pool of established 2002-2012 Harris Co. isolates (with defined *in vivo* phenotypes) in the Illumina HiSeq1000 NGS pipe-line will test the hypothesis that natural WNV population structure is host-dependent and a biological determinant of *in vivo* virulence.

1.6.4 Specific aim 3: Chapter 6

Identification of critical molecular determinants in the NS4B protein involved in modulation of host innate antiviral responses.

Hypothesis: *In vitro* expression of natural and infectious clone-derived NS4B mutations will exhibit both residue- and domain-dependent dysregulation of NS4B-linked antagonism of innate antiviral responses.

Rationale: *In vivo* studies have demonstrated the role of the flavivirus NS4B N-terminus in the inhibition of Jak-STAT signaling and Type-I IFN- α/β expression^{175,177,270}. Further studies have confirmed altered *in vitro* or *in vivo* phenotypes restricted to natural and engineered NS4B substitutions (*e.g.*, P38G/T116I, C102S, and E249G)^{174,178,180}. Based on these studies, integration of human alveolar A549 cell culture with the established NY99 infectious clone (NY99ic) reverse-genetics platform will provide an IFN-competent *in vitro* model to elucidate the impact of specific NS4B mutations and domain-dependent trends in Type-I IFN- β antagonism of host innate antiviral responses.

Table 1-1: Evolution of WNV lineage 1a—conserved amino acid substitutions

Lineage	Cluster	Genotype	E			NS1			2A	2B	NS3		4A	4B	NS5		
			126	159	291	70	99	206	224	103	175	249	85	11	274	314	898
			T	I	K	A	S	L	A	A	I	P	V	S	S	K	T
1a	1	-	T	.	N	.	.	.
	2	Eastern European	.	M	.	S	.	F	.	V	I
		Mediterranean	.	M	.	S	T/P
	3	-	P	.	T	.	.	.	T
	4	-	I	V	.	*	.	.	.	V	.	.	A
		NY99	I	V	.	*	.	.	.	V	.	.	A	.	.	.	*
		SE Coastal Texas	I	V	.	*	.	.	.	V	.	.	A
		NA/WN02	I	A	R	*	.	.	.	V	.	.	A	.	.	*	.
		SW/WN03	I	A	R	*	.	.	.	V	.	.	T	.	.	R	.
		MW/WN06	I	A	R	*	.	.	.	V	.	.	A/I	.	.	*	.
	5	-	V	.	.	.	T	.	.
	6	-
1b	-	-	A
2	-	-	.	.	*	.	A	H	.	N	.	.	.

E, envelope, NS, nonstructural; 4A, NS4A; 4B, NS4B. Amino acid changes in the WNV genome (shaded text) relative to the consensus sequence which define specific clusters or genotypes in lineages 1: clades 1a or 1b or lineage 2. Dots indicate no difference from consensus (top). Asterisks () indicate the presence of the indicated amino acid change in some but not all isolates.

Figure 1-1: Simplified, condensed distribution of global WNV isolates in a Neighbor-joining tree clustered based on phylogenetic lineage (1-6), clade in lineage 1 (1a or 1b), or cluster in lineage 1a (1-6). U.S. isolates are indicated in red boldface. Scale bar, branch length representing % nucleotide divergence.

[Modified from: **Mann, B.R., et al.** (2013) *Int J Environ Res Public Health*]¹⁴

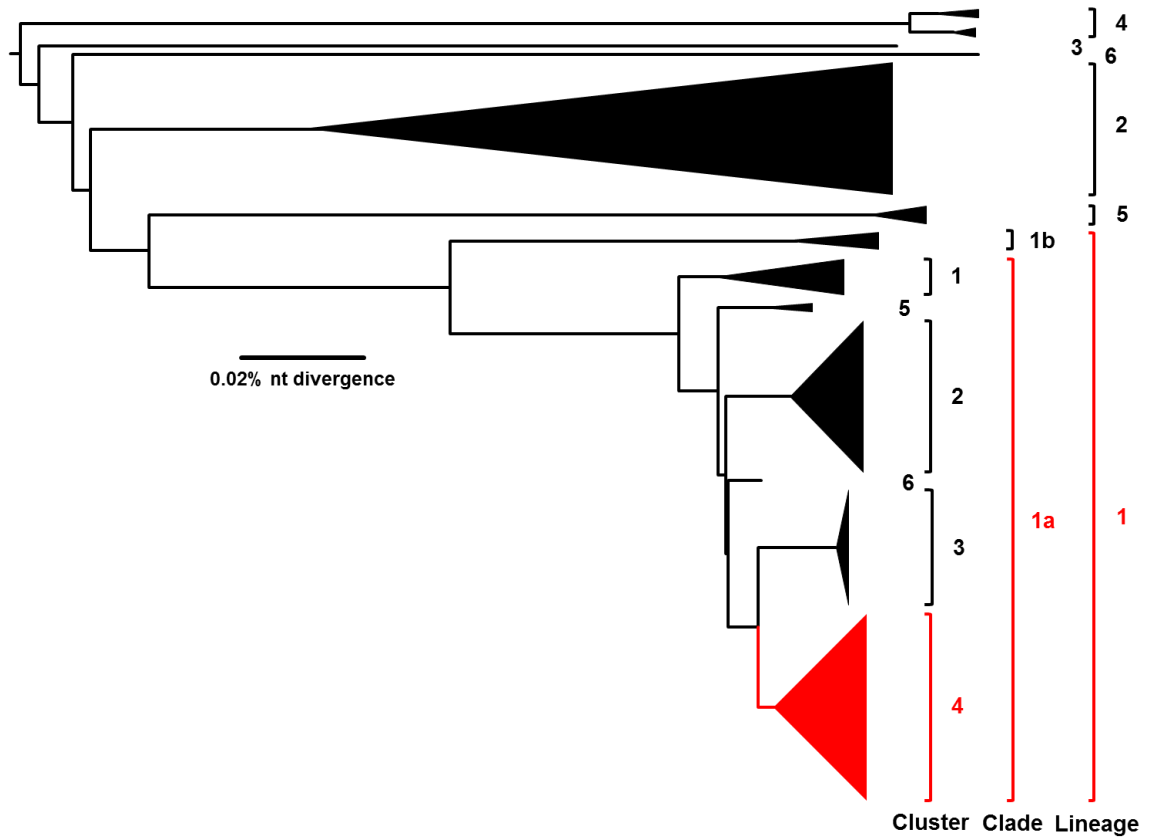


Figure 1-2: Bayesian phylogenetic tree depicting the simplified distribution of the North American genotypes within lineage 1a cluster 4. Posterior probabilities indicated ≥ 0.90 . Scale bar, divergence time in years.

[Modified from: **Mann, B.R., et al. (2013) *Int J Environ Res Public Health***]¹⁴

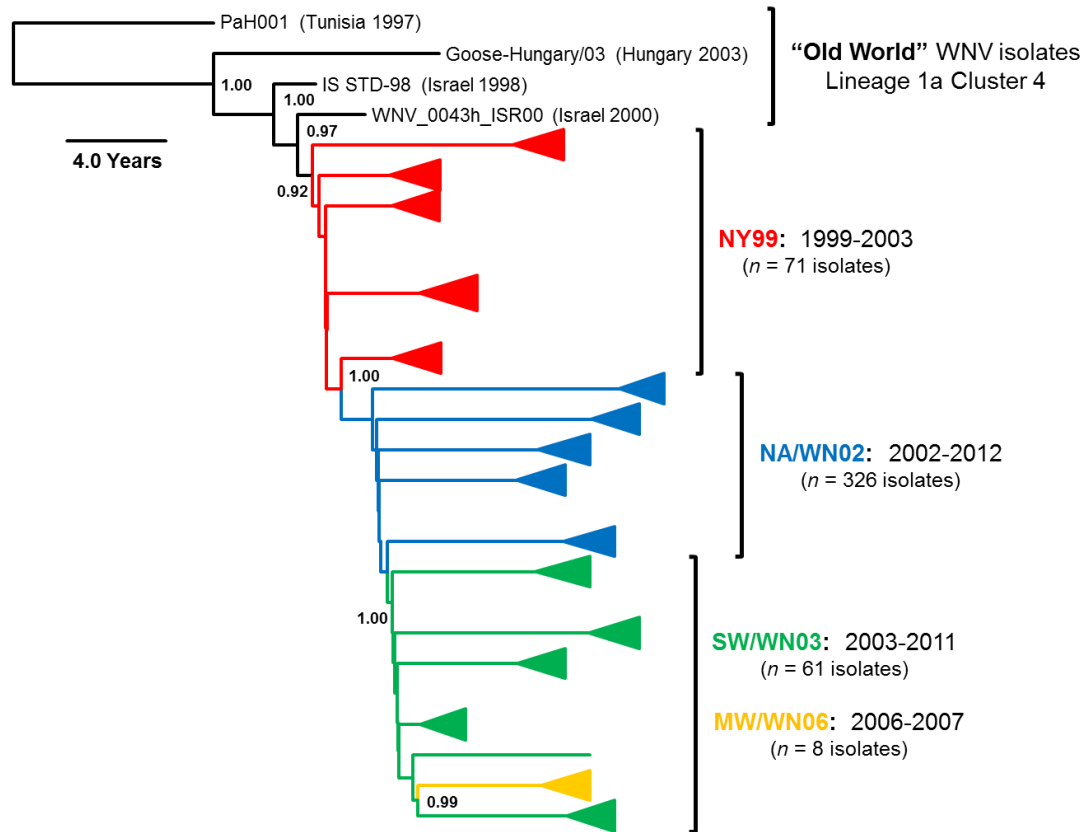


Figure 1-3: Natural WNV enzootic transmission cycle

[Modified from: CDC (2014) <http://www.cdc.gov/westnile/transmission/>]⁴⁸

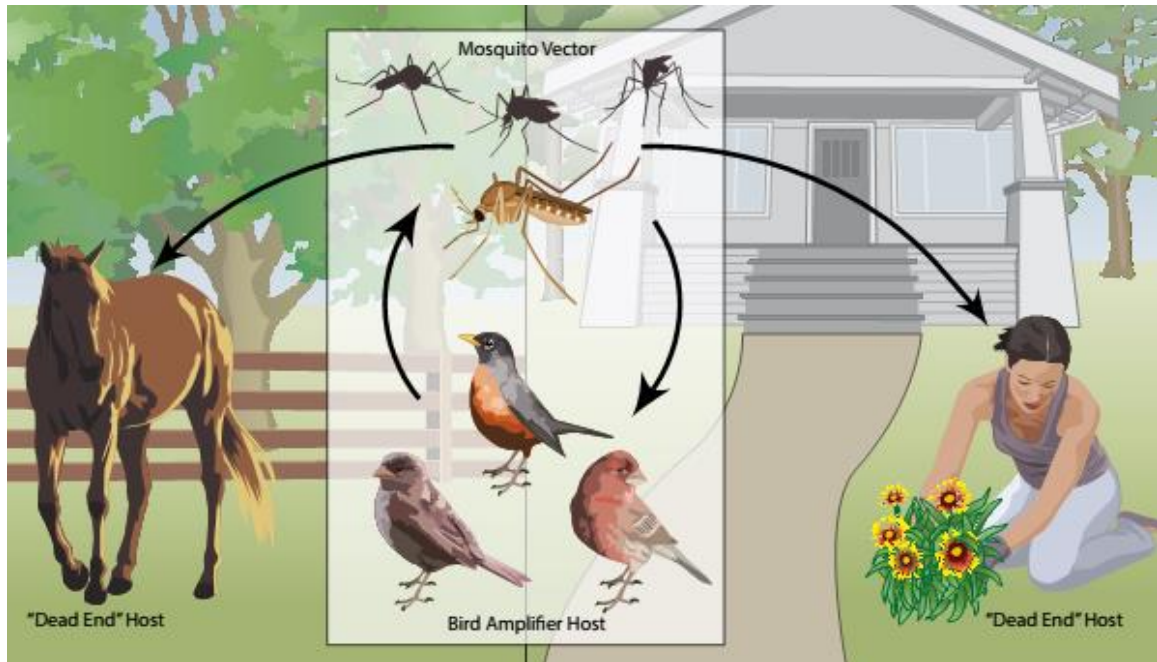


Figure 1-4: National Incidence of clinical WNV disease in the U.S., 1999-2013. Number of death indicated per year (boldface red). WNF, West Nile fever; WNND, West Nile neuroinvasive disease

[Epidemiological data derived from **CDC** (2014) <http://www.cdc.gov/westnile/>]⁴⁸

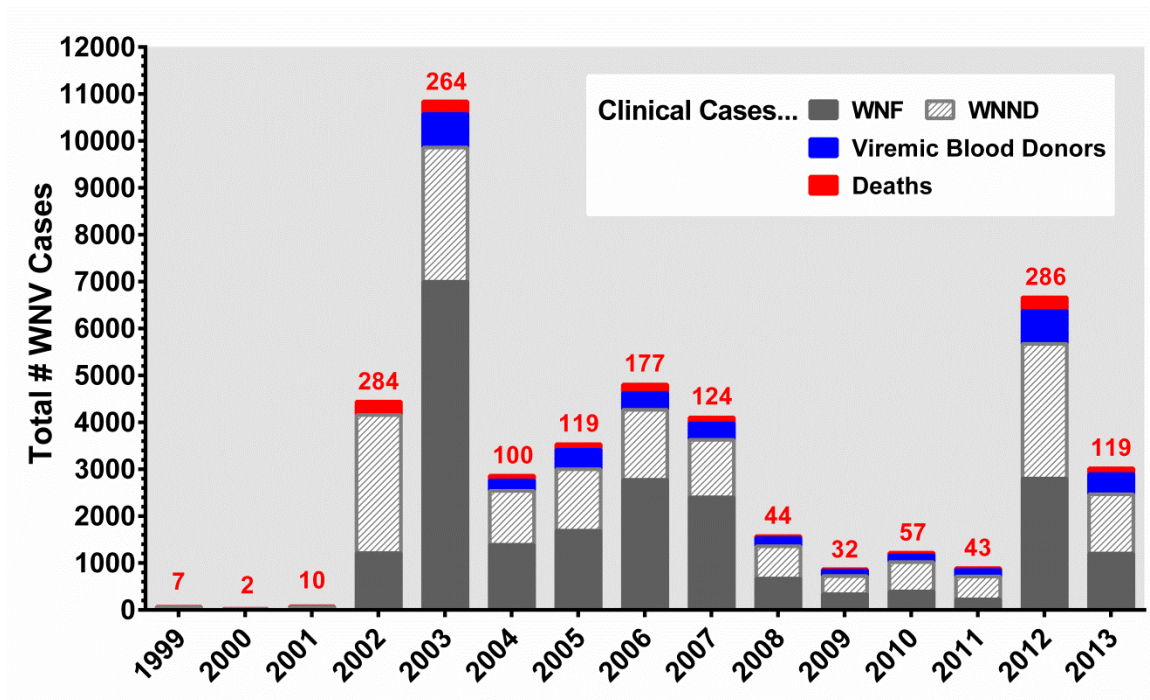


Figure 1-5: West Nile virus genome organization and processing.

[Adapted from: **Brinton, M.A.** (2002) *Ann Rev Microbiol* and **Perera and Kuhn** (2008) *Curr Opin Microbiol*^{131,132}

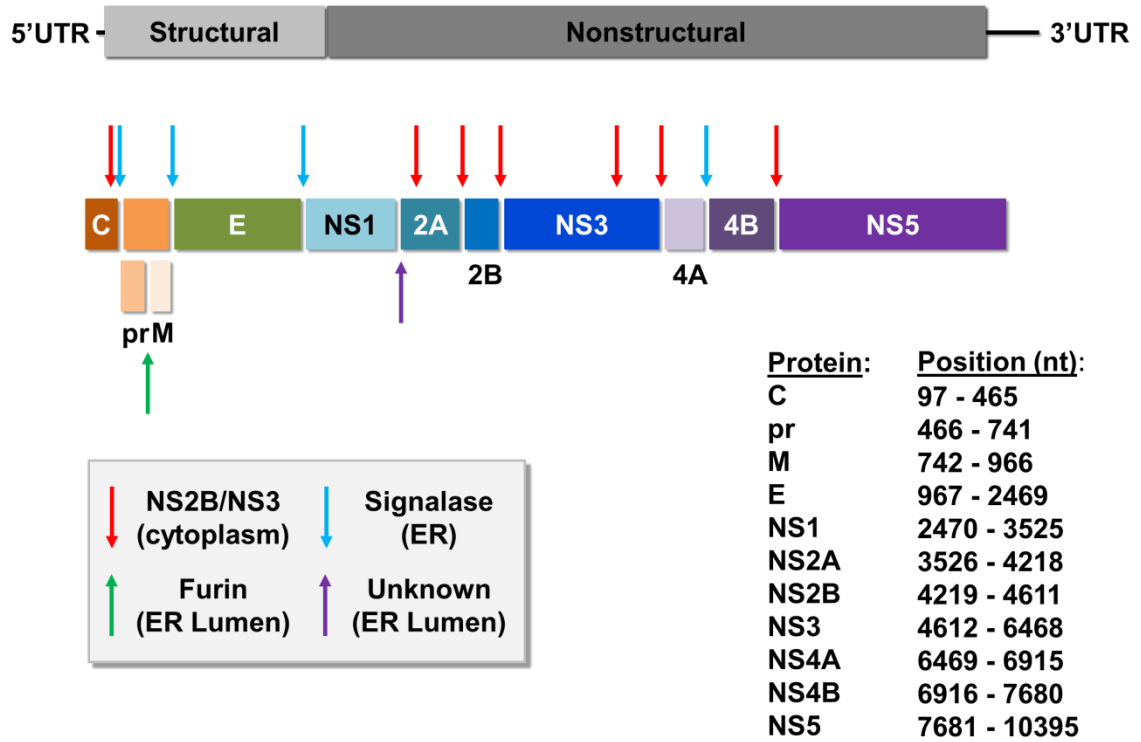


Figure 1-6: 5' and 3' untranslated region (UTR) structure and organization in the WNV genome, related to **Fig. 1-5**. Both **A)** conserved secondary RNA structures and **B)** long-distance base-pairings of RNA elements are indicated.

[Modified from: **Brinton, M.A. (2014) *Viruses***¹³³

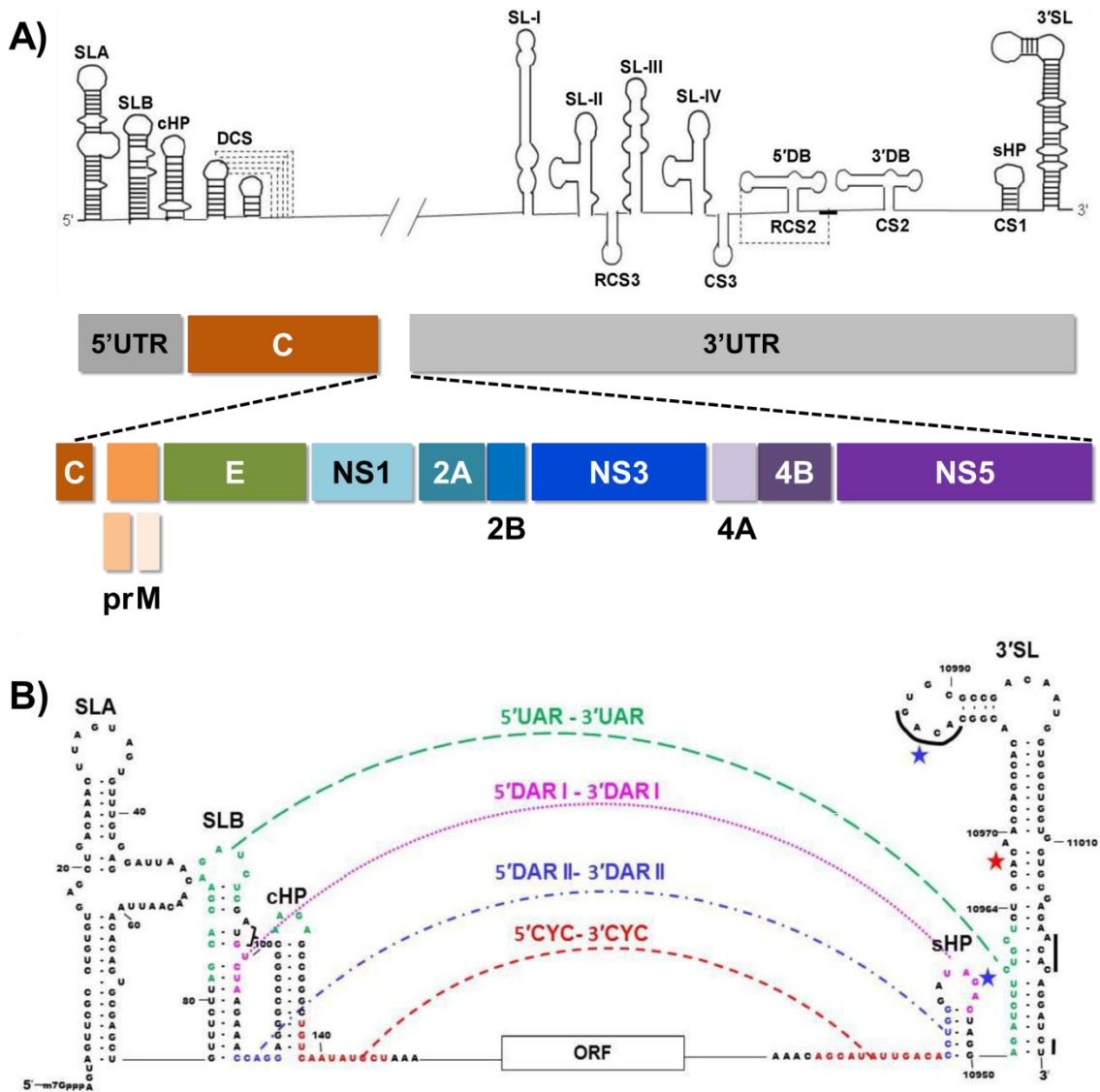


Figure 1-7: West Nile virus life cycle

[Modified from: **Suthar, M.S., et al. (2013) *Nat Rev Microbiol***²³⁷

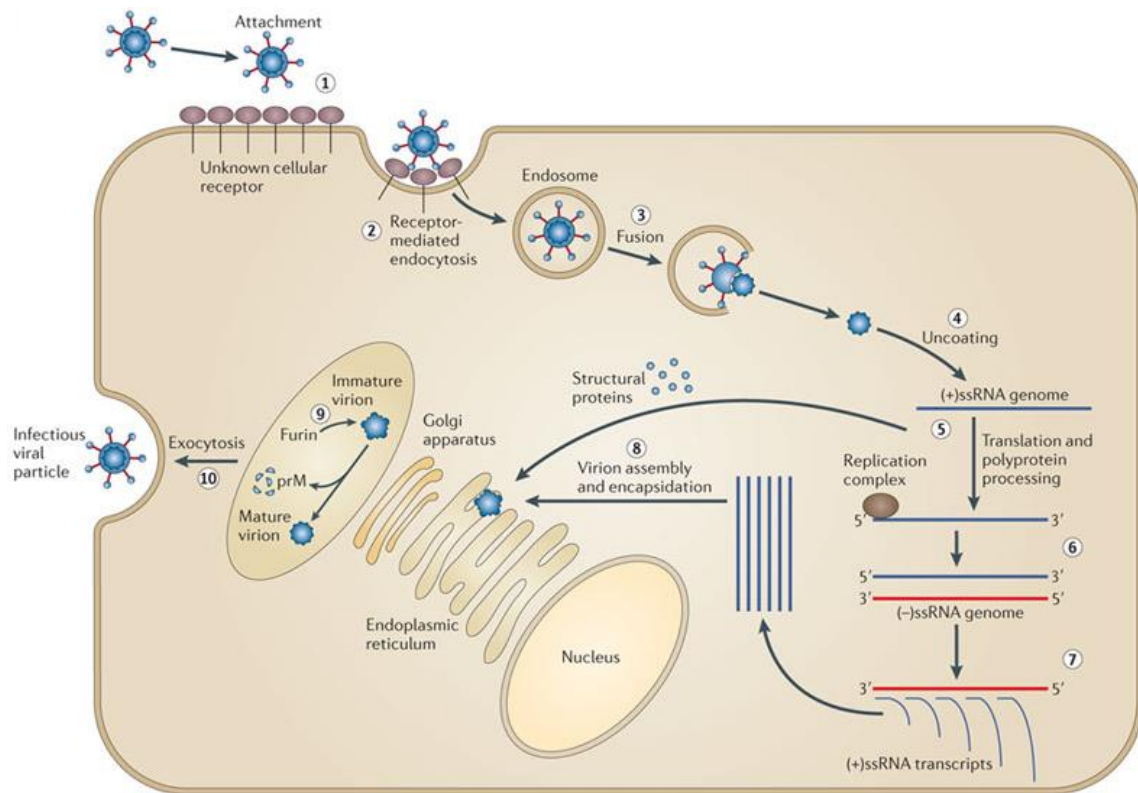
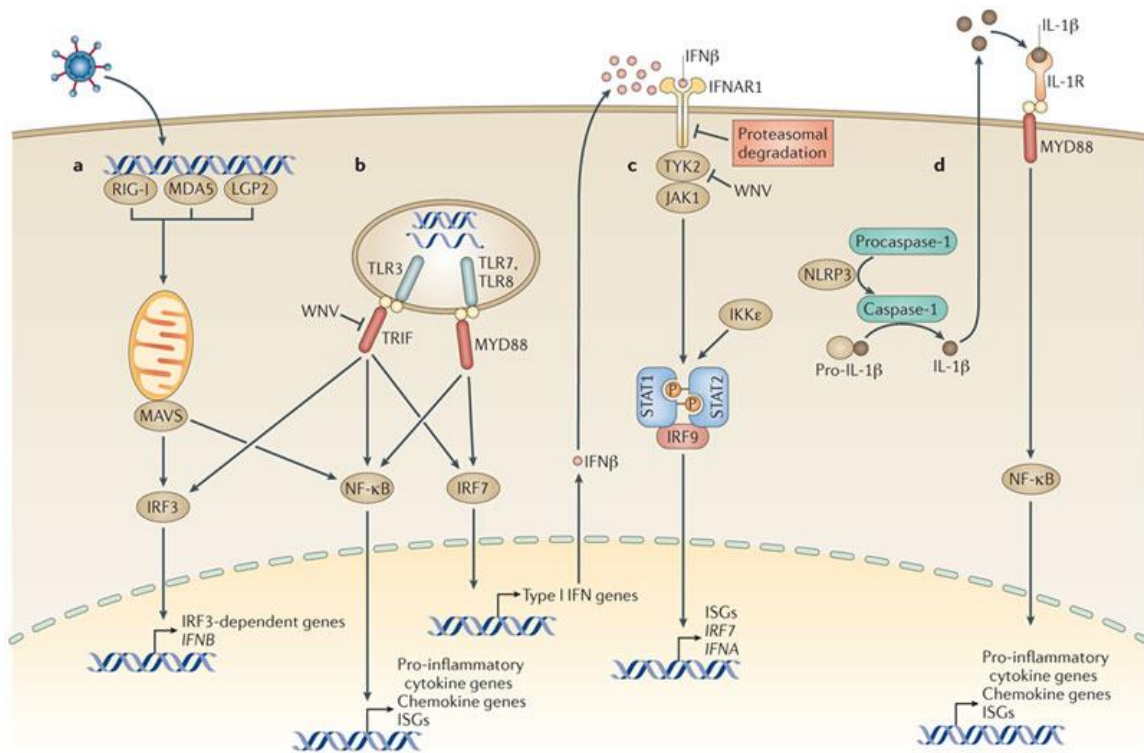


Figure 1-8: Innate immune responses to WNV infection

[Modified from **Suthar, M.S., et al. (2013) *Nat Rev Microbiol***]²³⁷



CHAPTER 2

Materials & Methods

2.1 BUFFERS, MEDIA, AND REAGENTS

2.1.1 Buffers

2.1.1.1 BUFFER RLT (QIAGEN RNeasy Mini Kit)

10 μ L β -mercaptoethanol added per 1mL buffer concentrate (prior to use)

2.1.2 Media

2.1.2.1 A549 AND VERO CELL CULTURE GROWTH MEDIA

MEM (L-glutamine, Earle's salts, Phenol Red; Gibco®)

8% (v/v) Bovine growth serum (BGS, heat-inactivated; HyClone™) **OR**

8% (v/v) Fetal bovine growth serum (FBS, heat-inactivated; HyClone™)

*2% BGS or FBS in “maintenance media”

2 mM L-Glutamine (200mM; Gibco®)

100 nM Non-essential amino acids (NEAA) (100x, 10 mM; Gibco®)

100 Units(U)/mL Penicillin (10,000 U/mL; Pen Strep, Gibco®)

100 μ g/mL Streptomycin (10,000 μ g/mL)

2.1.2.2 C6/36 CELL CULTURE GROWTH MEDIA

Minimal essential media (L-glutamine, Earle's salts, Phenol Red; Gibco®)

10% (v/v) FBS (heat-inactivated; HyClone™)

*2% FBS in “maintenance media”

0.1x Tryptose phosphate buffer (2x reagent stock)

100 nM Non-essential amino acids (NEAA) (100x, 10 mM; Gibco®)

1 mM Sodium pyruvate (100 mM; Gibco®)

100 Units(U)/mL Penicillin (10,000 U/mL; Pen Strep, Gibco®)

100 µg/mL Streptomycin (10,000 µg/mL)

2.1.2.3 2X MINIMAL ESSENTIAL MEDIA (MEM)

200 mL 10x MEM with Earle's Balanced salts (Sigma-Aldrich)

40 mL BGS or FBS (heat-inactivated; HyClone™)

25 mL 7.5% (w/v) Sodium bicarbonate (Gibco®)

5 mL 100x Non-essential amino acids (NEAA) (10 mM; Gibco®)

5 mL 10,000 U/mL or µg/mL Pen Strep (Gibco®)

Autoclaved de-ionized water (to 500 mL total volume)

2.1.2.4 1X MINIMAL ESSENTIAL MEDIA (MEM) AGAR

Equivalent volumes (*e.g.*, 250 mL each for a total 500 mL volume):

2% (w/v) Agar (autoclaved; Sigma-Aldrich)

2x (v/v) MEM (stored at 4°C)

[Optional] 1-3% (v/v) Neutral Red (≥90% dye content; Sigma-Aldrich)

2.1.3 Reagents

2.1.3.1 TRYPTOSE PHOSPHATE BUFFER

5.9% (w/v) Tryptose phosphate broth (TPB; Sigma-Aldrich)

Autoclaved de-ionized water (to 500 mL total volume)

2.2 CELL CULTURE METHODS

2.2.1 Cell-lines

Aedes (Ae.) albopictus mosquito larval (C6/36) cells (ATCC #CRL-1660)

Human alveolar basal Type II epithelial (A549) cells (ATCC #CCL-185)^{305,306}

Mammalian (*Cercopithecus aethiops*) kidney (Vero) cells (ATCC #CCL-81)

2.2.2 Cell culture protocols

Cell-lines were passaged in 8-10% BGS or FBS MEM growth media and re-seeded at 95% confluence until a maximum passage (P) from the ATCC seed stock: C6/36 (P < 30), A549 (P < 15), and Vero (P < 30). Both A549 and Vero cell-lines were maintained at 37°C, 5% CO₂ conditions with C6/36 cells incubated at 28°C (no CO₂).

2.3 WEST NILE VIRUS INOCULATION AND *IN VITRO* CELL CULTURE PASSAGE

2.3.1 General protocol

Inoculation of 95% confluent C3/36 or Vero cell-cultures with 100-200 µL virus seed stock (or prior serial passage depending on the isolate) in 2% BGS or FBS MEM maintenance media produced amplified working stocks harvested 2-4 dpi—upon onset of cytopathic effects (CPE)—in Vero cell culture or 3-4 dpi in C6/36 cell culture (no CPE). Both C6/36- and Vero-inoculated cell cultures were incubated at 28°C (no CO₂) and 37°C, 5% CO₂ conditions, respectively, in designated BSL3 facilities.

2.3.2 *In vitro* A549 passage: qRT-PCR and Bio-Plex® studies

Prior to the described qRT-PCR and Bio-Plex® studies, 95% confluent A549 cell cultures seeded in 6-well or 12-well ($6.9\text{-}8.5 \times 10^4$ cell/cm²) culture plates were inoculated with wild-type or NY99ic-derived viruses (indicated in **Tables 6-1** and **6-3**) at an MOI of 0.1 and incubated in 2% FBS MEM maintenance media at 37°C, 5% CO₂ for 6, 12, 24, and/or 36 hpi.

2.4 DETERMINATION OF VIRAL TITERS AND *IN VITRO* WNV PHENOTYPE

2.4.1 Plaque titration: Vero cell culture

In order to enumerate WNV infectivity titers, applied virus stocks or harvested samples were serially diluted 10^{-1} to 10^{-6} in PBS prior to the inoculation of individual wells in a 12-well culture plate—seeded with 95% confluent Vero cells—with 100 μ L of each dilute inoculum. Plates were then incubated at room temperature for 30 min. to 1 hour before overlaying with 2 mL 1x MEM agar. Following incubation at 37°C, 5% CO₂ for 2 dpi, all wells were overlaid with an additional 1 mL 1x MEM agar containing 1-3% Neutral Red. Plaques were then quantified following incubation for 24-48 hpi with a 100 pfu/mL limit-of-detection.

2.4.2 Temperature-sensitive (*ts*) *in vitro* phenotype

Based on standard Vero cell plaque titration (indicated above), inoculated 1x MEM agar overlaid 95% Vero cell cultures were incubated at either 37°C or 41°C for 48 hpi prior to staining with Neutral Red. Plaques were quantified 72 hpi with determination of viral titers at both 41°C and 37°C. Detected reductions in WNV titers (Δ) $>2.0 \log_{10}$ pfu/mL at 41°C versus 37°C defined a temperature-sensitive (*ts*) *in vitro* phenotype^{300,301}.

2.4.3 *In vitro* plaque morphologies

Plaque morphologies: small (*sp*, < 1.0 mm), medium (*mp*, 1.0-3.0 mm), and large (*lp*, > 3.0 mm) were determined at 72 hpi using the standard Vero cell plaque titration technique incubated at 37°C as described in prior, published studies^{57,174,179,180,295}.

2.5 SANGER SEQUENCING

2.5.1 Preparation of cDNA for sequencing

RNA was extracted from applied WNV stocks using the QIAamp Viral RNA Mini kit (Qiagen) following the 30 min. inactivation of infectious virus in AVL Buffer. Purified RNA suspensions were either processed immediately or stored at -20°C. Implementation of a “touch-down” RT-PCR protocol utilizing the Titan One-Tube RT-PCR system (Roche) and over-lapping primer sets (listed in **Table 2-1**) amplified cDNA fragments spanning the entire WNV genome (**Table 2-2**).

Titan One-Tube RT-PCR reaction mix (Roche):

1 µL Deoxynucleotide (dNTP) solution (10 mM; Sigma-Aldrich)
2.5 µL Dithiothreitol (DTT) solution (100 mM)
1.25 µL Dimethyl sulfoxide (DMSO, ≥99.9%; Sigma-Aldrich)
0.5 µL RNase Inhibitor, Human placenta (40,000 U/mL; New England Biolabs)
1 µL Enzyme Mix (reverse transcriptase, AMV)
5 µL Purified RNA
10 µL 5x RT-PCR reaction buffer (including 7.5 mM MgCl₂ and DMSO)
2 µL Forward (5') primer
2 µL Reverse (3') primer
Nano-Pure™ water (to total 50 µL reaction volume)

Isolation of individual cDNA fragments applied 2% agarose gel electrophoresis and ethidium bromide (EtBr)-based ultraviolet (UV) transillumination with subsequent purification of excised bands using the QIAquick PCR Purification kit (Qiagen). Purified cDNA samples were submitted with over-lapping sequencing primers to the UTMB Molecular Genomics Core for Sanger sequencing (**Table 2-3**).

Sanger sequencing reaction mix: (submitted for processing)

5 μ L Purified cDNA

1 μ L Forward (5') or Reverse (3') sequencing primer

2.5.2 Sanger sequence analysis

Resulting .AB1 output files were edited and assembled in the ContigExpress module of the VectorNTI Advance v11 program suite (Invitrogen) and exported as full-length FASTA consensus sequences.

2.6 PHYLOGENETIC AND SEQUENCE-BASED COMPARISONS**2.6.1 Sequence Alignments (general)**

Full-length, novel and/or published WNV FASTA consensus sequences were imported into BioEdit v7³⁰⁷ from either local or the GenBank.gov databases. Sequences were aligned using the Multiple Sequence Comparison by Log-Expectation (MUSCLE)^{308,309} or Clustal Omega^{310,311} algorithms on the EMBL-EBI server³¹² prior to trimming both the 5'- and 3'-ends to retain the encoded 10,299 nucleotide (nt) open reading frame (ORF). Dependent on the particular phylogenetic aim(s) (detailed in Chapters 3-5), inclusion of certain prototype sequences provided common phylogenetic out-groups in the described phylogenetic comparisons: IS-98 STD (AF481864)⁵⁴, IbAn-7019 (GQ851607), and/or AnD-27875 (GQ851606).

2.6.2 Neighbor-joining (NJ) phylogenetic comparisons

Importation of edited FASTA sequence alignments into Seaview v4.3.0 preceded NJ calculations using the Hasegawa-Kishino-Yano (HKY) 85 substitution model with 10,000 bootstrap replicates³¹³.

2.6.3 Maximum Likelihood (ML) phylogenetic comparisons

Maximum likelihood (ML) comparisons were conducted using the general-time-reversible (GTR) plus invariant (I) sites and a gamma (Γ) distribution (or GTR+I+ Γ) substitution model with 1,000 bootstrap replicates using RAxML-HPC Blackbox v7.3.2^{314,315} on the CIPRES Science Gateway tera-grid server³¹⁶.

2.6.4 Bayesian coalescent phylogenetic comparisons

Edited alignments were imported into BEAUti v1.7.4 to prepare the XML files required to run Bayesian comparisons in BEAST v1.7.4^{317,318}. In all applied phylogenetic runs, encoded XML files incorporated the GTR+I+ Γ_4 substitution model with applied taxa dates and an uncorrelated lognormal relaxed-clock model with or without Bayesian Skyline prior constraints. Resulting .log and directory tree (TRE) files were down-sampled from triplicate 50-100 million-state runs with 10% burn-in using both LogCombiner v1.7.4 and Tracer v1.5 with subsequent annotation of the optimized tree topology using TreeAnnotator v1.7.4^{317,318}.

2.6.5 Phylogenetic rendering

Inferred phylogenetic trees from NJ, ML, and Bayesian comparisons were all edited in FigTree v1.3.1 (www.mybiosoftware.com/phylogenetic-analysis/2407) and Microsoft PowerPoint 2010 in order to render images for publication.

2.6.6 Selection studies

Site-specific positive selection was performed on select, edited alignments with the single-likelihood ancestor counting (SLAC), fixed effect likelihood (FEL), and internal branches FEL (iFEL) methods on the Datamonkey server³¹⁹⁻³²². Positive selection at a specific, encoded amino acid codon was defined as the rate (d) of non-

synonymous (d_N) substitutions being greater than the rate of synonymous (d_S) substitutions (i.e., $d_N > d_S$) and a p -value < 0.05 in >1 applied method.

2.7 NEXT GENERATION SEQUENCING (NGS) PIPE-LINE

2.7.1 Preparation of vRNA for Illumina HiSeq1000 sequencing

Identical to the described Sanger sequencing protocol, RNA was extracted from applied WNV stocks using the QIAamp Viral RNA Mini kit (Qiagen) following the 30 minute inactivation of infectious virus in AVL Buffer. Purified RNA suspensions were stored at -80°C prior to submission to the UTMB Next Generation Sequencing core for processing (**Fig. 2-1**). cDNA libraries were constructed using random hexameric primers with the TruSeq RNA v2 kit (Illumina) and applied to the Illumina HiSeq1000 NGS system for sequencing.

2.7.2 NGS pipe-line: downstream data processing

Output 50 base-pair (bp) pair-end reads were edited using the Trimmomatic v0.22 and ILLUMINACLIP programs to remove adaptor sequences and low-quality ($Q < 35$) or ambiguous (N) bases at the ends of reads³²³. Processed FastQ contigs were then assembled either *de novo* via ABySS v1.3.7³²⁴ and BLAST or mapped to the either the prototype NY99-flamingo382-99 (AF196835) or attenuated TX1153 (AY712945) U.S. strains^{43,57}—as appropriate—using Bowtie2 v2.2.1³²⁵ with the “--very-sensitive-local” alignment setting. In order to limit sampling and single nucleotide variant (SNV) detection biases in applied NGS comparisons, PCR duplicates were removed with Picard-tools v1.110 using the MarkDuplicate.jar Java script (<http://picard.sourceforge.net>) with down-sampling of remaining mapped reads to $\sim 3,000$ -fold median depth-of-coverage via a Python v27 script (source-codes indicated below). Data file conversions in the described pipe-line (**Fig. 2-1**): FastQ, Sequence Alignment/Map (SAM), Binary

Alignment/Map (BAM), and vCard (VCF) files were conducted in Sequence Alignment/Map toolbox (SAMtools) v1.1.19³²⁶, SNVer v0.5.2³²⁷, or VarScan2 v2.3.6³²⁸.

2.7.3 Source codes

In all defined source-codes, XX represents the header in the processed FastQ, BAM, and SAM files (*e.g.*, XX equals 114 for WNV isolate TX114 [Bird114]). Source-codes were processed using a 3.60 GHz AMD FX(tm)-8150 Eight-Core processor with 8 gigabytes (GB) RAM on a 64-bit Windows 7 Home Premium operating system through the *Command Prompt*.

2.7.3.1 FASTQ CONTIG ALIGNMENT TO NY99 REFERENCE SEQUENCE

```
bowtie2 -p 2 --very-sensitive-local -x NY99 -1 XX_1 -2 XX_2  
-S XX 2> bowtie.report
```

2.7.3.2 CONVERSION FROM SAM TO BAM FILE FORMAT

Step One:

```
samtools view -Sb XX > XXa
```

Step Two:

```
samtools sort XXa XX
```

2.7.3.3 REMOVAL OF PCR DUPLICATES

Step One:

```
java -Xmx1200M -jar SortSam.jar INPUT=XX.bam  
OUPUT=XX_coordinate.bam SORT_ORDER=coordinate
```

Step Two:

```
java -Xmx1200M -jar MarkDuplicates.jar I=XX_coordinate.bam  
OUTPUT=XX_re.bam REMOVE_DUPLICATES=true  
OPTICAL_DUPLICATE_PIXEL_DISTANCE=0  
METRICS_FILE=XX.re.output
```

2.7.3.4 DOWN-SAMPLING TO NORMALIZE DEPTH-OF-COVERAGE

Step One:

```
java -Xmx1200M -jar SamToFastq.jar INPUT=XX_re.bam FASTQ=XX_1re  
SECOND_END_FASTQ=XX_2re
```

Step Two: (#### equals percent to be down-sampled)

```
python FASTQ_subsample_paired.py 0.#### XX_1re XX_2re XX_1dre  
XX_2dre
```

Step Three:

```
bowtie2 -p 2 --very-sensitive-local -x NY99 -1 XX_1dre -2 XX_2dre -S XX_re  
2> bowtie2.report
```

Step Four:

```
samtools view -Sb XX_re > XXa_re
```

Step Five:

```
samtools sort XXa_re XX_revsl
```

Step Six:

```
samtools index XX_revsl.bam XX_revsl.bam.bai
```

2.7.4 Bayesian phylogenetic comparisons

Full-length, consensus FASTA files were extracted from the indexed BAM files using SAMtools v0.1.19³²⁶. Bayesian coalescent phylogenetic comparisons applied identical parameters as defined above (in **Section 2.6**).

2.7.5 Shannon entropy calculations

Shannon entropy is a derivation of information theory that provides a quantifiable metric to evaluate the disorder (*i.e.*, genetic diversity) of a particular genomic position in the WNV genome. In effect, processed BAM files, normalized for depth-of-coverage, were applied to the bam2R function of the deepSNV v1.8.0 software

package³²⁹ in R (www.R-project.org) relative to the published, prototype NY99-flamingo382-99 (AF196835) sequence. Resulting tabulated reference and alternate nucleotide counts: A, C, G, T, and gap per position in the WNV genome were applied to a modified Shannon entropy equation in Microsoft Excel 2010 (see equation below)³³⁰:

$$\textbf{Shannon entropy } (S_n) = - \sum_{i=1}^n f_i (\ln f_i) \quad \dots \text{at } \textbf{each} \text{ nucleotide position}$$

n = A, C, G, T, and gap

f_i = proportion (0.00-1.00) of each nucleotide or gap at a particular position

Calculated Shannon entropy (S_n) values were averaged across the whole WNV genome, both 5' and 3'UTRs, and/or all 10 encoded genes resulting in transformed entropic indices ranging from 0.00 (minimum genetic diversity; fixed nucleotide position) to 1.61 (maximum genetic diversity; all elements [n] detected at 20% frequency).

2.7.6 Single nucleotide variant (SNV) detection

Single nucleotide variants (SNVs) were defined as alternative nucleotides: A, C, G, and T or insertion/deletion elements (*i.e.*, gaps) detected at $\leq 50\%$ mutational frequency relative to the NY99 reference sequence at a particular genomic position. SNVs were detected in each processed BAM file—normalized for depth-of-coverage—using both the VarScan2 v2.3.6³²⁸ heuristic model (strict 0.5% SNV detection limit) with SAMtools v0.1.19³²⁶ and the SNVer v0.5.2³²⁷ binomial-binomial, haploid detection model. Both platforms incorporated strand bias filters and applied limits for statistical significance (p -value < 0.05 ; Bonferroni-corrected $\alpha = 4.53 \times 10^{-6}$ in the SNVer model). Reported SNVs were confirmed with both detection platforms to limit identification of false-positive variants in either method³³¹.

2.8 DETERMINATION OF *IN VIVO* MOUSE NEUROINVASIVE PHENOTYPE

Natural WNV isolates with undetermined *in vivo* phenotype were screened in 3-4 week old outbred Swiss Webster mice for mouse neuroinvasive phenotype as defined in prior studies^{29,57,174}. Briefly, groups of five females were inoculated by the intraperitoneal (i.p.) route with either 10 or 5000 pfu of virus and monitored daily for clinical signs of illness. Isolates resulting in $\leq 50\%$ survival at the 10 pfu dose were designated “virulent” and all isolates with $\geq 50\%$ survival at the 5000 pfu dose were designated “attenuated.”

2.8.1 50% Lethal dose (LD₅₀) studies

Follow-up i.p. LD₅₀ (ipLD₅₀) studies at 10-fold doses of 10^3 to 10^{-1} pfu of virus ($n = 5$ mice per dose) were undertaken to confirm the *in vivo* neuroinvasive phenotype for all isolates identified as “attenuated” in **Section 2.8** (above) using the Spearman-Kärber method. Inclusion of both NY99-flamingo382-99 and TX1171 in these studies provided virulent and attenuated positive controls, respectively; in addition, mice inoculated with PBS (mock) served as negative controls. Statistical significance was determined via both the Log-rank Mantel-Cox and Gehan-Breslow-Wilcoxon tests at an $\alpha = 0.0001$ in GraphPad Prism v6.05 (GraphPad, San Diego, CA, USA).

2.8.2 Ethics statement

All described animal procedures were carried out according to the guidelines of the Committee on Care and Use of Laboratory Animals under an approved protocol in the UTMB Animal Biosafety Level 3 (ABSL3) research facilities.

2.9 QUANTITATIVE REVERSE-TRANSCRIPTASE PCR

2.9.1 RNA extraction

Following inoculation of 95% confluent A549 cell cultures (6.9×10^4 cells/cm²) with several WNV isolates or mutants at an MOI of 0.1, cell monolayers were harvested and processed at designated time-points using the Qiagen RNeasy Mini kit (as per the manufacturer's protocol). In brief, cell monolayers were washed with PBS followed by treatment with 600 μ L Buffer RLT and incubated at room temperature for 30 minutes to facilitate complete inactivation of infectious virus. Disrupted cell suspensions were processed immediately or stored at -80°C in 1.5 mL Eppendorf tubes under BSL2 lab conditions. Re-suspended cell homogenates were applied to Qiagen QIAshredder spin columns and processed according to the manufacturer's procedures. Flow-through lysates were mixed with 600 μ L 70% ethanol and applied to Qiagen RNeasy spin columns. Purified RNA was eluted using 30-60 μ L NanoPure™ water and stored at -80°C in 1.5 mL Eppendorf tubes.

2.9.2 cDNA synthesis

Purified RNA samples were diluted 50-fold in NanoPure™ water and quantified using 260 nm absorbance (A) with A260/A280 ratios ≥ 1.55 . In brief, 756 ng RNA was applied to the BioRad® iScript™ Advanced cDNA Synthesis kit and processed as indicated below:

Reaction mix:

4 μ L 5x iScript Advanced reaction mix (BioRad)

1 μ L iScript Advanced reverse-transcriptase (BioRad)

1 – 5 μ L quantified RNA (756 ng total)

NanoPure™ water (to 20 μ L total reaction volume)

Reaction protocol:

30 min. at 42°C

5 min. at 85°C

Hold at 4°C

cDNA were re-suspended with 60 µL NanoPure™ water (80 µL total volume) and stored at -20°C in 1.5 mL Eppendorf tubes.

2.9.3 SYBR® Green real-time PCR

Initial quantification of differential gene expression in the *in vitro* A549 platform (defined in **Specific Aim 3**) utilized SYBR® Green-based real-time PCR detection using the BioRad® PrimePCR™ Flavivirus Infections H96 panel (**Fig. 2-2**) on the BioRad® CFX Connect Real-Time PCR Detection system (**Table 2-4**). In brief, cDNA was applied to the PrimePCR™ detection panel as indicated below:

Reaction mix (per well):

10 µL 2x SsoAdvanced universal SYBR® Green supermix (BioRad)

8.5 µL NanoPure™ water

1.5 µL cDNA (approx. 9.45 ng/µL)

20 µL total reaction volume per well

2.9.4 Real-time PCR data management

Results from individual PrimePCR runs were imported into the BioRad® CFX Manager and pooled prior to normalization to *gapdh*, *thp*, and *hprt1* reference gene expression. In order to determine fold gene expression per applied WNV isolate, the $-\Delta\Delta C_T$ method was used for each individual gene target in the PrimePCR™ Flavivirus

Infections H96 panel relative to either mock or wild-type NY99 infection^{332,333}. Statistical significance for all identified fold-differences was determined using a Bonferroni-corrected Student's *t*-test (p -value < 0.05 ; $\alpha = 8.33 \times 10^{-3}$) to account for multiple comparisons at each individual time-point. Results were exported to GraphPad Prism v7 for graphical rendering.

2.9.5 NY99 infectious clone (NY99ic) studies: NS4B protein

Based on the results from this initial panel, the same procedures and protocols were applied to the comparison of fold gene expression between NY99ic isolates encoding unique amino acid substitutions in the NS4B protein. In contrast, a custom PrimePCR® plate was designed to enumerate fold-expression of 23 select genes relative to both mock and NY99ic infection at 36 hpi using the $-\Delta\Delta C_T$ method following initial normalization to the reference *gapdh* gene (**Fig. 2-3**)^{332,333}. Statistical significance applied a Bonferroni-corrected Student's *t*-test (p -value < 0.05 ; $\alpha = 4.167 \times 10^{-4}$) to limit detection of false-positive differences in fold gene expression. In addition, extracted mock, NY99ic-derived NS4B-P38G/T116I + NS3-N480H (P38G), and NY99ic RNA samples from the initial wild-type analyses were applied to the NY99ic-based comparisons to provide direct comparison of P38G-induced host gene expression in both experimental platforms.

2.10 BIO-PLEX PROTEIN QUANTIFICATION

2.10.1 Preparation of cytokine/chemokine samples

In-tandem with the extraction of host cell RNA for the described wild-type and NY99ic-based qRT-PCR studies, the cell culture supernatant was also harvested from infected A549 cell cultures at 6, 12, 24, and/or 36 hpi, aliquoted into 1.5 mL cryotubes, and stored at -80°C under BSL3 conditions.

2.10.2 BioRad® Bio-Plex Pro™ Human Cytokine 27-plex panel

Stored cell culture supernatants were thawed at room temperature and applied to the BioRad® Bio-Plex Pro™ Human Cytokine 27-plex panel (as per the manufacturer's protocol) using a vacuum manifold-based washing method. Detection of all 27 cytokines/chemokines in this panel was conducted on the Bio-Plex® 200 detection system with determination of raw protein concentrations (pg/mL) using the BioRad® Bio-Plex Manager™ software package. Outliers in the standard curves were removed prior to the determination of raw protein expression. Exported data files were processed in Microsoft Excel 2010 to confirm fold protein expression relative to mock, wild-type NY99, or NY99ic infection. Statistical significance for all identified differences in fold expression were conducted in GraphPad Prism v7 using a Kruskal-Wallis One-Way ANOVA and Dunn's *post-hoc* test ($\alpha = 0.05$) to account for multiple comparisons. Rendering of included graphics was conducted in both GraphPad Prism v7 and Microsoft PowerPoint 2010. Consistent with the described gene-based comparisons, harvested mock, NY99ic-derived P38G, and NY99ic supernatant samples from the initial platform were applied to the NY99ic-based cytokine/chemokine analyses as controls to enable direct comparison of P38G-induced host protein expression in both experimental platforms.

Table 2-1: WNV-specific reverse transcriptase (RT)-PCR primer sets

Primer Pair	Fragment Size (bp)	Primer*	Nucleotide Sequence
1	554	WN1F	5'-AGTAGTTCGCCTGTGTGA-3'
		WN553R	5'-CAGCAGCTGTTGGAAT-3'
2	1420	WN401F	5'-ACCGGAATTGCAGTCATGAT-3'
		WN1821R	5'-GCCATAGGTTGTTCCCTTCAAC-3'
3	959	WN862F	5'-CAGAGAGTTGTGTTTGTCTGTCTAT-3'
		WN1821R	5'-GCCATAGGTTGTTCCCTTCAAC-3'
4	891	WN1613F	5'-GGAACAGAGAGACGTTAATGGAGTT-3'
		WN2504R	5'-TCTTGCCGGCTGATGTCTATGGCA-3'
5	992	WN1613F	5'-GGAACAGAGAGACGTTAATGGAGTT-3'
		WN2605R	5'-CCAATCTGGAACTGATCGTAGAC-3'
6	1165	WN2418F	5'-TGGAGGAGTTTTGCTCTTC-3'
		WN3583R	5'-GGCCAAGAACACGACCAGAAGGCC-3'
7	1930	WN2881F	5'-CCGAGACCAAGGAATGTCCGACTC-3'
		WN4811R	5'-CTTGACACTGCCCCAGTATG-3'
8	2383	WN2881F	5'-CCGAGACCAAGGAATGTCCGACTC-3'
		WN5264R	5'-CTCAGTCTTCTGTTTATGGCCTC-3'
9	1710	WN5011F	5'-ACATCAGGCTACCAATAGT-3'
		WN6721R	5'-GTCGCGACTCCCAAGACAGC-3'
10	1522	WN5199F	5'-CGGCGCCGGTAAAACAAGG-3'
		WN6721R	5'-GTCGCGACTCCCAAGACAGC-3'
11	1729	WN5991F	5'-ACGTATCGGTAGAAATCCGT-3'
		WN7720R	5'-CTTTCCAAACCTCTCCCAAGGTGCGTCC-3'
12	1456	WN7087F	5'-ACGTCAGACTACATCAACACTT-3'
		WN8543R	5'-CGCCTGAGTCGTTCAATCC-3'
13	1650	WN7549F	5'-AATGGAGCAAGCTCTGTTTGG-3'
		WN9199R	5'-CAACTTCACGCAGGATGTAACC-3'
14	1556	WN8943	5'-GAGGAGCGCCAGAGAAGC-3'
		WN10499R	5'-GCAGCACCGTCTACTCAACTTC-3'
15	761	WN10268	5'-CGTGGGCAGAAAACATCCAGGTGGC-3'
		cWN11029	5'-AGATCCTGTGTTCTCGCACCAACC-3'

*Primers are designated Forward (F; 5') or Reverse (R; 3'). Exceptions include the cWN11029 primer which is designated "complimentary (c)" or anti-sense to the viral RNA strand.

Table 2-2: Touch-down Titan One-Tube RT-PCR reaction protocol

Step	Cycles	Temperature (°C)	Time	Modifications
1	1x	50	30 min.	-
		94	5 min.	-
2	10x	94	50 sec.	-
		61	45 sec.	-2°C every 2x cycle repeat*
		68	3 min.	-
3	18x	94	50 sec.	-
		51	45 sec.	-
		68	3 min.	+15 sec. every cycle repeat†
4	1x	68	7 min.	-
		4	∞	-

*Annealing temperature decreased from 61°C to 51°C over 10 cycles.
†Extension time increased from 3 min. to 7 min. 30 sec. over 18 cycles.

Table 2-3: WNV-specific Sanger sequencing primers, related to Table 2-1.

Primer Pair	Primer*	Nucleotide Sequence
1	WN1F	5'-AGTAGTTCGCCTGTGTGA-3'
	WN533R	5'-CAGCAGCTGTTGGAAT-3'
2	WN401F	5'-ACCGGAATTGCAGTCATGAT-3'
	WN974R	5'-GAAGTCTCTGTTGCTCATTCCAAGGC-3'
	WN1031F	5'-ATTTGGTTCTCGAAGGCGAC-3'
	WN1200R	5'-CGTTTGTCAATTGTGAGCTTC-3'
	WN1821R	5'-GCCATAGGTTGTTCCCTTCAAC-3'
3	WN862F	5'-CAGAGAGTTGTGTTTGTCTGCTAT-3'
	WN1200R	5'-CGTTTGTCAATTGTGAGCTTC-3'
	WN1821R	5'-GCCATAGGTTGTTCCCTTCAAC-3'
4	WN1613F	5'-GGAACAGAGAGACGTTAATGGAGTT-3'
	WN2504R	5'-TCTTGCCGGCTGATGTCTATGGCA-3'
5	WN1613F	5'-GGAACAGAGAGACGTTAATGGAGTT-3'
	WN2605R	5'-CCAATCTGGAAACTGATCGTAGAC-3'
6	WN2418F	5'-TGGAGGAGTTTTGCTCTTC-3'
	WN3283R	5'-TGTACCCTGGTCTCCTGT-3'
	WN3583R	5'-GGCCAAGAACACGACCAGAAGGCC-3'
7	WN2881F	5'-CCGGAGACCAAGGAATGTCCGACTC-3'
	WN3411F	5'-AGGAGCTGCACCTTACCACCAC-3'
	WN3583R	5'-GGCCAAGAACACGACCAGAAGGCC-3'
	WN4028F	5'-GCTTGAATCTGGATGTGTACAGG-3'
	WN4050R	5'-CCTGTACACATCCAGATTCAAGC-3'
	WN4811R	5'-CTTGACACTGCCCCAGTATG-3'
8	WN2881F	5'-CCGGAGACCAAGGAATGTCCGACTC-3'
	WN3411F	5'-AGGAGCTGCACCTTACCACCAC-3'
	WN3583R	5'-GGCCAAGAACACGACCAGAAGGCC-3'
	WN4028F	5'-GCTTGAATCTGGATGTGTACAGG-3'
	WN4050R	5'-CCTGTACACATCCAGATTCAAGC-3'
	WN4811R	5'-CTTGACACTGCCCCAGTATG-3'
	WN5264R	5'-CTCAGTCTTCTGTTTATGGCCTC-3'
9	WN5011F	5'-ACATCAGGCTCACCAATAGT-3'
	WN5417R	5'-GGAGACATCAGCCTG-3'
	WN5991F	5'-ACGTATCGGTAGAAATCCGT-3'
	WN6497R	5'-CCCAGAACCTCAATGAGCCCTATCTGAG-3'
	WN6721R	5'-GTCGCGACTCCCAAGACAGC-3'
10	WN5199F	5'-CGGCGCCGGTAAAACAAGG-3'
	WN5991F	5'-ACGTATCGGTAGAAATCCGT-3'
	WN6497R	5'-CCCAGAACCTCAATGAGCCCTATCTGAG-3'
	WN6721R	5'-GTCGCGACTCCCAAGACAGC-3'
11	WN5991F	5'-ACGTATCGGTAGAAATCCGT-3'
	WN6396S	5'-CTGGATTGACGCCAGGGTGACTC-3'
	WN6721R	5'-GTCGCGACTCCCAAGACAGC-3'
	cWN7282	5'-CCATGTAAGCATAGTGGC-3'
	WN7720R	5'-CTTTCCAAACCTCTCCAAGGTGCGTCC-3'

*Primers were designated Forward (F; 5') or Reverse (R; 3') with the initial nucleotide indicated relative to the prototype NY99-flamingo382-99 (NY99; AF196835) sequence. Exceptions include primers designated "complimentary (c)" or "sense (s)" to the viral RNA strand.

Table 2-3: WNV-specific Sanger sequencing primers, related to **Table 2-1** (Continued).

Primer Pair	Primer*	Nucleotide Sequence
12	WN7087F	5'-ACGTCAGACTACATCAACACTT-3'
	WN7549F	5'-AATGGAGCAAGCTCTGTTTGG-3'
	WN8060R	5'-ACTCCACTCTTCATGGTAA-3'
	WN8543R	5'-CGCCTGAGTCGTTCAATCC-3'
13	WN7549F	5'-AATGGAGCAAGCTCTGTTTGG-3'
	WN8330F	5'-ACGAGATGTATTGGGTGAGTCG-3'
	WN8543R	5'-CGCCTGAGTCGTTCAATCC-3'
	WN9199R	5'-CAACTTCACGCAGGATGTAACC-3
14	WN8943F	5'-GAGGAGCGCCAGAGAAGC-3'
	cWN9592	5'-CCACATCATCTGGGCCAATCAC-3'
	WN9725F	5'-CGCTCCACTTCCTCAATGC-3'
	WN10117R	5'-CCTCTGTTGTCATCCACTCTCC-3'
	WN10499R	5'-GCAGCACCGTCTACTCAACTTC-3'
15	WN10268S	5'-CGTGGGCAGAAAACATCCAGGTGGC-3'
	cWN11029	5'-AGATCCTGTGTTCTCGCACCACC-3'

*Primers were designated Forward (F; 5') or Reverse (R; 3') with the initial nucleotide indicated relative to the prototype NY99-flamingo382-99 (NY99; AF196835) sequence. Exceptions include primers designated "complimentary (c)" or "sense (s)" to the viral RNA strand.

Table 2-4: BioRad® PrimePCR™ real-time PCR protocol

Step	Cycles	Temperature (°C)	Time	Modifications
1	1x	95	2 min.	-
2	40x	95	5 sec.	-
		60	30 sec.	-
3	60x	95	5 sec.	-
		65	5 sec.	+0.5°C each additional step*
4	-	4	∞	-

*Extension temperature increased from 65°C to 95°C over 60 cycles.

Figure 2-1: Illumina HiSeq1000 Next Generation Sequencing (NGS) pipe-line

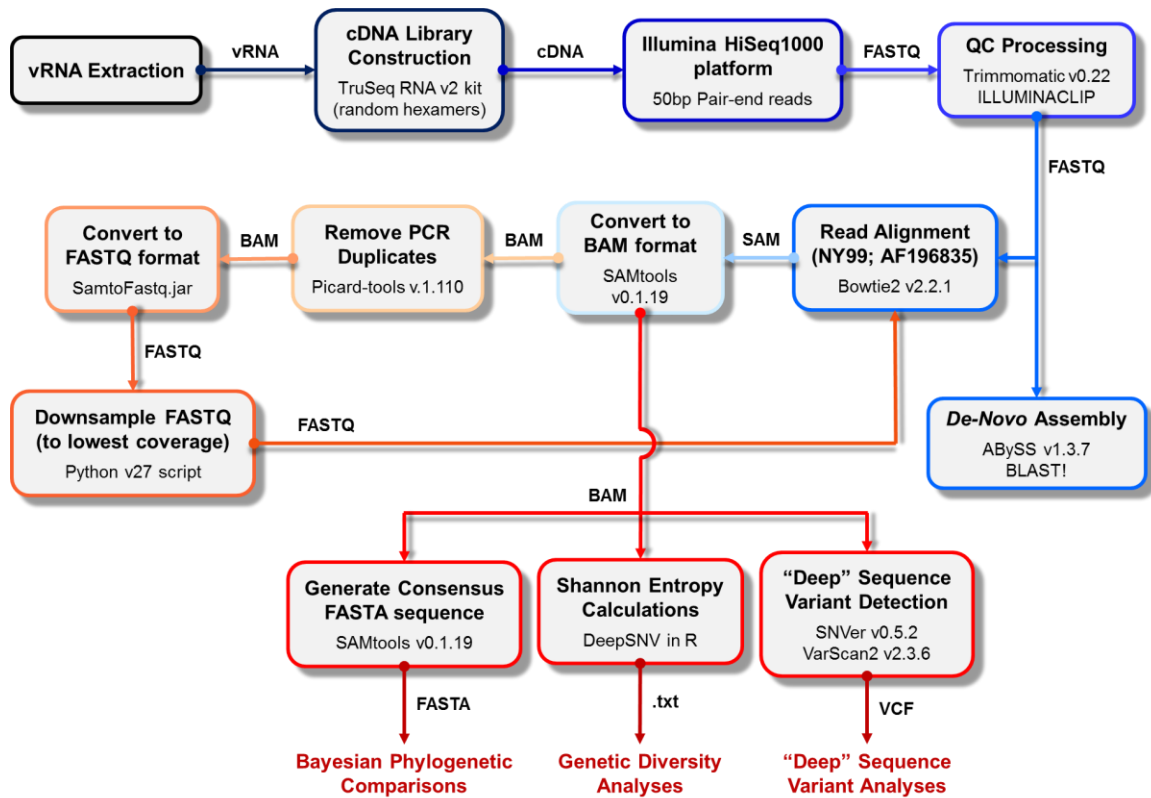


Figure 2-2: BioRad® PrimePCR™ Flavivirus Infections H96 panel

	1	2	3	4	5	6	7	8	9	10	11	12
A	<i>casp8</i>	<i>if144l</i>	<i>il1b</i>	<i>ly6e</i>	<i>tlr1</i>	<i>tbp</i>	<i>casp8</i>	<i>if144l</i>	<i>il1b</i>	<i>ly6e</i>	<i>tlr1</i>	<i>tbp</i>
B	<i>ccl5</i>	<i>ifih1</i>	<i>il1rl1</i>	<i>mapk11</i>	<i>tlr2</i>	<i>gapdh</i>	<i>ccl5</i>	<i>ifih1</i>	<i>il1rl1</i>	<i>mapk11</i>	<i>tlr2</i>	<i>gapdh</i>
C	<i>cd14</i>	<i>ifit1</i>	<i>il6</i>	<i>myd88</i>	<i>tlr3</i>	<i>hprt1</i>	<i>cd14</i>	<i>ifit1</i>	<i>il6</i>	<i>myd88</i>	<i>tlr3</i>	<i>hprt1</i>
D	<i>cd38</i>	<i>ifit3</i>	<i>il8</i>	<i>nfkb1</i>	<i>tlr4</i>	gDNA	<i>cd38</i>	<i>ifit3</i>	<i>il8</i>	<i>nfkb1</i>	<i>tlr4</i>	gDNA
E	<i>cd86</i>	<i>ifitm1</i>	<i>irf4</i>	<i>nfkb2</i>	<i>tnf</i>	PCR	<i>cd86</i>	<i>ifitm1</i>	<i>irf4</i>	<i>nfkb2</i>	<i>tnf</i>	PCR
F	<i>cxcl10</i>	<i>ifnb1</i>	<i>irf7</i>	<i>nos2</i>	<i>tnfrsf1a</i>	RQ1	<i>cxcl10</i>	<i>ifnb1</i>	<i>irf7</i>	<i>nos2</i>	<i>tnfrsf1a</i>	RQ1
G	<i>ddx58</i>	<i>il12a</i>	<i>isg15</i>	<i>oas3</i>	<i>tnfsf10</i>	RQ2	<i>ddx58</i>	<i>il12a</i>	<i>isg15</i>	<i>oas3</i>	<i>tnfsf10</i>	RQ2
H	<i>if144</i>	<i>il12b</i>	<i>lgals3bp</i>	<i>samd9</i>	<i>trim5</i>	RT	<i>if144</i>	<i>il12b</i>	<i>lgals3bp</i>	<i>samd9</i>	<i>trim5</i>	RT

Sample 1
Sample 2

Figure 2-3: Custom PrimePCR™ plate design

	1	2	3	4	5	6	7	8	9	10	11	12
A	<i>ccl5</i>	<i>il12a</i>	<i>nfk1b1</i>	<i>ccl5</i>	<i>il12a</i>	<i>nfk1b1</i>	<i>ccl5</i>	<i>il12a</i>	<i>nfk1b1</i>	<i>ccl5</i>	<i>il12a</i>	<i>nfk1b1</i>
B	<i>cxcl10</i>	<i>il1b</i>	<i>nfk1b2</i>	<i>cxcl10</i>	<i>il1b</i>	<i>nfk1b2</i>	<i>cxcl10</i>	<i>il1b</i>	<i>nfk1b2</i>	<i>cxcl10</i>	<i>il1b</i>	<i>nfk1b2</i>
C	<i>ddx58</i>	<i>il6</i>	<i>oas3</i>	<i>ddx58</i>	<i>il6</i>	<i>oas3</i>	<i>ddx58</i>	<i>il6</i>	<i>oas3</i>	<i>ddx58</i>	<i>il6</i>	<i>oas3</i>
D	<i>if144</i>	<i>il8</i>	<i>tlr3</i>	<i>if144</i>	<i>il8</i>	<i>tlr3</i>	<i>if144</i>	<i>il8</i>	<i>tlr3</i>	<i>if144</i>	<i>il8</i>	<i>tlr3</i>
E	<i>ifih1</i>	<i>irf3</i>	<i>tnf</i>	<i>ifih1</i>	<i>irf3</i>	<i>tnf</i>	<i>ifih1</i>	<i>irf3</i>	<i>tnf</i>	<i>ifih1</i>	<i>irf3</i>	<i>tnf</i>
F	<i>ifit1</i>	<i>irf7</i>	<i>tnfsf10</i>	<i>ifit1</i>	<i>irf7</i>	<i>tnfsf10</i>	<i>ifit1</i>	<i>irf7</i>	<i>tnfsf10</i>	<i>ifit1</i>	<i>irf7</i>	<i>tnfsf10</i>
G	<i>ifit3</i>	<i>isg15</i>	<i>trim25</i>	<i>ifit3</i>	<i>isg15</i>	<i>trim25</i>	<i>ifit3</i>	<i>isg15</i>	<i>trim25</i>	<i>ifit3</i>	<i>isg15</i>	<i>trim25</i>
H	<i>ifnb1</i>	<i>myd88</i>	<i>gapdh</i>	<i>ifnb1</i>	<i>myd88</i>	<i>gapdh</i>	<i>ifnb1</i>	<i>myd88</i>	<i>gapdh</i>	<i>ifnb1</i>	<i>myd88</i>	<i>gapdh</i>
	Sample 1			Sample 2			Sample 3			Sample 4		

CHAPTER 3

Dynamic Transmission of West Nile virus Across the United States-Mexican Border²

3.1 INTRODUCTION

West Nile virus remains a significant public health burden in the U.S. with >17,463 WNND and >22,094 WNF clinical cases and numerous equid infections reported from 1999-2013 alone⁴⁸. Despite continued WNV circulation in the U.S., few veterinary or human clinical cases have been confirmed in Mexico. However, serological screening of Mexican equid serum samples collected in 2002 and 2006-2007 identified evidence of wide-spread WNV transmission with anti-WNV antibodies observed in up to 62.5% of horses sampled in >14 Mexican states^{53,297,334-338}. Furthermore, the prototype Mexican WNV strain (TM171-03 [Mex03]; AY660002) was isolated in southeastern Tabasco State from a dead raven in 2003^{28,53}; however, human WNV infections in Mexico remain limited to eight confirmed clinical cases of WNF or WNND reported in the northern states of Chihuahua ($n = 4$), Nuevo Leon ($n = 2$), and Sonora ($n = 2$) with few additional WNV isolates collected from *Culex* (*Cx.*) spp. mosquito pools, horses, and birds between 2003 to 2004^{49,53,296,298,299,335}.

The current paradigm for WNV introduction into Mexico supports movement via seasonal bird migrations across the Gulf-of-Mexico into the Yucatan Peninsula⁴⁹; in turn, the absence of the E-V159A substitution in the prototype 2003 Tabasco strain places this initial introduction prior to the emergence of the NA/WN02 U.S. genotype. However,

² Based on the Copyright Transfer Agreement/License provided by the Copyright Clearance Center, Elsevier hereby grants permission to reproduce these materials published in Virology. The citation is: **Mann, B.R., McMullen, A.R., Guzman, H., Tesh, R.B., & Barrett, A.D.T. (2013).** Dynamic transmission of West Nile virus across the United States-Mexican border. *Virology* **436**(1), 75-80.

identification of the characteristic E-V159A, NS4A-A85T, and NS5-K314R amino acid substitutions in all 2003-2004 northern Mexico WNV isolates suggests the separate introduction(s) of the SW/WN03 U.S. genotype into northern Mexico via a unidirectional WNV transmission model across the Southwest border^{49,296,335,336}. In this chapter (**Subaim 1a**), nine novel WNV isolates collected from El Paso, Texas, U.S. and Ciudad Juarez, Chihuahua State, Mexico between 2005-2010 were sequenced and applied to in-depth phylogenetic analyses to investigate the potential role of dynamic WNV transmission along the U.S.-Mexican border on the evolution of WNV in the southwestern U.S.

3.2 RESULTS

3.2.1 Genomic sequence analysis: El Paso, Texas and Ciudad Juarez WNV isolates

Nine WNV isolates collected in El Paso, Texas, U.S., and Ciudad Juarez, Chihuahua State, Mexico were obtained from the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA) at the University of Texas Medical Branch (UTMB) and from the Texas State Department of Health Services in Austin (**Table 3-1**). Seven isolates were collected from 2005-2010 *Cx. spp.* mosquito pools in El Paso, Texas with one isolate each from 2005, 2007, and 2008, two from 2009, and two from 2010. In addition, two mosquito pool isolates, one each in 2008 and 2009, were collected from WNV-positive mosquito pools in Ciudad Juarez, Mexico. Six of the seven El Paso isolates (all but TX AR9-5282) were isolated from *Cx. tarsalis* mosquitoes. The two Ciudad Juarez isolates (including the TX AR9-5282 isolate) were cultured from *Cx. quinquefasciatus* mosquito pools.

Full-length, genomic consensus sequences were obtained for all nine El Paso and Ciudad Juarez isolates using over-lapping, primer-based RT-PCR of vRNA extracted

from Vero(p1) working stocks with the QIAmp Viral RNA Mini kit (Qiagen). Sanger sequencing of resulting cDNA fragments produced contig sequences aligned and edited—relative to the prototype NY99-flamingo382-99 U.S. strain (AF196835)—in ContigExpress of the VectorNTI Advance® v11 suite (Life Technologies, Grand Island, NY, US)^{58,63}.

Comparison of all nine isolates with the prototype NY99-flamingo382-99 (NY99) strain identified 43-72 nucleotide (nt) differences (0.39-0.65%) per 11,029 nt genome with increased nt divergence relative to the original TM171-03 Mexican isolate (**Table 3-2**)^{28,43,53}. In addition, all nine El Paso and Ciudad Juarez isolates encoded 12 of the 13 conserved nucleotide changes defining the NA/WN02 U.S. genotype^{55,58}; however, no single isolate encoded the SW/WN03 genotype C to U transition at position 3774 in the NS2A gene⁶³. Furthermore, all nine isolates differed at 36 residues in the encoded polyprotein relative the prototype NY99 strain with 5-14 (0.15-0.41%) substitutions per isolate (**Table 3-3**).

3.2.2 Evidence of regional WNV evolution on the US-Mexican border

Based on these sequence-dependent trends, integration of applied distance-based, maximum likelihood-based, and coalescent phylogenetic models confirmed the genotypic relationship of all sequenced El Paso and Ciudad Juarez isolates in the evolution of WNV on the U.S.-Mexican border. In total, 357 sequences—including the full-length, coding regions of the nine isolates sequenced in this chapter and the prototype IS-98 STD (Israel 1998; AF481864) isolate plus 347 published North American WNV isolates in the on-line GenBank.gov database (as of December 2011)—were assembled in BioEdit v7.0.9.0 and aligned using the MUSCLE algorithm on the EMBL-EBI server to compile a detailed North American sequence alignment³⁰⁷⁻³⁰⁹. Neighbor-joining (NJ) phylogenies applied the Hasegawa-Kishino-Yano 85 (HKY-85) substitution model with 10,000 bootstrap replicates in Seaview v4.3.0³¹³; maximum likelihood (ML) methods applied the

GTR+I+ Γ_4 substitution model with 1000 bootstrap replicates in RAxML v7.2.8 Blackbox on the CIPRES Science Gateway v3.1 teragrid server³¹⁴⁻³¹⁶. Paired Bayesian coalescent phylogenies also applied the GTR+I+ Γ_4 substitution model using a log-normal relaxed clock model with triplicate 50 million-state runs in the BEAST v1.6.2 software package (<http://beast.bio.ed.ac.uk>) which were down-sampled with 10% burn-in and annotated via LogCombiner v1.7.4 and TreeAnnotator v1.7.4^{317,318}. Inferred phylogenetic trees were edited in FigTree v1.3.1 ([www.mybiosoftware.com/phylogenetic -analysis/2407](http://www.mybiosoftware.com/phylogenetic-analysis/2407)) prior to rooting to the Israel IS-98 STD isolate as a common outgroup.

Inclusion of the novel 2005-2010 El Paso and Ciudad Juarez isolates retained the published, topologic distribution of the NY99, NA/WN02, and SW/WN03 genotypes (**Fig 3-1A**)⁶³. Furthermore, conserved nt divergence between all nine isolates indicates the co-circulation of three distinct genetic groups (**Figs. 3-1B and 1C**). Group 1 (0.06-0.19% divergence) consists of three 2009-2010 El Paso isolates: TX AR9-5282, TX AR9-7465, and TX AR10-5718; Group 2 (0.07-0.19%) includes the TX AR8-5947, TX AR8-6686, and TX AR9-6115 El Paso and Ciudad Juarez isolates collected from 2008-2009; and Group 3 (0.26-0.46%) consists of three 2005-2010 El Paso isolates: TX AR5-2686, TX AR7-6745, and TX AR10-6572 (**Fig. 3-1**). Each amino acid substitution conserved in >1 isolate remained limited to a single phylogenetic group excluding the E-V159A, NS4A-A85T, and NS5-K314R substitutions (**Table 3-3**). In particular, the Group 1 isolates lack the SW/WN03 genotype NS4A-A85T substitution whereas the Group 2 isolates do not encode the SW/WN03 genotype NS5-K314R substitution. All three substitutions are absent from the prototype TM171-03 isolate.

Based on these trends, distribution of the Group 1 isolates within the NA/WN02 genotype was not focused to a single monophyletic node; in contrast, both Group 2 and Group 3 exhibited restricted clustering within the SW/WN03 genotype in all applied phylogenetic methods (**Figs. 3B and 3C**). In addition, the Group 1 cluster exhibited consistent >70 internal bootstrap values and 100% posterior probabilities despite a

diverse temporal and geographic range in isolates including a 2003 Connecticut (WNV-1/US/BID-v4586/2003), 2008 New York (WNV-1/US/BID/v4622/2008), and 2009 Texas (TX7827) isolate sharing a common monophyletic lineage. Mean bootstrap values <15 and posterior probabilities ranging from 26-73%—dependent on the applied phylogenetic method—coupled with 0.53-0.73% sequence divergence failed to support the robust monophyletic relationships of the Group 1 cluster among other, published NA/WN02 genotype isolates.

In comparison, the Group 2 isolates formed a stable outgroup (43-100% posterior probabilities) of the SW/WN03 genotype with a 2003 Colorado (Colorado 3258) and 2007 New York (WNV-1/US/BID/v4093/2007) isolate. In particular, comparison of the 2007 New York and Group 2 isolates indicated 0.52-0.63% nt and 0.26-0.35% deduced amino acid divergence, with the conserved absence of the NS5-K314R substitution. The Group 3 El Paso isolates clustered within the SW/WN03 genotype with the conserved distribution of the 2005 TX AR5-2686 isolate as an outgroup to several 2003-2007 Arizona, California, Colorado, and Texas isolates. Both the TX AR7-6745 and TX AR10-6572 isolates formed stable monophyletic lineages with two 2006 and 2007 Texas isolates: 011WG-TX06EP and 013WG-TX07EP exhibiting 13-84% bootstrap values and posterior probabilities in addition to <0.29% nucleotide divergence among these four isolates. Significant sequence divergence (0.52-0.76%) between the Group 2 and Group 3 isolates relative to the prototype NY99 (0.39-0.54%) and TM171-03 (0.63-0.81%) strains supports the disparate emergence of these two phylogenetic clusters from a common ancestral SW/WN03 genotype strain.

3.3 DISCUSSION

In this chapter, inclusion of nine novel, sequenced 2005–2010 El Paso, Texas and Ciudad Juarez, Mexico isolates in applied phylogenetic comparisons of all published

North American strains provided a robust platform to investigate natural dynamics driving WNV evolution on the U.S.–Mexican border. Full-length genomic sequencing of both Ciudad Juarez isolates expanded the limited number of published WNV strains collected from wild birds ($n = 9$) and a single horse in northern and southeastern Mexico^{28,49}. In addition, these results support the co-circulation of three distinct genetic pools of WNV in the El Paso, Texas and Ciudad Juarez principalities exhibiting conserved phylogenetic clustering within the NA/WN02 (**Fig. 3-1B**) and SW/WN03 (**Fig. 3-1C**) U.S. genotypes. Furthermore, significant nucleotide divergence between the two Ciudad Juarez and 2008 El Paso isolates (Group 2; **Fig. 3-1C**) from both prototype NY99 (>0.43%) and TM171-03 (>0.69%) strains and additional 2005, 2007, and 2010 El Paso isolates (>0.52%) provides robust evidence for active WNV transmission on the U.S.-Mexican border between 2005-2010.

Emergence of the SW/WN03 genotype in the southwestern U.S. from 2003-2008 resulted in the rapid regional displacement of the NA/WN02 genotype⁶³; however, undetected circulation of either genotype in northern Mexico since the initial 2003 enzootic outbreak contrasts the continued detection and evolution of WNV in the U.S. Based on these epidemiological trends, surveillance of WNV transmission in Mexico remains limited to periodic local, state-wide, and/or national serum sampling campaigns of local equid populations and rare reports of human WNF and WNND cases^{53,296-299,334,337-339}. Possible explanations for the comparative absence of WNV activity in Mexico include (1) serologic cross-protection and/or competition with other endemic flaviviruses such as dengue or St. Louis encephalitis viruses^{299,334,340-342}; (2) under-reporting or clinical misdiagnosis of WNF under the dengue fever clinical umbrella; (3) passage of an attenuated WNV phenotype in resident *Culex* spp. mosquito populations or bird amplifying hosts^{28,109,143,343}; and/or (4) a range of other potential environmental and socioeconomic factors. However, inoculation of the prototype TM171-03 and 2004 northern Mexico Tecate strains into susceptible, indigenous bird hosts confirmed

comparable peak viremia, tissue tropism, and lethality despite attenuation of the TM171-03 strain (in particular) compared to the virulent NY99-flamingo382-99 strain in American crows^{109,343}. Critical differences in *Culex* spp. and wild bird speciation, distribution, and susceptibility to WNV infection offer additional, alternative explanations for the lack of clinical signs or symptoms while supporting circulation of a distinct, attenuated Mexican WNV phenotype.

To date, phylogenetic models of North American WNV evolution support the unidirectional introduction of WNV into the Yucatan Peninsula prior to 2003 with subsequent expansion into northern Mexico from the southwestern U.S between 2003-2004⁴⁹. In this chapter, recent (2005-2010) emergence of three independent genetic groups of WNV isolates with $\geq 0.52\%$ nucleotide divergence in the El Paso and Ciudad Juarez area support still undefined selective pressures distinct from the surrounding southwestern U.S. regions; furthermore, limited nucleotide and amino acid conservation (*i.e.*, E-V159A substitution) between both Ciudad Juarez isolates and published 2003-2004 northern Mexico and Tabasco TM171-03 strains precludes emergence of the Group 2 isolates from either northern (Ciudad Juarez) or southern (Tabasco) Mexican origin.

McMullen *et al.* (2011) demonstrated clear phylogenetic clustering of published 2003-2004 northern Mexico isolates within the SW/WN03 genotype distinct from the confirmed grouping of the prototype TM171-03 isolate in the NY99 genotype^{49,63}. Inferred monophyletic lineage of the Group 2 isolates with both 2003 Colorado and 2007 New York strains further supports this incongruent origin and the divergent evolution of this outgroup from an ancestral SW/WN03 genotype strain consistent with the second introduction of WNV into northern Mexico from 2003-2008 (**Fig. 3-2**). Robust phylogenetic clustering and limited sequence divergence ($<0.19\%$) between the Group 2 isolates compared to U.S. isolates of close temporal and geographic distribution reinforces the subsequent reintroduction of the 2008 El Paso isolate from Ciudad Juarez following circulation of an adapted SW/WN03 genotype strain in northern Mexico.

Based on these results, selective pressure from circulating Mexican WNV strains or (possibly) other related flaviviruses is a potential ecological stimulus contributing to the diverse genetic profile observed in the 2005-2010 El Paso and Ciudad Juarez isolates; in addition, inferred co-circulation of both the NA/WN02 and SW/WN03 genotypes further emphasizes the identified divergence of the 2008-2009 Group 2 isolates. In effect, this proposed model offers a novel, dynamic route in WNV transmission along the U.S.-Mexican border influencing the sustained evolution of WNV in the southwestern U.S. In conclusion, continued isolation and phenotypic characterization of southwestern U.S. and northern Mexico WNV isolates will be required to confirm the underlying dynamics driving the transmission and evolution of WNV in this region of North America.

Table 3-1: West Nile virus isolates described in the conducted genetic and phylogenetic analyses, 1998-2010.

Strain*	Location	Collection Year	Source†	GenBank accession no.
IS-98 ISD	Eilat, Israel	1998	White stork	AF481864
NY99-flamingo382-99	New York, NY, US	1999	Chilean flamingo	AF196835
TM171-03	Tabasco, Mexico	2003	Raven	AY660002
Colorado 3258	Colorado, US	2003	Magpie	DQ164203
WNV-1/US/BID-v4585/2003	Connecticut, US	2003	<i>Cx. salinarius</i>	HM488220
TX AR5-2686	El Paso, Texas, US	2005	<i>Cx. tarsalis</i>	JX015515
011WG-TX06EP	Texas, US	2006	Human	GQ507470
013WG-TX07EP	Texas, US	2007	Human	GQ507471
WNV-1/US/BID-v4093/2007	New York, US	2007	American crow	HM488201
TX AR7-6745	El Paso, Texas, US	2007	<i>Cx. tarsalis</i>	JX015516
WNV-1/US/BID-v4622/2008	New York, US	2008	American Crow	HM488237
TX AR8-6866	Ciudad Juarez, Mexico	2008	<i>Cx. quinquefasciatus</i>	JX015518
TX AR8-5947	El Paso, Texas, US	2008	<i>Cx. tarsalis</i>	JX015517
TX7827	Texas, USA	2009	Blue jay	JF415924
TX AR9-6115	Ciudad Juarez, Mexico	2009	<i>Cx. quinquefasciatus</i>	JX015520
TX AR9-5282	El Paso, Texas, US	2009	<i>Cx. quinquefasciatus</i>	JX015519
TX AR9-7465	El Paso, Texas, US	2009	<i>Cx. tarsalis</i>	JX015521
TX AR10-5718	El Paso, Texas, US	2010	<i>Cx. tarsalis</i>	JX015522
TX AR10-6572	El Paso, Texas, US	2010	<i>Cx. tarsalis</i>	JX015523

***Boldface**, strains sequenced in this chapter.

†El Paso, Texas and Ciudad Juarez, Mexico isolates were collected from *Culex* (*Cx.*) spp. mosquito pools.

Table 3-2: Nucleotide and amino acid divergence between El Paso and Ciudad Juarez WNV isolates

Strain*	Group 1					Group 2			Group 3		
	NY99	TM171-03	TX5282	TX7465	TX5718	TX6866	TX5947	TX6115	TX2686	TX6745	TX6572
NY99	-	0.41	0.61	0.64	0.65	0.43	0.46	0.54	0.41	0.39	0.53
TM171-03	0.09	-	0.87	0.89	0.92	0.69	0.73	0.81	0.67	0.63	0.77
TX5282	0.32	0.44	-	0.15	0.06	0.74	0.78	0.86	0.72	0.68	0.82
TX7465	0.26	0.38	0.12	-	0.19	0.79	0.83	0.92	0.76	0.76	0.90
TX5718	0.41	0.52	0.09	0.20	-	0.80	0.83	0.92	0.79	0.75	0.89
TX6866	0.26	0.38	0.52	0.47	0.61	-	0.07	0.17	0.53	0.52	0.65
TX5947	0.26	0.38	0.52	0.47	0.61	0.06	-	0.19	0.56	0.56	0.70
TX6115	0.35	0.47	0.61	0.55	0.70	0.15	0.09	-	0.64	0.63	0.76
TX2686	0.20	0.35	0.38	0.35	0.50	0.35	0.35	0.44	-	0.33	0.46
TX6745	0.15	0.26	0.29	0.23	0.38	0.29	0.29	0.38	0.17	-	0.26
TX6572	0.26	0.38	0.47	0.41	0.55	0.41	0.41	0.47	0.29	0.17	-

*Percent (%) nucleotide divergence from NY99 and TM171-03 is indicated above the diagonal with amino acid divergence shown below the diagonal. Shaded **boldface**, sequences with high nucleotide and/or amino acid similarity as indicated.

Table 3-3: Detected amino acid substitutions in applied El Paso and Ciudad Juarez WNV consensus sequences, 2005-2010

Gene*	Position	NY99†	Group 1			Group 2			Group 3		
			TX7465	TX5282	TX5718	TX6866	TX5947	TX6115	TX2686	TX6745	TX6572
C	119	A	V	V	V
	121	V	A	A	A
prM	140	V	.	I	I	A
E	51	A	T	T	T
	159	V	A	A	A	A	A	A	A	A	A
	168	S	T	.	.	.
	332	T	M
	467	A	S	.	.	.
NS1	314	R	.	.	.	K	K	K	.	.	.
NS2A	26	K	R	R	R
	46	F	.	.	S	L
	89	F	L	L	L
	92	A	T	.
	119	H	.	.	.	Y	Y	Y	.	.	.
	190	K	R
	224	A	.	.	.	V	V	V	.	.	.
NS2B	78	R	W	.	.
	116	L	M	M	.	.	.
	119	V	L	L	L
NS3	155	I	M	.	.
	249	P	.	L	L
	258	V	.	.	.	I	I	I	.	.	.
	355	Y	.	.	.	F	F	F	.	.	.
	518	Y	C	.	.

*C, capsid; prM, pre-membrane; E, envelope; NS, nonstructural.

†All amino acid substitutions indicated relative to the prototype NY99-flamingo382-99 strain (NY99) [AF196835]. Dots “.” indicate no difference from the NY99 consensus sequence. Shaded regions, conserved amino acid substitutions within Groups 1-3.

Table 3-3: Detected amino acid substitutions in applied El Paso and Ciudad Juarez WNV consensus sequences, 2005-2010
(Continued)

Gene*	Position	NY99†	Group 1			Group 2			Group 3		
			TX7465	TX5282	TX5718	TX6866	TX5947	TX6115	TX2686	TX6745	TX6572
NS4A	85	A	.	.	.	T	T	T	T	T	T
	135	V	.	M	M
NS4B	240	I	M	.	.
	249	E	G
NS5	11	V	.	.	.	A
	44	R	K
	202	Y	.	.	.	F	F	F	.	.	.
	314	K	R	R	R	.	.	.	R	R	R
	374	Y	.	.	C
	414	M	.	.	I
	560	D	E	.	.	.
	860	A	T	T

*C, capsid; prM, pre-membrane; E, envelope; NS, nonstructural.

†All amino acid substitutions indicated relative to the prototype NY99-flamingo382-99 strain (NY99) [AF196835]. Dots “.” indicate no difference from the NY99 consensus sequence. Shaded regions, conserved amino acid substitutions within Groups 1-3.

Figure 3-1: Bayesian-inferred 70% majority-rule phylogenetic trees of A) all published, full-length North American WNV isolates, 1999-2010; B) Group 1 isolates in the NA/WN02 genotype; and C) SW/WN03 genotype isolates. Boldface red, isolates sequenced in this chapter.

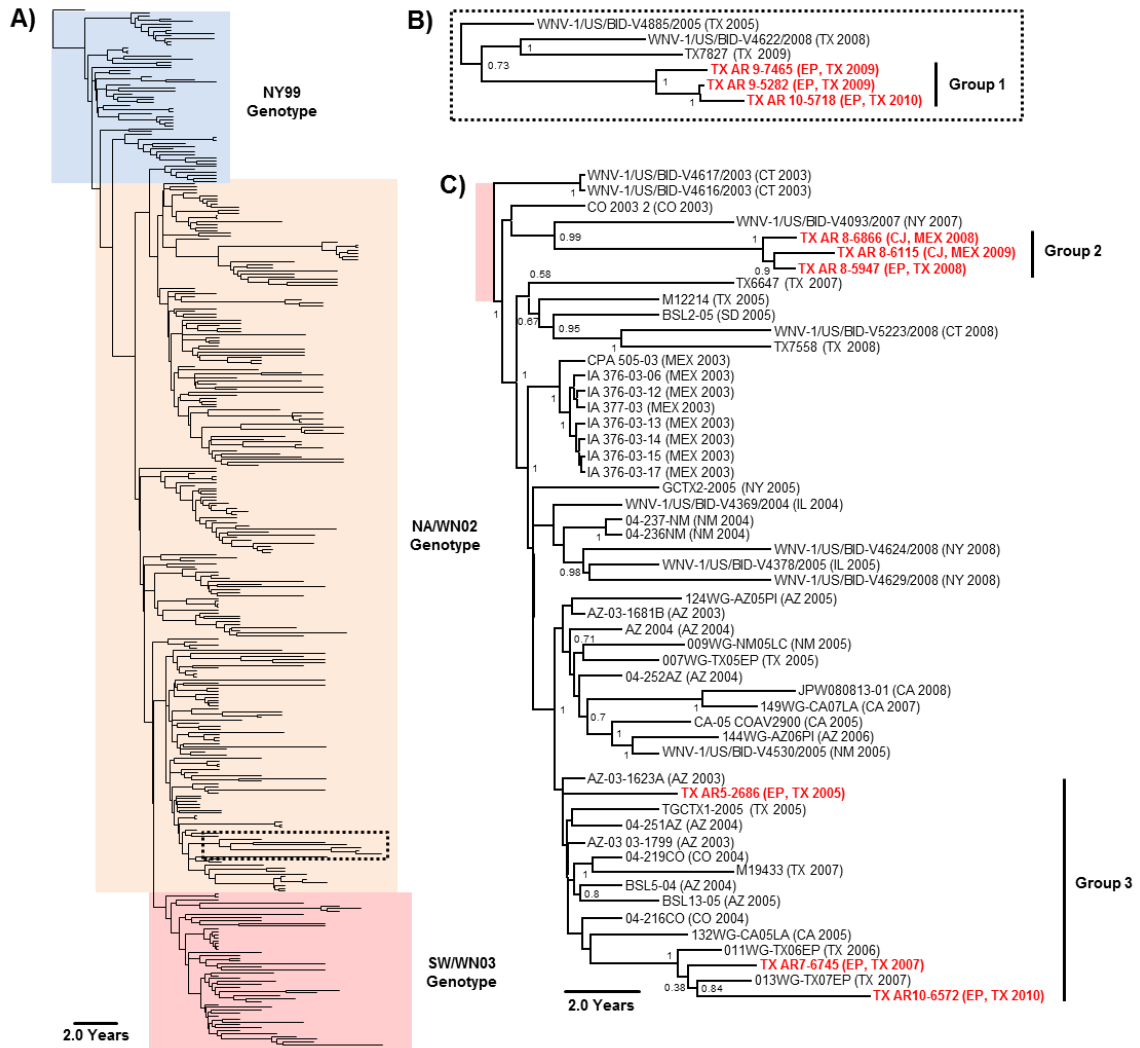
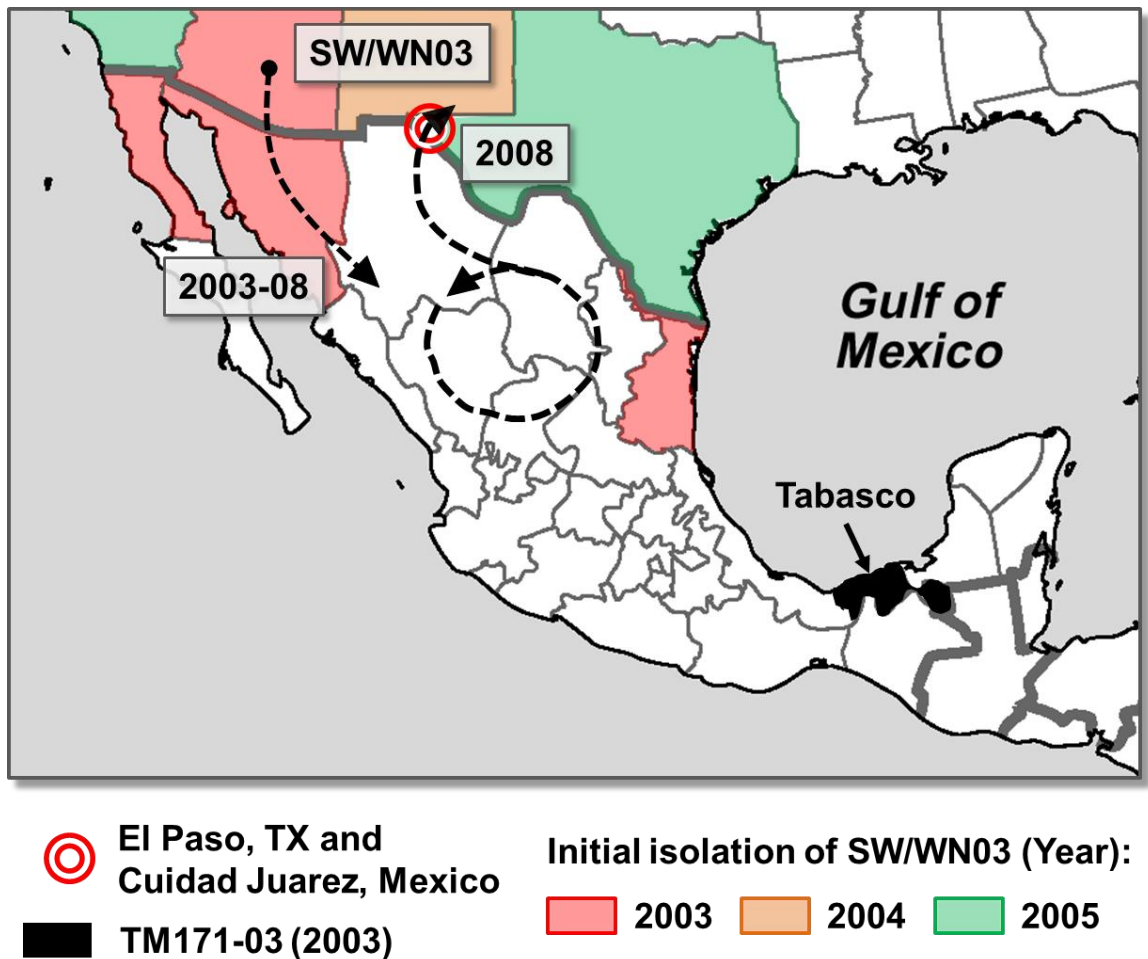


Figure 3-2: West Nile virus reintroduction model on the U.S.-Mexican border. Dashed arrows, transmission of the SW/WN03 genotype from the southwestern U.S. to northern Mexico between 2003-2008 with U.S. reintroduction in 2008³



³ Imaged adapted from “North America Political Outline Map” image at <http://www.travelsworlds.com/north-america-political-outline-map.html/mexico-physical-map-blanknorth-america-political-outline-map-vpjzbowj>. Cited on: September 30, 2014.

CHAPTER 4

Continued Evolution of WNV, Houston, Texas, USA, 2002-2012⁴

4.1 INTRODUCTION

Following introduction of WNV into the U.S. in 1999, local transmission of the original NY99 genotype in resident *Culex* spp. mosquito and wild bird populations fueled the geographic expansion of WNV from the northeast across the continental U.S., north into Canada, and south into Central and South America^{42,43,49-51,53,344}. Subsequent introduction into Texas in 2002 resulted in 105 confirmed human infections, high mortality rates among local passeriformes and corvids, and a 31.2% seroconversion rate among resident birds of Harris Co., Texas (Houston metropolitan area) alone³⁴⁵. Uninterrupted SLEV surveillance in local *Culex* (Cx.) spp. mosquito populations by the Harris Co. Public Health and Environmental Services: Mosquito Control Division since 1964 provided an ideal infrastructure for expanded WNV detection in the mosquito vector and wild bird reservoir on a major bird migratory flyway^{345,346}. Routine collections of WNV-positive birds and mosquito pools have provided an established platform to investigate WNV diversity and evolution on a fine-geographic scale (*i.e.*, Houston) comparable with similar surveillance programs in Midwestern and northeastern U.S. foci^{42,44,59,68,69,347,348}; however, Harris Co., TX represents a distinct ecosystem from these prior studies, namely a warm year-round climate with unique resident mosquito and avian species.

⁴ Emerging Infectious Diseases is published by the Centers for Disease Control and Prevention (CDC), a U.S. Government agency. Therefore, all materials published in Emerging Infectious Diseases are in the public domain and can be used without permission. Proper citation, however, is required. This citation is: **Mann, B.R., McMullen, A.R., Swetnam, D.S., Salvato, V., Reyna, M., Guzman, H., Bueno, R., Jr., Dennett, J.A., Tesh, R.B., Barrett, A.D.T. (2013)** Continued evolution of West Nile virus, Houston, Texas, USA, 2002-2012. *Emerg Infect Dis* **19**(9), 1418-1427.

Previous studies of characterized WNV isolates collected in 2002-2009 Harris Co. vector-surveillance operations confirmed the rapid displacement of the NY99 genotype with the novel NA/WN02 genotype, in 2002, and subsequent emergence of the SW/WN03 genotype in the southwestern U.S., in 2003^{55,58,63}. Currently, both the NA/WN02 and SW/WN03 genotypes still appear to co-circulate; however, endemic transmission of WNV in the U.S. since 2006 has demonstrated a dramatic decline in confirmed, clinical WNV incidence with <1100 human cases reported during 2008-2011⁴⁸. Furthermore, despite identification of regional, heterogeneous WNV populations, a relative stasis in WNV evolution has been observed in Harris Co. consistent with the logistic molecular clock model and decreasing viral growth rate proposed on a national scale^{60-62,68}. The 2012 WNV transmission season represented a major departure from this *status quo* with >5,600 human infections reported nation-wide⁴⁸. Incidence of clinical WNV disease in Texas alone accounted for 32.9% of total, reported U.S. cases—including 844 WNND cases and 89 deaths—with >994 confirmed cases in the greater Dallas/Fort Worth, Texas metropolitan area and 101 cases in Harris Co. alone^{48,349}.

Therefore, in this chapter, inclusion of 14 novel, full-length 2010-2012 Harris Co. isolates and three 2012 Collin/Denton Co. (Dallas/Fort Worth, TX) isolates together with published 2002-2009 phylogenetic studies provided an ideal model to elucidate critical, genotypic determinants contributing to divergent WNV incidence in the 2012 Texas outbreak. Using this platform, termed the Harris Co. paradigm, fine-scale phylogenetic and geospatial trends identified distinct genetic patterns linked to WNV evolution on a local, regional, and national scale.

4.2 RESULTS

4.2.1 West Nile virus collection in Harris Co., Texas

In total, 14 WNV isolates—cultured from WNV-positive, dead bird brain tissue—were obtained from the WRCEVA at UTMB; all applied isolates were collected during 2010-2012 routine vector-surveillance efforts in Harris Co., TX by the Harris Co. Public Health and Environmental Services: Mosquito Control Division (**Table 4-1**). Two isolates were collected from house sparrows (*Passer domesticus*)—one each in 2010 and 2011—with 12 additional isolates obtained from WNV-positive birds collected in 2012: 10 from blue jays (*Cyanocitta cristata*) and one each from a loggerhead shrike (*Lanius ludovicianus*) and house sparrow.

In all applied Harris Co. phylogenetic and geospatial analyses, all 14 2010-2012 WNV isolates were combined with 28 additional, published Harris Co. isolates collected during 2002-2009 from *Culex* (Cx.) or *Aedes* (Ae.) spp. mosquito pools ($n = 12$), birds ($n = 13$), or clinical human samples ($n = 3$)^{57,58,63}. Genomic sequences were determined for WNV isolates from each year (except 2004): 2002 ($n = 3$), 2003 ($n = 8$), 2005 ($n = 2$), 2006 ($n = 3$), 2007 ($n = 4$), 2008 ($n = 1$), and 2009 ($n = 7$). Sample coverage was restricted on the basis of sequence and sample availability. Furthermore, the Kuritz strain (TVP8533; AY289214) was included as a common Texas outgroup (southeast coastal Texas genotype), and the TX114 isolate (Bird114 from June 2002; GU27998) served as the prototypic member of the NA/WN02 genotype.

4.2.2 Divergent WNV evolution in Harris Co., Texas, 2010-2012

Full-length, genomic consensus sequences were obtained for all 14 novel 2010-2012 Harris Co. isolates using over-lapping, primer-based RT-PCR of vRNA extracted from Vero(p1) working stocks—*i.e.*, Vero[p2] overall as isolations were conducted in Vero cell culture—with the QIAmp Viral RNA Mini kit (Qiagen). Sanger sequencing of resulting cDNA fragments produced output contig sequences aligned and edited (relative to the prototype NY99-flamingo382-99 [NY99; AF196835] U.S. strain) in ContigExpress of the VectorNTI Advance® v11 suite (Life Technologies, Grand Island, NY, US)^{43,58,63}.

Sequences were assembled in BioEdit v7.0.9.0 using the encoded 10,299 nt open reading frame (ORF) for all 42 2002-2012 Harris Co. isolates and aligned via the MUSCLE algorithm on the EMBL-EBI server³⁰⁷⁻³⁰⁹.

4.2.2.1 NUCLEOTIDE CHANGES

Initial sequence comparisons between the 14 2010-2012 Harris Co. isolates and the prototype NY99 strain identified 45-79 nucleotide (nt) differences (0.41-0.72%) per 11,029 nt genome; furthermore, each isolate encodes 8 of the 13 nt changes characteristic of the NA/WN02 genotype⁵⁸. In addition, novel U to C transitions at positions 7015 in the NS4B gene and 8811 in the NS5 gene remained conserved in 11 of the 12 Harris Co. isolates from 2012 (excluding TX8604).

4.2.2.2 AMINO ACID SUBSTITUTIONS

Overall, the 14 2010-2012 Harris Co. isolates differed at 48 unique residues in the encoded 3,433 aa polyprotein (relative to NY99) with 2-10 substitutions (0.06-0.29%) per isolate (**Table 4-2**). Each isolate encoded the E-V159A substitution characteristic of the NA/WN02 genotype⁵⁸; however, the signature NS4A-A85T SW/WN03 genotype substitution was confirmed in the 2011 TX8349 isolate only⁶³. Furthermore, 11 of the 48 deduced amino acid substitutions were conserved in >1 isolate.

Based on these trends, single-likelihood ancestor counting (SLAC), fixed effect likelihood (FEL), and internal branches FEL (iFEL) methods on the Datamonkey server were applied to screen for potential site-specific positive selection—defined as $d_N > d_S$ (p -value < 0.05) in >1 method³¹⁹⁻³²². In effect, potential positive selection was identified in all applied methods at position H119 (H119Y) in the NS2A protein for the 2012 TX8546 isolate alone.

4.2.2.3 PHYLOGENETIC ANALYSES

Reconstruction of ancestral topologies among all 42 published 2002-2012 Harris Co. WNV isolates applied neighbor-joining (NJ), maximum likelihood (ML), and relaxed-clock Bayesian coalescent methods in order to confirm phylogenetic relationships. In brief, applied NJ methods were processed in Seaview v4.3.0 using the HKY-85 substitution model and 10,000 bootstrap replicates³¹³. Maximum likelihood analyses utilized the GTR+I+ Γ substitution model and 1,000 bootstrap replicates via RAxML-HPC Blackbox v7.3.2 on the CIPRES Science Gateway v3.1 teragrid server³¹⁴⁻³¹⁶. Bayesian-inferred coalescent phylogenies were produced in BEAST v1.6.2 (<http://beast.bio.ed.ac.uk>) using the GTR+I+ Γ_4 substitution model with applied taxa dates, uncorrelated lognormal relaxed clock model, and Bayesian Skyline prior constraints; output .log and .tre files were down-sampled from triplicate 50 million-state runs using both LogCombiner v1.7.4 and TreeAnnotator v1.7.4 in the BEAST v1.6.2 program suite^{317,318}. Inferred phylogenetic trees were then edited in FigTree v1.3.1 (www.mybiosoftware.com/phylogenetic-analysis/2407).

Based on this approach, all three methods produced consistent tree topologies rooted to the common IS-98 STD strain (AF481864) outgroup⁵⁴. Inclusion of the 2010-2012 isolates retained the phylogenetic clustering of the 2002-2009 Harris Co. isolates in the six monophyletic lineages (Groups 1-6) proposed by McMullen *et al.* (**Fig. 4-1**)⁶³. In addition, conserved nt and aa divergence among the 14 2010-2012 isolates indicated the emergence of 4 novel, monophyletic lineages of avian isolates (Groups 7-10) (**Table 4-2**); furthermore, using a Bonferroni-corrected Fischer's exact test ($\alpha = 0.05$), Groups 7-10 demonstrated a significant (p -value ≤ 0.044) relationship between year-of-collection and identified monophyletic lineage with more 2012 isolates clustering within Groups 8-10.

Group 7 (0.57% nt divergence from NY99) included the 2009 TX7827 and 2010 TX8092 isolates. Group 8 consisted of the 2003 TX1461 (Bird1461) outgroup and three

2012 isolates: TX8560, TX8571, and TX8599 (0.01-0.04%). Group 9 (0.40-0.45%) included three 2012 isolates: TX8546, TX8559, and TX8567. Group 10 included the 2006 TX6276 isolate and 2012 TX8551, TX8562, TX8589, and TX8590 isolates (0.00-0.61%). Overall, the 37 Harris Co. isolates in Groups 1-10 differed at 80 nt positions and 27 shared substitutions with >1 isolate (**Table 4-3**). In particular, five substitutions: prM-V156I, NS2A-M90V, NS2A-L95F, NS4B-I240M, and NS4B-E249G were conserved in >1 phylogenetic group. Furthermore, the signature NS4A-A85T SW/WN03 genotype substitution was identified in the 2011 TX8349 and other Group 5 isolates alone. Five additional 2005-2012 Harris Co. isolates clustered as outliers to Groups 1-10; in particular, both the 2012 TX8572 and TX8604 isolates were outliers to Groups 8-10 (>0.44% nt divergence) with increased divergence between these two isolates (1.13% nt and 0.38% aa).

4.2.2.4 GEOSPATIAL COMPARISONS

Superimposition of these phylogenetic groupings for 38 of the 42 Harris Co. isolates on a vector-borne WNV incidence map of Harris Co.—provided by the Harris Co. Public Health and Environmental Services: Mosquito Control Division (based on reported collection information in the 268 mosquito control operational areas)—highlighted the shift in WNV circulation patterns across the greater Houston region over the past decade (**Fig. 4-2**). Four isolates were excluded due to unknown collection location. Based on this information, mosquito pool isolates (M1, M2, M3, etc.) demonstrated robust genetic homogeneity limited to nine of the 268 operational areas suggesting WNV over-seasoning in resident mosquito populations. In particular, isolates M1 (2005) and M3 (2007) clustered within Group 5 (SW/WN03 genotype), indicating transmission of the same virus strain in Houston across multiple years. In contrast, bird-derived isolates (B1, B2, etc.) illustrated wide-spread incidence of similar genetic

signatures in 2002-2011 isolates—highlighted in dispersal of several 2005-2011 Group 5 isolates. Group 7-10 isolates demonstrated similar geographic distributions but limited monophyletic support with Group 1-6 isolates; furthermore, Group 8 isolates remained geographically restricted compared to the more pervasive Group 10 isolates. In conclusion, both phylogenetic and geospatial analysis of the novel 2012 WNV isolates indicated a closer genetic relationship with 2003-2006 Harris Co. strains than more recent 2007-2011 isolates.

4.2.3 Texas 2012 WNV outbreak: Novel WNV introduction event

Inclusion of the 14 2010-2012 Harris Co. isolates in an additional phylogenetic analysis with 358 published North American isolates provided an *in silico* platform to link the potential role of active virus transmission in North America on the recent WNV evolution dynamics observed in Harris Co., Texas. NJ, ML, and Bayesian relaxed clock methods (as described above) all produced consistent overall tree topologies with retention of the published NY99, NA/WN02, and SW/WN03 genotypes (**Fig. 4-3A**)^{58,63}. Groups 7-10 exhibited conserved clustering within the NA/WN02 genotype and robust (≥ 0.90 posterior probabilities) monophyletic support for shared lineage with several 2006-2009 New York (NY) and Connecticut (CT) isolates (**Fig. 4-3, panels B-F**); furthermore, the 2010 TX8092 isolate demonstrated consistent monophyletic clustering within the NA/WN02 genotype with the 2009 TX7827 and additional 2008 NY (WNV-1/US/BID-v4622/2008) isolates (**Fig. 4-3G**). In contrast, the novel 2011 TX8349 Harris Co. isolate retained topologic distribution within the SW/WN03 genotype. Outside of the identified Group 7-10 monophyletic clusters, all inferred basal node topologies exhibited poor statistical support (≤ 0.70 posterior probabilities) within the NA/WN02 genotype.

Principal support for the identified monophyletic lineages is based on limited nt divergence ($< 0.65\%$) and retention of unique aa substitutions between 2010 and 2012 Harris Co. isolates and those from the northeastern U.S. compared with published isolates

from Texas and the surrounding southwestern U.S.; in particular, the Group 9 lineage (**Fig. 4-3E**) encoded several conserved substitutions: NS2A-T52I, NS2A-L95F, NS3-S334T, and NS4B-S14I with a single NS2A-H119Y substitution shared with the 2008 NY (WNV-1/US/BID/v4097/2008) outgroup and the 2012 TX8546 isolate. Each isolate also encoded the signature E-V159A substitution with the conserved absence of the NS4A-A85T and NS5-K314R substitutions supporting monophyletic distribution and ancestral lineage within the NA/WN02 versus either the NY99 and/or SW/WN03 genotypes.

4.2.4 Harris Co. paradigm—Model for WNV evolution in the U.S.

In order to confirm application of this proposed Harris Co. paradigm outside the greater Houston metropolitan region, three additional WNV isolates, collected from *Cx.* spp. mosquito pools in the recent 2012 Dallas/Fort Worth, TX outbreak, were obtained from the Texas State Department of Health Services (**Table 4-1**). Following Sanger sequencing, all three isolates were applied to Bayesian phylogenetic analysis with all 42 2002-2012 Harris Co. isolates (using aforementioned methodologies). Both Collin Co. isolates: TX AR12-1648 and TX AR12-10674 exhibited limited nt divergence ($\leq 0.30\%$) and robust monophyletic clustering (≥ 0.98 posterior probabilities) with several Group 9 and Group 10 isolates in the Harris Co. paradigm (**Fig. 4-4**). Furthermore, the TX AR12-1468 Denton Co. isolate demonstrated shared monophyletic lineage with the non-clustering TX8572 isolate. In conclusion, these results supported application of the Harris Co. paradigm as a relevant *in silico* model for WNV evolution in Texas and, possibly, on a national scale.

4.3 DISCUSSION

Surveillance of WNV transmission in Harris Co., Texas provides an ideal model for elucidating genotypic dynamics of endemic and epidemic WNV evolution on a fine geographic scale. The southeastern coastal region of Texas serves as a temporary roosting site on the major Central and Mississippi flyways for migratory birds in transit between more temperate and tropical regions of the Americas; consequently, in addition to resident bird populations, this region hosts more avian species (and possible WNV wild bird hosts) than anywhere else in the U.S.³⁴⁶. Prior applications of the proposed Harris Co. paradigm confirmed the emergence of both the NA/WN02 and SW/WN03 U.S. genotypes in 2002 and 2003, respectively, as a surrogate model for WNV evolution on a national scale^{58,63}. In this chapter, inclusion of 14 novel 2010-2012 Harris Co. isolates in this paradigm supported the emergence of four monophyletic lineages (Groups 7-10) distinct from prior established Group 1-6 phylogenetic clusters (**Fig. 4-1**)⁶³. Furthermore, these isolates exhibited closer ancestral lineage with 2002-2003 Harris Co. isolates compared to more recent 2007-2009 strains; however, despite confirmed circulation of the SW/WN03 genotype (Group 5) in the southwestern U.S. during 2003-2009, the 2011 TX8349 isolate alone clustered with several 2005-2008 Group 5 isolates with the conserved expression of the signature NS4A-A85T substitution.

Restriction of several, conserved substitutions to individual monophyletic groups highlighted increased genetic heterogeneity among circulating 2010-2012 isolates than in earlier 2002-2009 Harris Co. WNV populations. Geographic reconstruction of all applied Harris Co. isolates ($n = 42$)—based on mosquito control operational areas—supported a heterogeneous transmission model with limited correlation between fine-scale geographic dispersion and sequence divergence over time (**Fig. 4-2**)^{68,347}; however, regional phylogenetic foci among Group 2, 4, and 5 *Culex* and *Aedes* spp. mosquito pools and Group 8 avian isolates reinforced potential homogeneous and trans-seasonal WNV transmission in localized vector populations published in other studies^{42,44,59,69}.

Prior to the results of this chapter, a relative stasis in WNV evolution had been observed in Harris Co. and the northeastern U.S. following the emergence and confirmed co-circulation of the NA/WN02 and SW/WN03 genotypes^{59-62,69}; in particular, fine-scale geographic phylogenetic analyses in suburban Chicago and Illinois identified the maintenance and active evolution of heterogeneous WNV populations in local, resident mosquito and bird vectors from the initial 1999 U.S. introduction until 2008, but there has been no published results more recently^{68,347,348}. Based on these trends, applied phylogenetic comparisons between all 14 2010-2012 Harris Co. isolates highlighted the significant divergence of recent local WNV populations from reported, historical dynamics with the emergence of three distinct monophyletic lineages (Groups 8-10) in the 2012 outbreak alone (**Fig. 4-1**). Furthermore, despite shared geographic distribution and environmental conditions, the 2010 and 2012 Harris Co. isolates exhibited >0.42% sequence divergence relative to Group 1-6 isolates; in addition, co-circulation of these novel genetic signatures was confirmed in *Culex* spp. mosquito pools ($n = 3$) collected in the greater Dallas/Fort Worth region during the 2012 Texas WNV outbreak. Based on these observations, introduction of a novel or existing U.S. strain into circulating greater Houston WNV populations—and possibly Texas as a whole—since 2010 offers an alternative mechanism driving emergence of these divergent genetic signatures; however, these results do not exclude possible emergence of related existent virus populations as dominant regional strains in the recent epidemic WNV transmission season.

In order to test these hypotheses, 14 2010-2012 Harris Co. isolates were included in a comprehensive Bayesian phylogenetic analysis with 358 published North American WNV isolates (**Fig. 4-3**). Shared lineage of Group 7-10 isolates with several 2006-2009 northeastern U.S. strains within the NA/WN02 genotype was identified with robust monophyletic support (≥ 0.90 posterior probabilities) (**Fig. 4-3, panels B-G**). Limited sequence divergence ($\leq 0.65\%$) and conserved expression of unique aa substitutions relative to 0.8-1.0% observed genetic diversity between published Group 1-6 isolates

from the southwestern U.S. provided principal evidence in support of these monophyletic lineages⁶³.

Based on these results, the 2010 and 2012 Harris Co. isolates exhibited inferred lineage from ancestral strains circulating in the northeastern U.S. consistent with a single or multiple introduction event(s) during 2010-2012 in the greater Houston region. However, poor statistical confidence (≤ 0.70 posterior probabilities) for all inferred basal node topologies in the NA/WN02 genotype limits direct comparison of independent evolution between Harris Co. isolates within different monophyletic lineages (**Fig. 4-3**); furthermore, consistent phylogenetic clustering of the 2011 TX8349 isolate in the SW/WN03 genotype (Group 5) indicated co-circulation of this genotype in Harris Co. until the 2012 transmission season. However, limited clustering of sequenced 2010-2012 Harris Co. isolates within the SW/WN03 genotype may be an artifact of inherent biases in sample collection.

Sampling bias is a recognized constraint in phylogenetic and paired geospatial analyses. Unfortunately, the predominance of WNV surveillance has remained restricted to few regional foci; of the 372 full-length WNV isolates applied in this chapter, 31.5% ($n = 117$) were collected in CT, 22.0% ($n = 82$) in NY, 16.4% ($n = 61$) in TX, and 8.1% ($n = 30$) in Illinois during 1999-2012. In contrast, a single isolate has been published for each of the mid-western/central Michigan, North Dakota, and South Dakota U.S. states—all of which reported increased incidence of clinical, human WNV disease ($n \geq 89$ cases) in the 2012 epidemic season not observed since 2007⁴⁸. Based on these biases, increased local and regional WNV vector-surveillance efforts, including collection of isolates, are needed to provide critical, unbiased insight into the underlying dynamics driving WNV evolution and transmission in U.S. host populations.

In effect, the Harris Co. Public Health and Environmental Services: Mosquito Control Division, founded in response to a 1964 SLEV outbreak, has provided a model infrastructure for WNV surveillance in resident mosquito and bird populations in the

greater Houston region³⁴⁵. Uninterrupted collection/processing of WNV-positive birds and mosquito pools following the introduction of WNV into Harris Co. in 2002 has sustained a conduit for the scientific investigation of real-time disease outbreaks with the direct translation of findings towards optimized vector-borne disease control and prevention. Incorporation of this paradigm in public health directives across the U.S. would provide a proactive approach towards detection and response to clinical outbreak scenarios of endemic and exotic pathogens in the U.S.

In conclusion, application of on-going vector-surveillance efforts in Harris Co. from 2002-2012 highlighted periodic shifts in the genotypic signatures of circulating WNV populations as a comprehensive *in silico* model for WNV evolution—designated the Harris Co. paradigm. Texas represents a unique, surrogate geographic model to sample viral gene-flow across the U.S. from migrant bird and local ornithophilic *Culex* spp. mosquito populations; in this regard, previous *in silico* phylogenetic applications of this proposed paradigm confirmed the local retention of both regional (*e.g.*, southwestern and northeastern U.S.) and national genotypic trends—supporting annual representative, regional WNV populations in Harris Co., TX^{55,58,63}. However, (since 2004) active surveillance has been restricted to resident mosquito populations with isolation of bird-derived WNV isolates from reported dead-birds only. Isolation of all 14 2010-2012 Harris Co. isolates in this chapter from birds biases detection to bird-selected WNV populations; however, retention of these detected *in silico* relationships with three additional *Cx.* spp. isolates from Collin/Denton Co., Texas support extrapolation of the Harris Co. paradigm to both bird and mosquito hosts in the 2012 Texas outbreak. Overall, continued WNV surveillance is needed to validate the local, regional, and national circulation of the annual genetic shifts identified in the Harris Co. paradigm as a potential, national model for WNV evolution in the U.S.

Table 4-1: West Nile virus isolates described in the conducted Harris Co., Texas and U.S. sequence and phylogenetic analyses, 1998-2012

Strain*	Map Code†	Location	Zip Code‡	Source§	Collection Year	GenBank Accession No.
IS-98 STD	-	Eilat, Israel	-	White Stork	1998	AF481864
NY99 (382-99)	-	New York, NY	-	Chilean Flamingo	1999	AF196835
Kuritz	-	Beaumont, TX	-	Human	2002	AY289214
TX114	B1-1	Harris Co., TX	77043	Blue Jay	2002	GU827998
TX 2002 1	-	Harris Co., TX	-	Human	2002	DQ164198
TX 2002 2	-	Harris Co., TX	-	Human	2002	DQ164205
TX1153	B2-1	Harris Co., TX	77077	Mourning Dove	2003	AY712945
TX1171	B3-1	Harris Co., TX	77030	Blue Jay	2003	AY712946
TX1175	B4-3	Harris Co., TX	77346	Blue Jay	2003	GU828000
TX1461	-	Harris Co., TX	-	Avian	2003	AY712947
TX 2003	-	Harris Co., TX	-	Human	2003	DQ164199
v4095	M10-2	Harris Co., TX	77093	<i>Cx. quinquefasciatus</i>	2003	GU828002
v4369	M11-2	Harris Co., TX	77039	<i>Cx. quinquefasciatus</i>	2003	AY712948
v4380	M12-2	Harris Co., TX	77093	<i>Cx. quinquefasciatus</i>	2003	GU828001
M12214	M1-5	Harris Co., TX	77020	<i>Cx. quinquefasciatus</i>	2005	JF415914
TX5058	B5	Harris Co., TX	77057	Blue Jay	2005	JF415929
M6019	M2-6	Harris Co., TX	77026	<i>Cx. quinquefasciatus</i>	2006	JF415930
TX5810	B6-6	Harris Co., TX	77345	Common Grackle	2006	JF415915
TX6276	B7-10	Harris Co., TX	77373	Northern Mockingbird	2006	JF415916
M19433	M3-5	Harris Co., TX	77020	<i>Cx. quinquefasciatus</i>	2007	JF415919
TX6647	B8-5	Harris Co., TX	77084	Blue Jay	2007	JF415917
TX6747	B9	Harris Co., TX	77346	Blue Jay	2007	JF415918
TX7191	B10	Harris Co., TX	77005	Blue Jay	2007	JF415920
TX7558	B11-5	Harris Co., TX	77375	Blue Jay	2008	JF415921
M20122	M4-4	Harris Co., TX	77026	<i>Ae. albopictus</i>	2009	JF415928
M20140	M5-4	Harris Co., TX	77021	<i>Ae. albopictus</i>	2009	JF415926
M20141	M6-4	Harris Co., TX	77021	<i>Cx. quinquefasciatus</i>	2009	JF415927
M37012	M7-4	Harris Co., TX	77021	<i>Cx. quinquefasciatus</i>	2009	JF415922
M37906	M8-4	Harris Co., TX	77021	<i>Cx. quinquefasciatus</i>	2009	JF415923
M38488	M9-4	Harris Co., TX	77004	<i>Ae. albopictus</i>	2009	JF415925
TX7827	B12-7	Harris Co., TX	77060	Blue Jay	2009	JF415924
TX8092	B13-7	Harris Co., TX	77084	House Sparrow	2010	KC333374
TX8349	B14-5	Harris Co., TX	77016	House Sparrow	2011	KC333375
TX8546	B15-9	Harris Co., TX	77065	Blue Jay	2012	KC333376
TX8551	B16-10	Harris Co., TX	77449	Blue Jay	2012	KC333377
TX8559	B17-9	Harris Co., TX	77506	Blue Jay	2012	KC333378
TX8560	B18-8	Harris Co., TX	77062	Blue Jay	2012	KC333379
TX8562	B19-10	Harris Co., TX	77450	Blue Jay	2012	KC333380
TX8567	B20-9	Harris Co., TX	77065	Blue Jay	2012	KC333381
TX8571	B21-8	Harris Co., TX	77059	Blue Jay	2012	KC333382
TX8572	B22	Harris Co., TX	77080	Blue Jay	2012	KC333383
TX8589	B23-10	Harris Co., TX	77049	Loggerhead Shrike	2012	KC333384
TX8590	B24-10	Harris Co., TX	77339	Blue Jay	2012	KC333385
TX8599	B25-8	Harris Co., TX	77058	Blue Jay	2012	KC333386
TX8604	B26	Harris Co., TX	77021	House Sparrow	2012	KC333387
TX AR12-1486	-	Denton Co., TX	-	<i>Cx. quinquefasciatus</i>	2012	KC711057
TX AR12-1648	-	Collin Co., TX	-	<i>Cx. quinquefasciatus</i>	2012	KC711058
TX AR12-10674	-	Collin Co., TX	-	<i>Cx. restuans</i>	2012	KC711059

*Boldface, strains in bold were sequenced in this chapter.

†Map code designations include B (bird) and M (mosquito) for isolates collected in Harris County (Co.), Texas with phylogenetic relationships (Groups 1-10) indicated after the hyphen (see Fig. 4-1).

‡Zip codes indicated for region of isolate collection within Harris Co., Texas where available (see Fig. 4-2).

§Mosquito (M) isolates were collected from either *Culex* (Cx.) spp. or *Aedes* (Ae.) *albopictus* mosquito pools.

Table 4-2: Detected amino acid substitutions in novel Harris Co., TX isolates, 2010-2012

		C*					M	E					NS1							
Strain†	Year	36	76	104	114	119	150	51	93	123	159	400	35	41	51	167	236	308	326	
NY99	1999	S	T	K	M	A	M	A	R	T	V	S	Y	G	K	M	I	I	D	
TX8092	2010	.	I	.	.	.	T	.	.	.	A	I	.	.	.	
TX8349	2011	K	.	A	
TX8560	2012	A	
TX8571	2012	A	
TX8599	2012	A	
TX8546	2012	A	
TX8559	2012	A	
TX8567	2012	S	.	.	.	N	A	.	.	A	.	.	.	V	.	
TX8589	2012	A	
TX8551	2012	.	.	R	A	V	.	.	
TX8562	2012	.	.	R	A	V	.	.	
TX8590	2012	.	.	R	A	F	V	.	.	
TX8572	2012	N	.	.	I	A	
TX8604	2012	T	.	.	A	.	H	.	R	.	.	.	G	

*C, capsid; prM, pre-membrane; E, envelope and NS, nonstructural.

†Amino acid changes are relative to the prototype NY99 strain [AF196836] indicated in boldface. Dots indicate no change from the NY99 isolate. Shaded regions, conserved amino acid substitutions present in >1 Harris Co. isolate.

Table 4-2: Detected amino acid substitutions in novel Harris Co., TX isolates, 2010-2012 (Continued)

Strain†	Year	NS2A*								NS2B			NS3			NS4A	
		43	52	58	90	95	119	188	224	26	103	120	180	334	549	59	85
NY99	1999	V	T	V	M	L	H	R	A	I	V	V	E	S	Y	L	A
TX8092	2010	.	.	.	V	F
TX8349	2011	I	.	.	.	F	.	.	V	I	T
TX8560	2012	D
TX8571	2012	D
TX8599	2012	V	.	.	D
TX8546	2012	.	I	.	.	F	Y	T	.	.	.
TX8559	2012	.	I	.	.	F	T	.	.	.
TX8567	2012	.	I	.	.	F	T	H	.	.
TX8589	2012	.	.	I	.	.	.	K
TX8551	2012	K	.	.	.	I
TX8562	2012	K	.	.	.	I
TX8590	2012	K	.	.	.	I
TX8572	2012
TX8604	2012

*NS, nonstructural: NS2A, NS2B, NS3, and NS4A.

†Amino acid changes are relative to the prototype NY99 strain [AF196836] indicated in boldface. Dots indicate no change from the NY99 isolate. Shaded regions, conserved amino acid substitutions present in >1 Harris Co. isolate.

Table 4-2: Detected amino acid substitutions in novel Harris Co., TX isolates, 2010-2012 (Continued)

Strain†	Year	NS4B							NS5						
		14	23	24	26	200	240	244	25	33	42	49	280	402	445
NY99	1999	S	V	K	N	I	I	L	T	I	H	V	K	I	L
TX8092	2010	.	A	.	.	V
TX8349	2011
TX8560	2012
TX8571	2012
TX8599	2012
TX8546	2012	I
TX8559	2012	I
TX8567	2012	I
TX8589	2012	M
TX8551	2012	.	.	R	I	.	.	.
TX8562	2012	.	.	R	I	.	.	.
TX8590	2012	I	.	.	M
TX8572	2012	V	I	.	.	.	E	T	.
TX8604	2012	.	.	.	T	T	Y

*NS, nonstructural: NS4B and NS5.
†Amino acid changes are relative to the prototype NY99 strain [AF196836] indicated in boldface. Dots indicate no change from the NY99 isolate. Shaded regions, conserved amino acid substitutions present in >1 Harris Co. isolate.

Table 4-3: Identified amino acid substitutions in Harris Co., TX isolates, 2002-2012

Group	Strain†	Year	C*						prM				
			32	39	76	104	109	119	5	88	122	150	156
	NY99	1999	R	D	T	K	T	A	N	R	V	M	V
1	TX114	2002
	TX1153	2003	I
	TX1171	2003	I
2	v4095	2003	D
	v4369	2003	D
	v4380	2003	D
3	TX 2003	2003
	TX1175	2003	S
4	M20122	2009
	M37906	2009
	M37012	2009
	M38488	2009
	M20140	2009	I
	M20141	2009	I
5	TX7558	2008
	TX6647	2007	I	.	.
	M12214	2005
	TX8349	2011
	M19433	2007
6	TX 2002 1	2002
	TX 2002 2	2002
	M6019	2006	.	N	K	.	.	.
	TX5810	2006
7	TX7827	2009	I
	TX8092	2010	.	.	I	T	.
8	TX1461	2003
	TX8560	2012
	TX8571	2012
	TX8599	2012
9	TX8546	2012
	TX8559	2012
	TX8567	2012	S
10	TX6276	2006
	TX8589	2012
	TX8551	2012	.	.	.	R
	TX8562	2012	.	.	.	R
	TX8590	2012	.	.	.	R

*C, capsid; prM, pre-membrane.

†Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. Dots indicate no change from the NY99 isolate. Shaded regions, substitutions conserved in >1 isolate of one or more phylogenetic groups (described in **Fig. 4-1**).

Table 4-3: Identified amino acid substitutions in Harris Co., TX isolates, 2002-2012
(Continued)

Group	Strain†	Year	E*										NS1	
			49	70	79	93	123	159	400	431	460	475	41	167
	NY99	1999	E	T	E	R	T	V	S	V	I	N	G	M
1	TX114	2002	A
	TX1153	2003	A
	TX1171	2003	A
2	v4095	2003	A
	v4369	2003	A
	v4380	2003	A
3	TX 2003	2003	A
	TX1175	2003	A
4	M20122	2009	A	.	F	L	.	.	.
	M37906	2009	A	.	.	L	.	.	.
	M37012	2009	.	I	.	.	.	A	.	.	L	.	.	.
	M38488	2009	.	I	.	.	.	A	.	.	L	.	.	.
	M20140	2009	A	.	.	L	.	.	.
	M20141	2009	A	.	.	L	.	.	.
5	TX7558	2008	A
	TX6647	2007	A
	M12214	2005	A
	TX8349	2011	.	.	.	K	.	A
	M19433	2007	A
6	TX 2002 1	2002	.	.	D	.	.	A
	TX 2002 2	2002	A
	M6019	2006	A	.	.	.	S	.	.
	TX5810	2006	A
7	TX7827	2009	A
	TX8092	2010	A	I
8	TX1461	2003	A
	TX8560	2012	A
	TX8571	2012	A
	TX8599	2012	A
9	TX8546	2012	A
	TX8559	2012	A
	TX8567	2012	N	A	A	.
10	TX6276	2006	K	A	.	.	M	.	.	.
	TX8589	2012	A
	TX8551	2012	A
	TX8562	2012	A
	TX8590	2012	A	F

*E, envelope; NS, nonstructural: NS1.

†Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. Dots indicate no change from the NY99 isolate. Shaded regions, substitutions conserved in >1 isolate of one or more phylogenetic groups (described in Fig. 4-1).

Table 4-3: Identified amino acid substitutions in Harris Co., TX isolates, 2002-2012
(Continued)

Group	Strain†	Year	NS1*		NS2A									
			236	308	43	52	58	90	95	98	102	118	119	124
	NY99	1999	I	I	V	T	V	M	L	R	Q	Y	H	I
1	TX114	2002
	TX1153	2003
	TX1171	2003
2	v4095	2003
	v4369	2003
	v4380	2003
3	TX 2003	2003
	TX1175	2003
4	M20122	2009	H	.	.
	M37906	2009
	M37012	2009
	M38488	2009
	M20140	2009
	M20141	2009
5	TX7558	2008
	TX6647	2007	V
	M12214	2005	V
	TX8349	2011	.	.	I	.	.	.	F
	M19433	2007
6	TX 2002 1	2002	H	.	.	.
	TX 2002 2	2002	H	.	.	.
	M6019	2006	G
	TX5810	2006	G
7	TX7827	2009	V
	TX8092	2010	V
8	TX1461	2003
	TX8560	2012
	TX8571	2012
	TX8599	2012
9	TX8546	2012	.	.	.	I	.	.	F	.	.	.	Y	.
	TX8559	2012	.	.	.	I	.	.	F
	TX8567	2012	.	V	.	I	.	.	F
10	TX6276	2006
	TX8589	2012	.	.	.	I
	TX8551	2012	V
	TX8562	2012	V
	TX8590	2012	V

* NS, nonstructural: NS1 and NS2A

† Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. Dots indicate no change from the NY99 isolate. Shaded regions, substitutions conserved in >1 isolate of one or more phylogenetic groups (described in Fig. 4-1).

Table 4-3: Identified amino acid substitutions in Harris Co., TX isolates, 2002-2012
(Continued)

Group	Strain†	Year	NS2A*			NS2B				NS3				
			137	188	224	26	41	103	120	106	160	180	213	328
	NY99	1999	A	R	A	I	A	V	V	V	S	E	I	E
1	TX114	2002
	TX1153	2003
	TX1171	2003
2	v4095	2003
	v4369	2003
	v4380	2003
3	TX 2003	2003
	TX1175	2003
4	M20122	2009
	M37906	2009
	M37012	2009
	M38488	2009	A	.	.	M	.
	M20140	2009	V
	M20141	2009	V
5	TX7558	2008	V
	TX6647	2007
	M12214	2005	.	.	V
	TX8349	2011	.	.	V
	M19433	2007	A	.	.	.
6	TX 2002 1	2002
	TX 2002 2	2002
	M6019	2006
	TX5810	2006
7	TX7827	2009
	TX8092	2010	F
8	TX1461	2003	D	.	K
	TX8560	2012	D	.	.
	TX8571	2012	D	.	.
	TX8599	2012	.	.	.	V	D	.	.
9	TX8546	2012
	TX8559	2012
	TX8567	2012
10	TX6276	2006
	TX8589	2012	.	K
	TX8551	2012	.	K	I
	TX8562	2012	.	K	I
	TX8590	2012	.	K	I

* NS, nonstructural: NS2A, NS2B, and NS3.

†Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. Dots indicate no change from the NY99 isolate. Shaded regions, substitutions conserved in >1 isolate of one or more phylogenetic groups (described in Fig. 4-1).

Table 4-3: Identified amino acid substitutions in Harris Co., TX isolates, 2002-2012
(Continued)

Group	Strain†	Year	NS3*					NS4A			NS4B			
			334	355	365	549	562	59	85	135	14	23	24	33
	NY99	1999	S	Y	S	Y	R	L	A	V	S	V	K	L
1	TX114	2002
	TX1153	2003
	TX1171	2003
2	v4095	2003
	v4369	2003
	v4380	2003
3	TX 2003	2003
	TX1175	2003
4	M20122	2009
	M37906	2009
	M37012	2009
	M38488	2009
	M20140	2009
	M20141	2009
5	TX7558	2008	T
	TX6647	2007	K	.	T
	M12214	2005	.	.	G	.	.	.	T
	TX8349	2011	I	T
	M19433	2007	T
6	TX 2002 1	2002
	TX 2002 2	2002
	M6019	2006	F
	TX5810	2006
7	TX7827	2009	.	H
	TX8092	2010	A	.	.
8	TX1461	2003	M
	TX8560	2012
	TX8571	2012
	TX8599	2012
9	TX8546	2012	T	I	.	.	.
	TX8559	2012	T	I	.	.	.
	TX8567	2012	T	.	.	H	I	.	.	.
10	TX6276	2006
	TX8589	2012
	TX8551	2012	R	.
	TX8562	2012	R	.
	TX8590	2012

* NS, nonstructural: NS3, NS4A, and NS4B.

†Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. Dots indicate no change from the NY99 isolate. Shaded regions, substitutions conserved in >1 isolate of one or more phylogenetic groups (described in Fig. 4-1).

Table 4-3: Identified amino acid substitutions in Harris Co., TX isolates, 2002-2012
(Continued)

Group	Strain†	Year	NS4B*						NS5					
			173	176	200	240	241	249	21	49	91	200	296	312
	NY99	1999	V	I	I	I	T	E	K	V	M	R	H	D
1	TX114	2002
	TX1153	2003	G
	TX1171	2003	G	.	.	.	L	.	.
2	v4095	2003	Y	.
	v4369	2003	A	Y	.
	v4380	2003	Y	.
3	TX 2003	2003	E
	TX1175	2003	G
4	M20122	2009	.	.	.	M
	M37906	2009	.	.	.	M
	M37012	2009	.	.	.	M
	M38488	2009	.	.	.	M
	M20140	2009	.	.	.	M
	M20141	2009	.	.	.	M
5	TX7558	2008
	TX6647	2007	R
	M12214	2005
	TX8349	2011
	M19433	2007	.	.	.	M	V	.	.	.
6	TX 2002 1	2002	.	.	.	M
	TX 2002 2	2002	I	.	.	M
	M6019	2006	.	.	.	M
	TX5810	2006	.	.	.	M
7	TX7827	2009
	TX8092	2010	.	.	V
8	TX1461	2003
	TX8560	2012
	TX8571	2012
	TX8599	2012
9	TX8546	2012
	TX8559	2012
	TX8567	2012
10	TX6276	2006	.	V	.	M
	TX8589	2012	.	.	.	M
	TX8551	2012	I
	TX8562	2012	I
	TX8590	2012	I

* NS, nonstructural: NS4B and NS5.

†Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. Dots indicate no change from the NY99 isolate. Shaded regions, substitutions conserved in >1 isolate of one or more phylogenetic groups (described in Fig. 4-1).

Table 4-3: Identified amino acid substitutions in Harris Co., TX isolates, 2002-2012
(Continued)

Group	Strain†	Year	NS5*								
			314	395	445	515	619	640	661	688	804
	NY99	1999	K	M	L	Q	A	L	S	A	A
1	TX114	2002
	TX1153	2003	V
	TX1171	2003	D	V
2	v4095	2003	.	.	.	P
	v4369	2003
	v4380	2003
3	TX 2003	2003
	TX1175	2003
4	M20122	2009
	M37906	2009	.	I
	M37012	2009
	M38488	2009
	M20140	2009
	M20141	2009
5	TX7558	2008
	TX6647	2007
	M12214	2005
	TX8349	2011
	M19433	2007	R
6	TX 2002 1	2002
	TX 2002 2	2002
	M6019	2006
	TX5810	2006	P	T	.	.
7	TX7827	2009
	TX8092	2010
8	TX1461	2003	S
	TX8560	2012
	TX8571	2012
	TX8599	2012
9	TX8546	2012
	TX8559	2012
	TX8567	2012
10	TX6276	2006
	TX8589	2012
	TX8551	2012
	TX8562	2012
	TX8590	2012	.	.	M

* NS, nonstructural: NS5.

†Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. Dots indicate no change from the NY99 isolate. Shaded regions, substitutions conserved in >1 isolate of one or more phylogenetic groups (described in **Fig. 4-1**).

Figure 4-1: Bayesian-inferred, 50% majority-rule, coalescent phylogenetic tree of published, full-length Harris Co., TX isolates, 2002-2012. Boldface, isolates sequenced in this chapter. Strain names link geographic map code (e.g. B1, B2, M1, M2)—refer to **Fig. 4-2**. Year-of-collection annotated in parentheses. Scale bar indicates divergence time (in years).

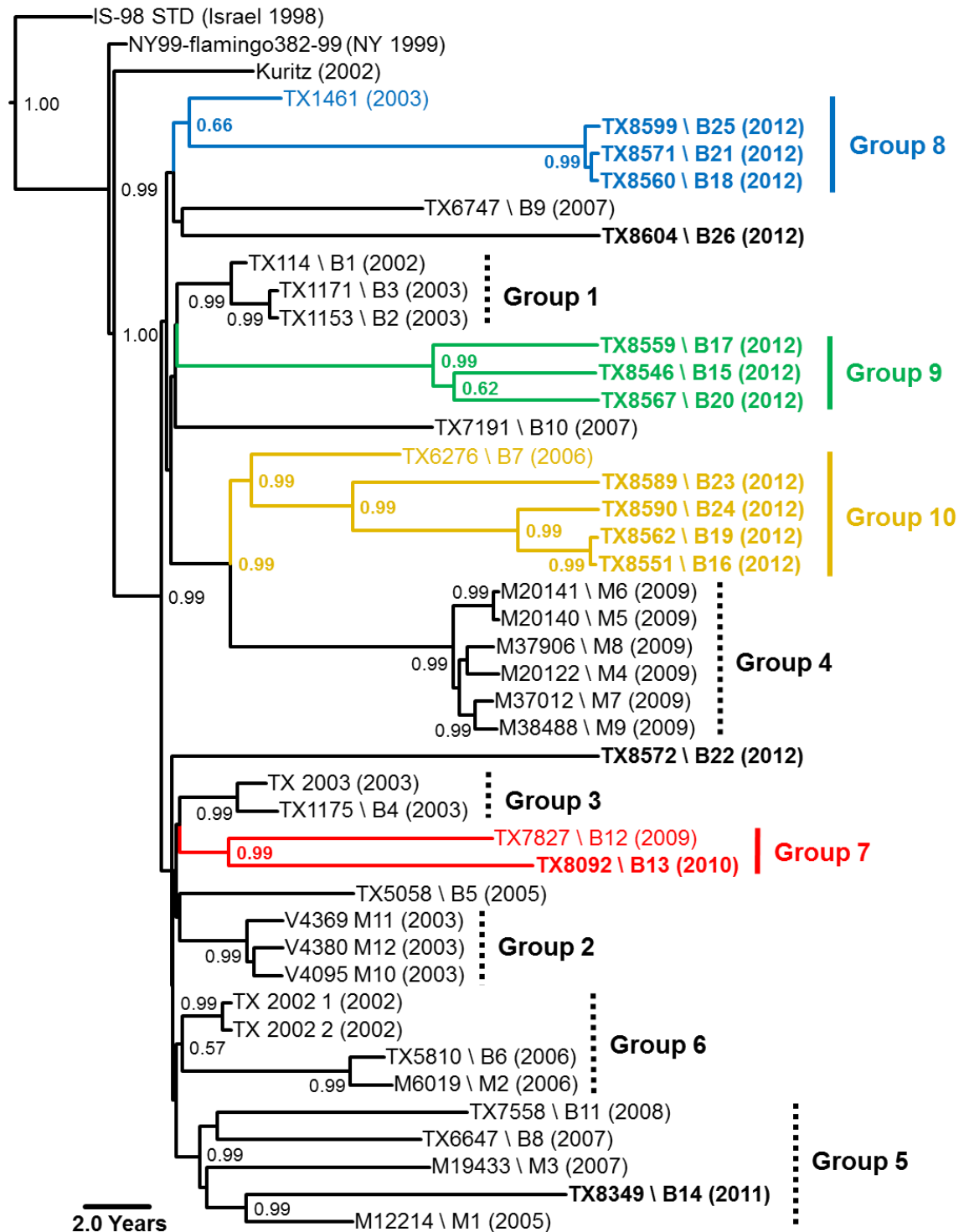


Figure 4-2: Incidence of vector-borne WNV in Harris Co., Texas, 2002-2012, indicating the cumulative distribution of confirmed avian and mosquito (*Culex* and *Aedes* spp.) isolates. Small numbers, reference codes for each of the 268 mosquito control operational areas. Colored symbols indicate 2002-2012 Harris Co. isolates designated in **Table 4-1** with open-symbols representing isolates clustering in Group 7 (red), Group 8 (blue), Group 9 (green), Group 10 (gold), or non-clustering isolates (black).

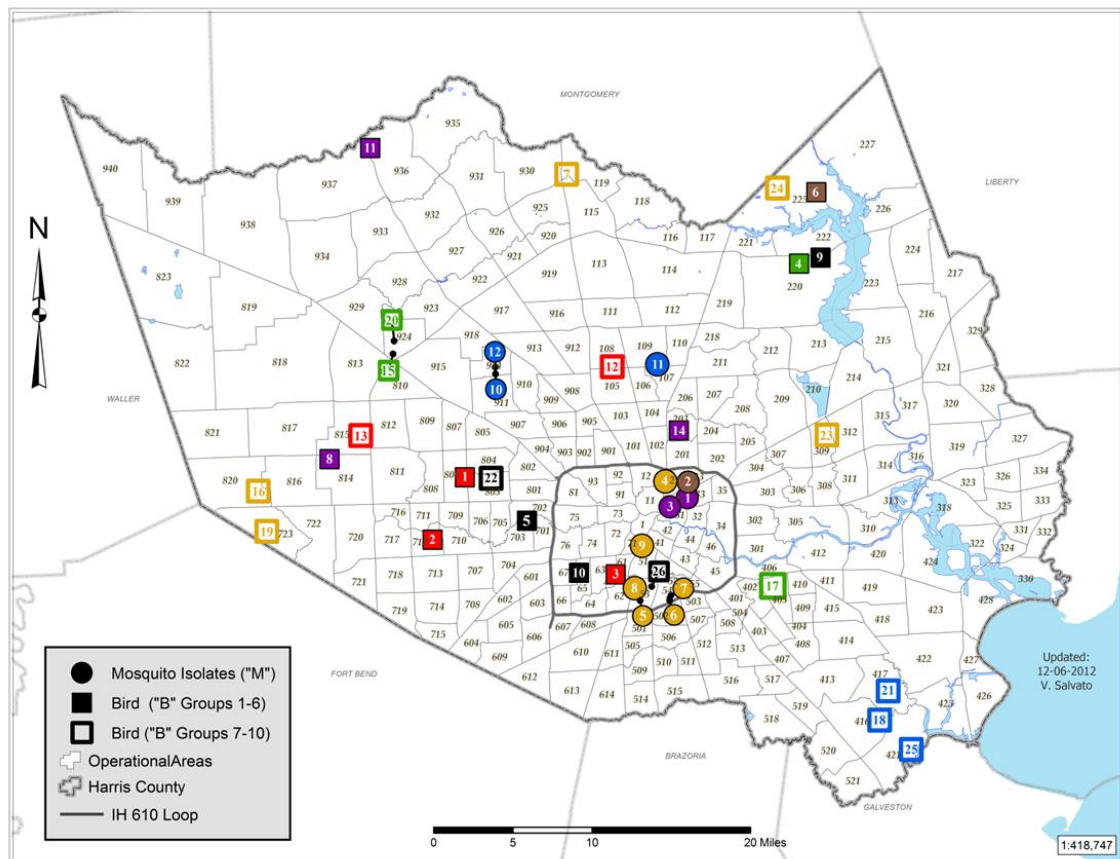


Figure 4-3: Evolution of WNV in North America, 1999-2012. Bayesian coalescent tree of A) all published North American isolates clustering in the NY99 (blue), NA/WN02 (orange), or SW/WN03 (red) U.S. genotypes containing inferred monophyletic lineages B-G of the novel 2010-2012 Harris Co. isolate (indicated in boldface) color-coded (Panels B-G) based on **Fig. 4-1** designations. Posterior probabilities (range 0.00-1.00) indicated along branches. Scale bar, divergence time (in years).

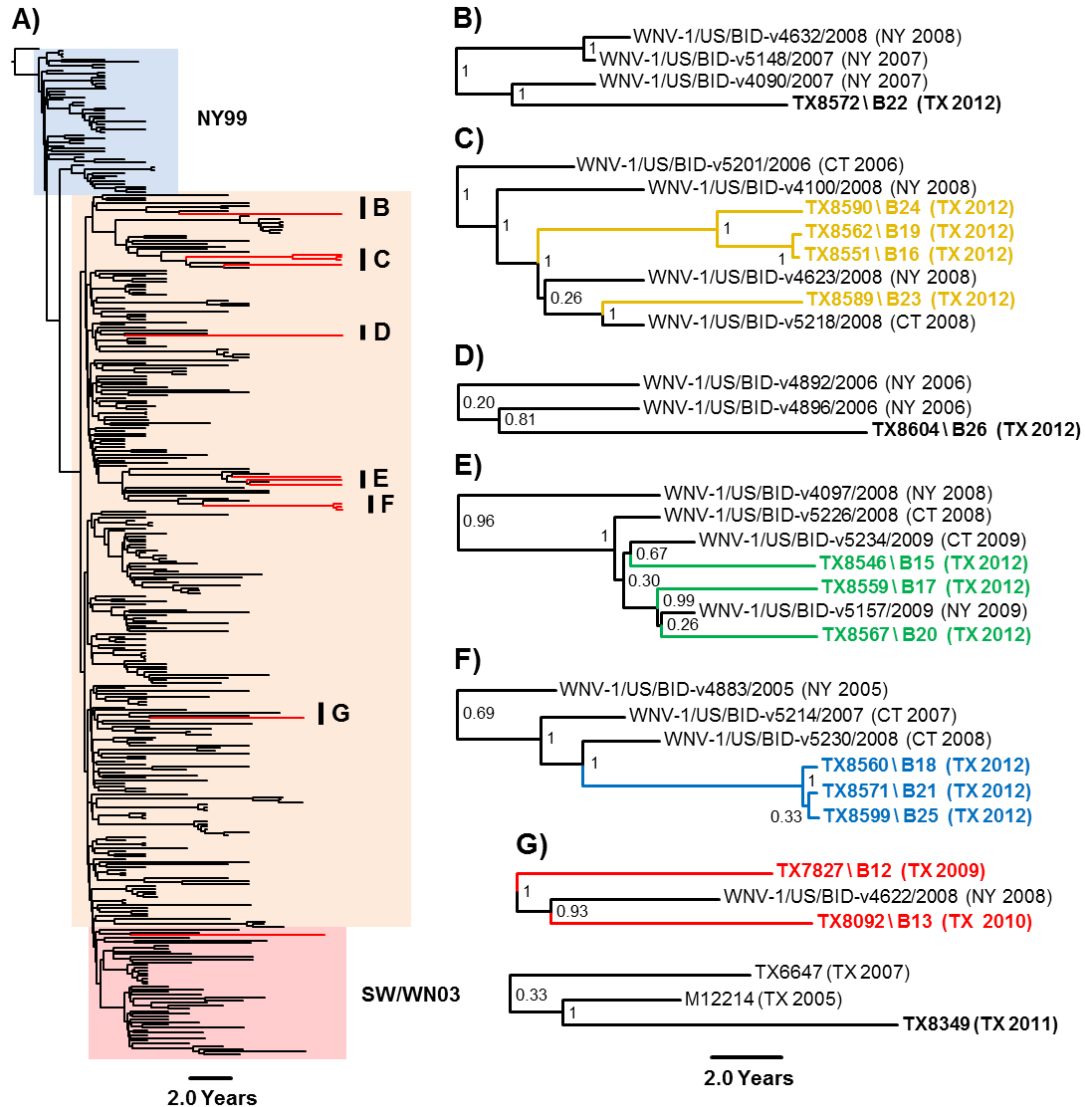
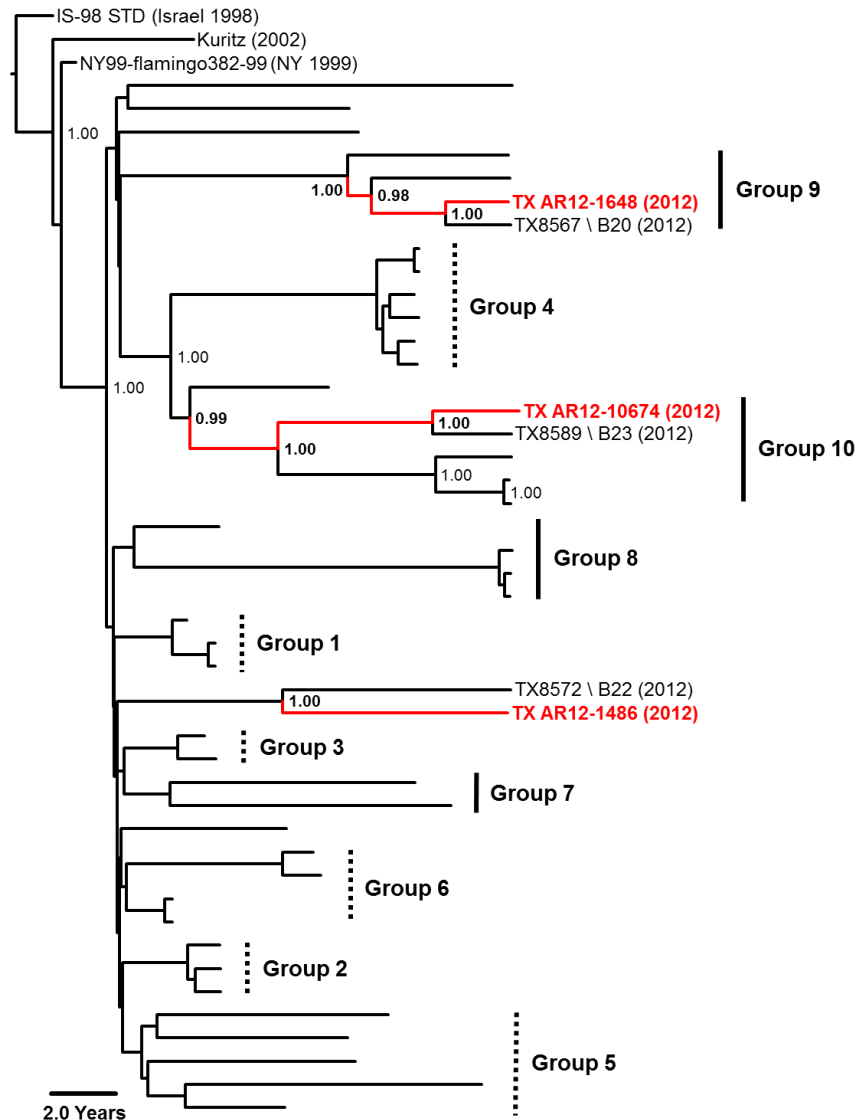


Figure 4-4: Bayesian phylogenetic support for expanded application of the Harris Co., Texas paradigm as a model for WNV evolution during the Dallas/Fort Worth, Texas outbreak, 2012. Boldface red, novel 2012 Collin and Denton Co. isolates sequenced in this chapter. Posterior probabilities ≥ 0.90 indicated. Scale bar, divergence time (in years).



CHAPTER 5

Natural Determinants of “Deep” WNV Population Structure

5.1 INTRODUCTION

In-depth, consensus-based phylogenetic studies of all published U.S. isolates have identified limited sequence differences ($< 1.4\%$ nt divergence)¹⁴ despite increasing regional heterogeneity and the accumulation of novel genotypic signatures in recent isolates—as confirmed in **Specific Aim 1** of this dissertation^{63,350,351}. In short, WNV continues to evolve in the U.S. Prior studies have published on the impact of both the principal *Culex* spp. mosquito vector and diverse range of avian hosts (corvids in particular) in the selection, adaptation, and expansion of WNV populations. Furthermore, parallel, haplotype reconstruction of artificial, sub-clonal WNV populations *in vitro* and *in vivo* has provided an overall perspective of consensus factors in the genetic flux and overall fitness of dominant WNV quasispecies structure^{116,120-122,302}.

Next generation sequencing (NGS) technologies provide an excellent tool to investigate the impact of specific genetic determinants on the structure, maintenance, and plasticity of natural WNV population dynamics below the level of consensus sequencing. In particular, integration of downstream single nucleotide variant (SNV) detection algorithms facilitates the identification of “rare” alternate nucleotides: A, C, G, or U—termed “deep” sequence variants—present in $< 50\%$ of the total RNA population per genomic position with limits-of-detection $\leq 1\%$ in multiple, robust platforms^{327,328}. Based on these principles, several NGS models have been developed from and applied to significant, viral human pathogens: HIV, hepatitis B virus (HBV), hepatitis C virus (HCV), and poliomyelitis virus (among others)^{330,352-355}; however, the bulk of these studies have remained restricted to viruses that are transmitted human-to-human. To date,

such studies have not been undertaken on arthropod-borne viruses that have complex transmission cycles involving multiple insect and vertebrate enzootic hosts which undergo compounding adaptive and selective pressures.

In this chapter, integration of the Illumina HiSeq1000 NGS pipe-line with the Harris Co. paradigm (defined in **Chapter 4**) provided an established, well-characterized pool of natural bird ($n = 14$) and mosquito pool ($n = 6$) isolates to define both host- and genotype-dependent genetic determinants involved in both RNA population structure and SNV selection as potential drivers in the evolution of natural WNV isolates. Prior applications of the Harris Co. paradigm confirmed the emergence, extinction, and sustained circulation of multiple U.S. genotypes, including the southeast coastal Texas, NA/WN02, and SW/WN03 genotypes as a surrogate model for WNV evolution on a national scale^{55,57,58,63,351}; inclusion of representative isolates from all 10 published, regional Harris Co. phylogenetic groups—described in **Chapter 4, Table 4-3** and **Fig. 4-1**—provides a robust platform to identify potential determinants driving the continued evolution of WNV in the U.S.^{63,351}. Paired application of a modified statistical test, Shannon entropy index,—developed to elucidate antiviral treatment-induced heterogeneity in HBV populations—provided a population-based metric to enumerate the mutational flux (*i.e.*, genetic diversity) in natural WNV quasispecies^{330,356}. Based on these principles, paired isolate information (including isolation host, phylogenetic relationships, or collection dates) for established Harris Co. strains and dual SNV detection supported the in-depth elucidation of critical host- and genotype-dependent determinants in the evolution of natural U.S. WNV populations.

5.2 RESULTS

Incorporation of the Illumina HiSeq1000 platform in the described NGS protocol (**Fig. 2-1**) provided a robust pipe-line to probe underlying host, spatial, and genotypic

factors involved in WNV population structure as an applied, natural model for arboviral evolution and adaptation in the U.S. Inclusion of 19 published or novel 2002-2012 Harris Co. isolates and one surrounding Montgomery Co., Texas isolate from distinct genotypic backgrounds (based on **Fig. 4-1**) provided a link with established phylogenetic relationships (**Table 5-1**)^{63,351}. In brief, vRNAs were extracted from the initial isolation or working stock (as indicated in **Table 5-1**) using the QIAmp Viral RNA Mini kit (Qiagen) and converted into cDNA via random hexameric primer amplification with the Illumina TruSeq RNA v2 kit (Illumina). Resulting 50 base-pair (bp) pair-end reads generated on the Illumina HiSeq1000 platform were processed to remove adaptor sequences and low-quality ($Q < 35$) or ambiguous (N) bases. Raw FastQ reads were then aligned to either the prototype NY99-flamingo382-99 (NY99; AF196835) or naturally attenuated TX1153 (AY712945) reference sequences with subsequent removal of PCR duplicates and down-sampling as indicated^{43,57}. Following NGS downstream processing, median raw depth-of-coverage ranged from 3,100- to 50,900-fold per genomic position for all natural isolates relative to NY99 (**Table 5-2**). Overall, contiguous coverage was confirmed in $\geq 99.4\%$ of all applied isolates extending from nucleotide position 13 to $\geq 10,976$ of the WNV genome.

5.2.1 Bridging traditional Sanger and NGS methodologies

5.2.1.1 CONSENSUS SEQUENCE COMPARISONS

Initial comparisons between NGS-indexed and published Sanger-based consensus sequences provided a means to link identified differences in vRNA population structure with published phylogenetic relationships. Extraction of NGS-derived FASTA sequences from processed Binary Alignment/Map (BAM) files in SAMtools v0.1.19 for all 21 applied isolates (*i.e.*, 20 Texas isolates and NY99 as reference) confirmed limited divergence from published sequences deposited in the GenBank.gov database. Overall, a

total of 16 nucleotide discrepancies were identified (encoding seven aa substitutions) in seven of these 21 NGS-indexed sequences including the prototype NY99 strain (**Table 5-3**). In particular, the extracted FASTA sequence for NY99 diverges from the published consensus at eight nucleotide positions (**Table 5-4**). Identification of five of these eight discrepancies in the NGS-indexed consensus alone (positions 1106 [aa M-N47I] and 7170) or derived infectious clone (NY99ic) backbone—indicated transitions at positions 1428, 3880 (aa NS2A-H119Y), and 4922 (aa NS3-K104R)—support the origin of these nt changes and aa substitutions through *in vitro* cell culture adaptation¹¹²; in contrast, conservation of three synonymous nt transitions at positions 7015, 8811, and 10851 in >98.8% of all published North American isolates indicated potential sequencing errors in the original prototype NY99-flamingo382-99 sequence⁴³.

5.2.1.2 PHYLOGENETIC COMPARISONS

Bayesian coalescent modeling of all NGS-indexed consensus sequences examined the impact of these detected discrepancies on published phylogenetic relationships in the Harris Co. paradigm; in brief, FASTA sequences were extracted from the BAM files of 21 applied NGS-indexed isolates including four additional isolates: IbAn-7019 (Nigeria 1980), AnD-27875 (Senegal 1993), Kuritz (Beaumont, TX 2002; SE coastal Texas genotype), and TM171-03 (Tabasco, Mexico 2003). Indexed consensus sequences were edited in BioEdit v7.0.9.0 to include the encoded ORF and aligned via the MUSCLE algorithm on the EMBL-EBI server³⁰⁷⁻³⁰⁹. Inferred topologies applied the GTR+I+ Γ_4 substitution model and log-normal relaxed clock model in BEAST v1.6.2 (<http://beast.bio.edi.ac.uk>) with down-sampling of triplicate 50 million-state runs in LogCombiner v1.7.4 and TreeAnnotator v1.7.4^{317,318}. Both non-U.S. strains: IbAn-7019 and AnD-27875 served as common “Old World” lineage 1a phylogenetic outgroups;

furthermore, both the Kuritz and TM171-03 strains served as members of the SE coastal Texas and NY99 genotypes, respectively.

No topological differences were identified in the applied Bayesian phylogenies with the retention of published genotypic trends for all isolates in the NY99, NA/WN02, and SW/WN03 U.S. genotypes (**Fig. 5-1**)—as described in **Chapter 4, Fig. 4-1**^{63,351}. In addition, both newly reported 2010 TX8102 and 2011 TX8375 isolates clustered with the published 2010 TX8092 isolate in Group 7 of the NA/WN02 genotype with <0.19% nt divergence and 3-7 conserved substitutions (**Table 5-5**). Based on these results, application of the described NGS pipe-line to the Harris Co. paradigm reproduced phylogenetic relationships consistent with published genotypic trends without loss of consensus information obtained in traditional Sanger sequencing technologies; furthermore, these results supported the integration of published genotypic relationships as robust clustering metrics in downstream genetic diversity and SNV detection analyses.

5.2.2 “Deep” NGS populations retain consensus-based trends

In-depth detection of positional or genome-wide variation in applied, total vRNA populations provided a semi-quantitative criterion to assess the impact (direct or indirect) of defined environmental factors or stimuli in WNV quasispecies structure (*i.e.*, genetic diversity). Modified application of Shannon entropy to haploid, WNV-derived NGS datasets quantified the relative plasticity of a specific nucleotide position or encoded genomic region in the total RNA population structure^{330,331}. In this described NGS pipe-line, aligned BAM files were indexed using the bam2R function of the deepSNV v1.8.0 software package³²⁹ in R relative to the prototype NY99 strain. Tabulated reference or alternate allelic counts: A, C, G, U, or gap per genomic position were then applied to a modified Shannon entropy index defined by Nishijima *et al.* (2012)³³⁰. Resulting entropic indices ranged from 0.00 (minimum diversity; fixed nucleotide) to 1.61 (maximum

diversity; each nucleotide detected at 25%) for a specific nucleotide position or as an average across a delineated genomic region—as described in **Section 2.7.5**.

Prior to in-depth comparisons, initial application of this modified entropic index in the defined NGS pipe-line re-evaluated the published phylogenetic relationships in the Harris Co. paradigm using the total RNA population structure and confirmed parameters: (1) genotype ($\alpha = 0.017$), (2) regional groupings ($\alpha = 0.0011$), and (3) year-of-isolation ($\alpha = 0.0011$) with determined statistical significance via Bonferroni-corrected Mann Whitney U-test (**Table 5-1** and **Fig. 5-1**). Based on this model, isolates that clustered in the SW/WN03 genotype (Group 5) exhibited higher mean Shannon entropies (p -value < 0.05) for all regions of the WNV genome, excluding the prM gene ($p = 0.055$) and 5'UTR ($p = 0.236$). Furthermore, all regional groups (in general) in the NA/WN02 genotype demonstrated distinct entropic signatures ($p < 0.05$); however, the RNA population structure of Group 1, Group 2, and Group 5 isolates were indistinguishable for all genomic regions except the NS5 gene and 5'UTR ($p < 0.05$). In addition, Group 3 and Group 7 isolates exhibited similar genome-wide entropic trends, excluding the NS2A ($p < 0.01$) and NS5 genes ($p < 0.05$).

From a temporal perspective, no correlation existed between confirmed *low* or *high* annual diversity trends relative to reported 2002-2012 WNV incidence in Harris Co., TX mosquito or human populations (**Fig. 5-2**)⁴⁷; however, 2003 (Groups 1 and 3), 2010 (Group 7), and 2012 isolates all exhibited similar mean Shannon entropies for the entire genome—excluding the NS3 and NS4B genes (p -value < 0.05). In effect, isolates sequenced from these years exhibited comparable diversity profiles consistent with published, consensus sequence-based phylogenetic trends—described in **Chapter 4**³⁵¹. Based on these results, genotypic and phylogenetic trends derived from consensus-based methodologies are retained in the total WNV population structure.

5.2.3 Host-dependent factors in natural WNV diversity

Previous published *in vitro* and *in vivo* studies with laboratory-derived, sub-clonal WNV populations to elucidate host-fitness trade-offs in WNV isolates modified the original Trade-Off Hypothesis with the *Culex* spp. mosquito vector serving as the principal amplifying host and passage in birds providing strong, purifying selection on WNV populations—despite no apparent fitness costs^{116,120-122}. Based on these studies, indexed Shannon entropies were pooled for all applied bird and mosquito-derived WNV isolates (**Fig. 5-3**) in order to confirm natural, host-specific bias in positional or regional genetic diversity. Overall, total RNA populations sampled from natural bird isolates demonstrated higher mean Shannon entropic trends for all genomic regions (except in the prM gene) with significant increases in region-specific diversity in the NS2A ($p = 0.034$) and NS4B ($p = 0.021$) genes and the 3'UTR ($p = 0.028$). Pooled mosquito isolates exhibited position-dependent increases in mean Shannon entropies (>0.050 ; equivalent to $>0.9\%$ total mutational frequency of a single, minor allele at a particular genomic position) restricted to 10 nucleotide positions: 211, 728, 2390, 2993, 7096, 7431, 7614, 7749, 7860, and 10888 in the encoded C, prM, E, NS1, NS4B, and NS5 genes and the 3'UTR. In comparison, positions 93, 9420, and 10888 alone in the 5'UTR, NS5 gene, and 3'UTR, respectively, of bird isolates exhibited mean Shannon entropy values > 0.05 (**Fig. 5-3**). Consolidation of these results indicated host-specific bias in WNV population structure with higher mean genetic diversity in bird isolates and position-dependent selection in mosquito pool isolates.

5.2.4 Determinants driving WNV population diversity: SNV detection

Integration of both the SNVer v0.5.2 (binomial-binomial)³²⁷ and VarScan2 v2.3.6 (heuristic; strict 0.5% SNV frequency cut-off)³²⁸ SNV detection models provided a robust method to detect and confirm significant— $\alpha = 4.53 \times 10^{-6}$ and $p\text{-value} < 0.05$, respectively—SNVs representing down to 0.86% of the total WNV population structure below the level of consensus; furthermore, all reported SNVs were confirmed with both

detection platforms to limit the identification of false-positive variants. Overall, 294 SNVs were detected in all 21 applied, natural WNV isolates (relative to the prototype NY99 sequence); 159 significant ($p < 0.05$) SNVs were identified representing 140 unique positions in the WNV genome (**Tables 5-2 and 5-6**). Mutational frequencies—the percentage of alternate nucleotides detected—ranged from 0.86-48.63% of the total depth-of-coverage for all 159 SNVs identified at these positions. In particular, 47 of the detected SNVs (29.6% or 36.2%, excluding SNVs in the 5'/3'UTRs, respectively) encoded non-synonymous amino acid substitutions in up to 48.63% of the total vRNA population (**Table 5-7**). Mean SNV detection rates varied from gene-to-gene with an overall genome-wide rate of 6.87×10^{-4} mutations/site/genome or 8 SNVs per WNV isolate. Seven positions: 93, 732, 3409, 6871, 10408, 10435, and 10888 in the WNV genome encoded 26 SNVs conserved in >1 isolate at 1.00 to 35.71% mutational frequencies (**Fig. 5-4A**); both alternate nucleotides at positions 3409 and 6871 represent conserved, non-synonymous mutations in the encoded NS1 (R314G) and NS4A (V135M) WNV proteins.

No biases in significant SNV detection were observed relative to host (mosquito or bird, $p > 0.205$), year-of-isolation ($p > 0.397$), or phylogenetic relationship ($p > 0.511$); furthermore, there was no statistical deviation in the number of significant SNVs detected between encoded genomic regions. Despite no host-dependent bias, no single SNV was identified in >1 mosquito isolate; in contrast, all seven conserved SNVs were detected in multiple bird isolates (**Fig. 5-4B**). Consistent with both the region- and position-specific elevations in mean Shannon entropies for bird isolates, all SNVs identified in >2 isolates were present in either the 5' or 3'UTRs, including peak diversities at positions 93 and 10888 (**Figs. 5-3 and 5-4B**). In particular, the C to U transition at position 93 in the 5'UTR was conserved in 7 out of 15 (46.7%) bird isolates at 2.51 to 35.71% mutational frequencies. Exceptions included positions 10408 and 10888 in the 3'UTR; in particular, the C to U transition at position 10408 was confirmed in the consensus sequence for

seven bird ($n = 5$) and mosquito ($n = 2$) isolates, and a single mosquito pool isolate (M19433) encoded the U to A transversion at position 10888 shared with additional bird isolates. Based on these results, SNV detection in natural mosquito isolates supports random positive selection of alternate nucleotide frequencies restricted to individual pooled samples; in contrast, bird isolates exhibit both region- and position-specific trends in SNV selection.

5.2.5 Phenotypic link to WNV population diversity

Previously published studies utilizing natural isolates in the Harris Co. paradigm identified variable mouse neuroinvasive phenotype in outbred 3-5 week old NIH Swiss or Swiss Webster mice^{29,57,174}. In order to confirm the *in vivo* phenotype for all applied isolates, all isolates were screened in female 3-4 week old, outbred Swiss Webster mice via intraperitoneal (i.p.) inoculation with either 5000 or 10 pfu virus doses. Isolates resulting in $\leq 50\%$ survival at the 10 pfu dose were designated “virulent” and all isolates with $\geq 50\%$ survival at the 5000 pfu dose were designated “attenuated”. Follow-up i.p. 50% lethal median dose (ipLD₅₀) studies at dosages of 10^3 to 10^{-1} pfu of virus ($n = 5$ mice per virus dose) confirmed all identified “attenuated” *in vivo* phenotypes relative to NY99 (virulent), TX1171 (attenuated), or PBS (mock) inoculation. Based on these data, isolates were grouped according to confirmed *in vivo* phenotype: virulent or attenuated and/or isolation host-type; furthermore, mean Shannon entropies were calculated per isolate with designations of *low* (blue; mean Shannon entropies ≤ 0.0025) to *high* (red; ≥ 0.0035) genetic diversity for each specific genomic region (**Fig. 5-5**).

No consistent entropic trends were identified among confirmed virulent mosquito pool isolates; in contrast, all attenuated isolates—both bird- and mosquito pool-derived—exhibited *low* genetic diversity trends for all identified genomic regions, excluding the 3’UTR, compared to *high* genetic diversity trends in the C gene and both 5’/3’UTRs of virulent bird isolates (**Figs. 5-3 and 5-5**). Bonferroni-corrected Mann-Whitney U tests

confirmed a significant decrease in the mean Shannon entropies of attenuated isolates for all genomic regions excluding the 3'UTR (p -value = 0.708) relative to all virulent isolates (Fig. 5-6). In addition, the absence of conserved SNV populations between natural mosquito pool isolates (both attenuated and virulent) is consistent with these trends; furthermore, 6 of the 7 conserved SNVs (excluding the SNV at position 10888) remained restricted to virulent bird isolates alone with 58.3% of these isolates encoding the C to U transition at position 93 in the 5'UTR in particular. In conclusion, the overall relationship between WNV population structure and *in vivo* neuroinvasive phenotype represents a continuum; however, consistent entropic and SNV detection trends support a link between *in vivo* attenuation and homogeneous WNV populations defined through limited genetic variation.

5.3 DISCUSSION

Implementation of initial consensus-based comparisons in the defined NGS pipeline provided a benchmark to confirm published trends in the Harris Co. paradigm and bridge novel dynamics in WNV population structure with published, consensus-based phylogenetic relationships. Limited differences between NGS-derived and published consensus sequences for all applied isolates did not alter published phylogenetic relationships^{63,351}; however, detected discrepancies in the reference NY99-flamingo382-99 sequence at positions 7015, 8811, and 10851 supported potential sequencing errors in this prototype U.S. strain consistent with >98.8% of all published U.S. isolates⁴³. Based on these results, observed host-, genotype-, and phenotype-dependent trends can be attributed to intrinsic differences in the underlying population structure of WNV isolates in the Harris Co. paradigm.

Published efforts to link the evolution and fitness of WNV populations to natural, molecular determinants or host-dependent stimuli remain limited. Prior studies, involving

infectious clone-derived viruses, incorporated sub-clonal or haplotype reconstruction in several alphaviral and flaviviral systems (including chikungunya virus (CHIKV), dengue fever virus [DENV], St. Louis encephalitis virus [SLEV], Venezuelan equine encephalitis virus [VEEV], and WNV) following *in vitro* or *in vivo* population bottlenecks to define host- or passage-induced dependencies in the fitness and genetic plasticity of arbovirus populations^{118,119,302,303,357-360}. Based on these studies (and others), the Trade-Off Hypothesis serves as the current paradigmatic (yet controversial) model for WNV evolution and fitness in the U.S.; in brief, stochastic expansion of WNV diversity in the *Culex* spp. mosquito vector is restricted through strong, purifying selection in avian hosts resulting in limited WNV host-adaptation and evolution—despite no apparent fitness costs^{116,119,120,122}. However, utilization of infectious clone-derived viruses raises the question^{116,361,362}: do consensus-based cross-sections reflect the true depth of WNV population structure? Integration of a Shannon entropic index in the Illumina HiSeq1000 NGS pipe-line allowed detection of all allelic constituents: A, C, G, U, and “gap” in applied RNA samples providing a comprehensive, population-based metric to enumerate the intrinsic diversity in WNV populations in the Harris Co. model; however, due to lower depth-of-coverage in the encoded *N*- and *C*-termini—inherent in NGS processing of viral RNA—this index is biased towards genetic diversity in either the positive- or minus-strand in the respective 5’UTR (nts 13-70) and 3’UTR (nts 10900-11007) terminal genomic regions.

Based on this approach, circulating U.S. genotypes and regional phylogenetic groups in Harris Co. (Houston), TX exhibited distinct entropic signatures consistent with published consensus-based phylogenetic trends described in **Chapters 3-4**^{63,351}. Retention of these *in silico* relationships with additional U.S. isolates collected in other geographic regions supports the conservation of these identified WNV diversity trends within endemic bird or *Culex* spp. mosquito populations outside the Harris Co. paradigm^{68,69,347,348}; however, confirmation is needed. No correlation was detected

between reported WNV surveillance efforts—in mosquitoes (**Fig. 5-3**)—and the observed diversity of paired, sampled WNV isolates; in effect, the epidemiological role of WNV incidence in local *Culex* spp. mosquito pools as an annual indicator of WNV diversity is limited.

Detection of “deep” SNVs in this chapter provided an additional, paired method to confirm entropic trends in WNV quasispecies structure. In total, 159 significant ($p < 0.05$) SNVs were identified at 140 unique genomic positions among all 21 sampled isolates resulting in a mean SNV detection rate of 7-8 SNVs/genome (or 6 to 7×10^{-4} mutations/site) at $\geq 0.86\%$ positional frequencies (**Table 5-6**); this rate is consistent with reported, annual U.S. WNV mutation rates (10^{-4} to 10^{-3} substitutions/site) supporting the comparable, cumulative impact of genetic drift, error-prone flaviviral replication, and host-specific pressures on SNV selection^{13,62,64,363}. However, direct evidence for the temporal selection or fixation of SNV populations is limited. In general, detected SNVs in the encoded polyprotein sequence represented novel WNV mutations or incomplete fixation of the NGS-derived consensus sequence; thus, the potential structural or phenotypic roles of individual SNVs are undefined. Exceptions included multiple SNVs detected in the NS4A gene and both the 5’/3’UTRs: C93U, G6871A (NS4A-V135M), C10408U, U10414C, and C10435U identified in published North American WNV isolates collected from humans, birds, and mosquito pools^{58,69,351}. Furthermore, both conserved C93U and U10888A SNVs reside in critical 5’/3’UTR RNA elements involved in either 5’ Type I cap-dependent translation and/or formation of the 5’-3’ panhandle structure required for initiation of flavivirus minus-strand RNA synthesis^{202,206,364,365}. Further research is needed to elucidate the role of individual or synergistic SNV populations as potential, underlying components in host-specific WNV adaptation or flavivirus replication using combined infectious clone- and haplotype-based *in vitro* approaches.

Consolidation of these entropic trends and detected SNV populations among applied bird- and mosquito pool-derived isolates supports a novel perspective to the Trade-Off Hypothesis highlighting host-dependent biases in population structure, at least for isolates from Harris Co., TX. For all regions of the WNV genome (except the prM gene), natural bird isolates exhibited higher genetic diversity with significant region-specific biases linked to the NS2A and NS4B genes and 3'UTR. These results do not, in fact, contradict the Trade-Off Hypothesis. Despite higher overall entropic trends, bird-derived isolates exhibited conserved dynamics in WNV quasispecies structure with evidence of regional, host-dependent SNV selection in the prM, NS1, and NS4A genes and both 5'/3'UTRs; in contrast, unique diversity trends and detected SNV populations were confirmed in all applied mosquito pool isolates. For example, all positional peaks in mean mosquito pool Shannon entropies (excluding positions 7614 and 10888; **Fig. 5-3**) represented encoded SNVs (detected at 8.86-13.04% mutational frequencies) restricted to a single isolate (M6019). Despite these trends, implementation of “pooled” mosquito isolates in the applied NGS platform incorporates inherent bias in these diversity comparisons with potential mixing of multiple, distinct isolates in pooled isolation techniques. However, no evidence of mixed mosquito-derived virus populations was detected as defined by confirmation of a single, dominant consensus sequence for all applied mosquito pool isolates. Overall, these trends support the independent, stochastic expansion of total WNV populations within the *Culex* spp. mosquito vector with subsequent selection of conserved WNV population structure in the avian host consistent with the Trade-Off Hypothesis^{116,119,120,122}.

Despite these host-dependent trends, both *Culex* spp. and bird isolates exhibited a continuum in overall or region-specific genetic diversity; in effect, determination of the *in vivo* neuroinvasive phenotype for all applied Harris Co. isolates identified a novel link between WNV neurovirulence and population structure. Natural WNV isolates with *attenuated* phenotype demonstrated *low* overall genetic diversity. This is not an original

concept. Pioneering studies confirmed the de-limiting impact of quasispecies structure on *in vitro* RNA virus pathogenesis^{114,366,367}; furthermore, *in vivo* attenuation and reduced fitness of both engineered high-fidelity and/or ribavirin-resistant virus populations has been linked to homogenous quasispecies structure with compromised adaptation to host selection pressures^{304,368-370}. However, this is the first report supporting *low* genetic diversity in natural virus populations as a positive correlate of *in vivo* attenuation. Exceptions exist, as in all biological models, but limited WNV diversity in this model has been mapped to several genomic regions including the 5'UTR of non-neuroinvasive bird isolates consistent with linked entropic trends and SNV populations. In turn, all detected SNVs in either the 5' or 3'UTRs (excluding the prolific U10888A mutation) remained restricted to virulent bird isolates. Besides potential implications in modulated WNV replication and/or translation dynamics, explicit phenotypic advantages for SNV selection in UTR elements remains unclear; however, the conserved C93U SNV in the 5'-upstream initiation AUG region (AUR) of stem-loop II (SLII) resides in the putative OAS1 binding domain and may modulate host RNaseL degradation of WNV vRNAs³⁷¹. In conclusion, further investigation is needed to both elucidate the impact of intra-host diversity and selection in natural WNV populations and explore the novel link between *in vivo* virulence and the diversity of natural virus populations.

Table 5-1: WNV isolates in the described Next Generation Sequence (NGS) analyses.

Strain	Location	Source*	Passage History†	Collection Year	Phylogenetic Group‡	GenBank Accession No.
NY99-flamingo382-99	New York, NY	Chilean Flamingo	V (p2)	1999	-	AF196835
TX114	Harris Co., TX	Blue Jay	V	2002	1	GU827998
TX1153	Harris Co., TX	Mourning Dove	V	2003	1	AY712945
TX1171	Harris Co., TX	Blue Jay	V	2003	1	AY712946
TX1175	Harris Co., TX	Blue Jay	V	2003	3	GU828000
TX1576	Montgomery Co., TX	Blue Jay	V	2003	3	GU827999
v4095	Harris Co., TX	<i>Cx. quinquefasciatus</i>	V-E6	2003	2	GU828002
v4369	Harris Co., TX	<i>Cx. quinquefasciatus</i>	V-E6	2003	2	AY712948
M6019	Harris Co., TX	<i>Cx. quinquefasciatus</i>	C V	2006	6	JF415930
M19433	Harris Co., TX	<i>Cx. quinquefasciatus</i>	V	2007	5	JF415919
TX7558	Harris Co., TX	Blue Jay	V	2008	5	JF415921
M20122	Harris Co., TX	<i>Ae. albopictus</i>	V	2009	4	JF415928
M38488	Harris Co., TX	<i>Ae. albopictus</i>	V	2009	4	JF415925
TX8092	Harris Co., TX	House Sparrow	V	2010	7	KC333374
TX8102	Harris Co., TX	Blue Jay	V	2010	-	-
TX8349	Harris Co., TX	House Sparrow	V	2011	5	KC333375
TX8375	Harris Co., TX	American Crow	V	2011	-	-
TX8551	Harris Co., TX	Blue Jay	V	2012	10	KC333377
TX8560	Harris Co., TX	Blue Jay	V	2012	8	KC333379
TX8567	Harris Co., TX	Blue Jay	V	2012	9	KC333381
TX8572	Harris Co., TX	Blue Jay	V	2012	-	KC333383

*Mosquito isolates were collected from either *Culex* (*Cx.*) *quinquefasciatus* or *Aedes* (*Ae.*) *albopictus* mosquito pools.

†Isolates were passaged in Vero (V), Vero E6 (V-E6), or C6/36 (C) cell-lines in the order indicated with successive passages in the same cell-line designated.

‡Phylogenetic group designations based on published Harris County, Texas phylogenetic analyses^{63,351}.

Table 5-2: Depth-of-coverage and detected single nucleotide variants (SNVs)

Strain	Median depth-of-coverage (x1000)*	Genomic region†		Detected SNVs (#)		
		Start (5')	Stop (3')	Total	p-Value Filter‡	≥1% Frequency¶
NY99-flamingo382-99	40.9	13	10991	5	4	4
TX114	14.5	13	10978	15	5	4
TX1153	25.1	13	10993	8	4	4
TX1171	7.0	13	10993	9	8	8
TX1175	21.4	13	10993	2	2	2
TX1576	23.7	13	10991	14	5	5
v4095	15.3	13	10990	3	0	0
v4369	32.3	13	10994	12	5	4
M6019	19.7	13	10991	16	11	11
M19433	41.9	13	11007	4	2	2
TX7558	10.2	13	10987	24	13	12
M20122	10.6	13	10984	2	2	2
M38488	3.1	13	10976	17	9	9
TX8092	10.4	13	10991	3	3	3
TX8102	20.8	13	10993	36	14	13
TX8349	11.4	13	10993	35	15	12
TX8375	32.0	13	10991	22	18	18
TX8551	50.9	13	11007	16	9	9
TX8560	44.7	13	10993	40	23	21
TX8567	28.8	13	10993	9	5	5
TX8572	33.4	13	10991	2	2	2
Total				294	159	150

*Depth-of-coverage, raw median number of unique reads (A, C, G, or U) at each position in the WNV genome.

†Region of the output NGS sequence with a Phred (Q)-like base quality threshold >1000 after removal of PCR duplicates and down-sampling to 3000x depth-of-coverage.

‡Number of detected alternate minor alleles at Bonferroni-adjusted p-value ≤ 0.05 with both the SNVer and VarScan2 SNV detection methods (including all SNVer detected SNVs at position 10888).

¶Detected SNVs which represent ≥1% of all nucleotide reads at a specific position.

Table 5-3: Consensus differences between the NGS-indexed and published NY99-flamingo382-99 (NY99) sequences

		M		E		NS2A			NS3			NS4B			NS5		3'	
		1106‡	1428	1442‡	2466	3880‡	4146	4803	4922‡	6138	6426	6996	7015	7170	7938	8811	9352	10851
Strain†	Year																	
NY99 (382-99)	1999	A	C	U	C	C	A	C	A	C	C	C	U	C	U	U	C	A
382-99 (Vero p2)	1999	U	U	.	.	U	.	.	G	.	.	.	C	U	.	C	.	G
NY99ic (Vero p1)	-	.	U	.	.	U	.	.	G	C	.	G
TX114	2002	.	.	C	U	.	G	U	.	U	U	U	C	.	C	C	U	G
TX1153	2003	.	.	C	U	.	G	U	.	U	U	U	C	.	C	C	U	G
TX8092	2010	.	.	C	U	.	G	U	.	U	U	U	C	.	C	C	U	G
TX8349	2011	.	.	C	U	.	G	U	.	U	U	U	C	.	C	C	U	G
TX8551	2012	.	.	C	U	.	G	U	.	U	U	U	C	.	C	C	U	G

*M, membrane; E, envelope; NS, nonstructural; 3', 3' untranslated region (UTR). Nucleotide positions are indicated which differ between the indexed NGS and published NY99-flamingo382-99 (NY99; AF196836) consensus sequences relative to other published 2002-2012 US strains.

†NY99 infectious clone (NY99ic) is derived from the prototype NY99-flamingo382-99 strain^{43,112}.

‡Four nucleotide changes encode amino acid substitutions: 1106 (M-N47I), 1442 (E-V159A), 3880 (NS2A-H119Y), and 4922 (NS3-K104R).

Table 5-4: NGS and published consensus sequence discrepancies

Strain	Year	Nucleotide		Location†		
		Reference	Alternate	Position	UTR/Gene	Substitution
NY99-flamingo382-99	1999	A	U	1106	E	N47I
	.	C	U	1428	E	-
	.	C	U	3880	NS2A	H119Y
	.	A	G	4922	NS3	K104R
	.	U	C	7015	NS4B	-
	.	C	U	7170	NS4B	-
	.	U	C	8811	NS5	-
	.	A	G	10851	3'UTR	-
TX1171	2003	U	G	8279	NS5	L200R
	.	A	C	9743	NS5	D688A
	.	U	G	11000	3'UTR	-
TX1175	2003	U	G	11015	3'UTR	-
TX1576	2003	U	G	11015	3'UTR	-
M6019	2006	A	U	2993	NS1	N175I
TX8551	2012	C	U	7215	NS4B	A115V
TX8567	2012	U	C	7515	NS4B	-

*E, envelope; NS, nonstructural; UTR, untranslated region. Both the consensus reference and NGS-identified alternate nucleotides are indicated for each indicated West Nile virus strain relative to its position in the genomic sequence.

†Differences which encode amino acid substitutions are highlighted in red boldface.

Table 5-5: Encoded amino acid substitutions in the novel 2010 and 2011 Harris Co., TX WNV isolates

Strain†	Year	C	M	E	NS1	NS2A	NS2B	NS3		NS4B		NS5	
		76	150	159	167	90	103	177	235	23	202	314	653
NY99wt (382-99)	1999	T	M	V	M	M	V	A	A	V	I	K	F
TX8092	2010	I	T	A	I	V	F	.	.	A	V	.	.
TX8102	2010	.	T	A	I	V	F	.	.	A	V	.	.
TX8375	2011	.	T	A	I	.	.	V	V	.	.	Q	S

*C, capsid; M, membrane; E, envelope; NS, nonstructural; 2A, NS2A; 2B, NS2B. Amino acid changes indicated relative to the prototype NY99-flamingo382-99 strain AF196836.
†Isolates sequenced in this study indicated in boldface.

Table 5-6: Single nucleotide variants (SNVs) detection in natural WNV populations, p -value < 0.05 (Bonferroni-corrected)

Strain	Year	Position*	Gene	Residue	Mutation		SNV Count		Reference Count		Depth-of-Coverage	SNV Frequency (%)	p-value
					REF	SNV	+ Strand	- Strand	+ Strand	- Strand			
NY99-flamingo382-99 (382-99)	1999	1106	E	I47N	T	A	330	262	1162	825	2579	22.95%	0.00E+00
		2760	NS1	Y97	C	T	411	703	1038	1277	3429	32.49%	0.00E+00
		3880	NS2A	Y119H	T	C	298	156	1263	510	2227	20.39%	0.00E+00
		7170	NS4B	G85	T	C	257	201	1085	688	2231	20.53%	0.00E+00
TX 114	2002	93	5'UTR	-	C	T	67	20	425	134	646	13.47%	2.22E-16
		3384	NS1	G305	A	T	19	2	843	110	974	2.16%	1.07E-10
		4932	NS3	K107	G	A	28	4	2819	419	3270	0.98%	5.78E-07
		6871	NS4A	V135M	G	A	21	18	1137	1907	3083	1.27%	2.27E-11
		10408	3'UTR	-	C	T	15	18	1348	1631	3012	1.10%	2.43E-08
		10777	3'UTR	-	T	C	4	10	150	189	353	3.97%	1.22E-10
TX 1153	2003	3840	NS2A	I105	T	C	15	21	1465	1764	3265	1.10%	5.50E-09
		8424	NS5	T248	C	T	71	134	1704	1362	3271	6.27%	0.00E+00
		<i>9471</i>	<i>NS5</i>	<i>S597</i>	<i>T</i>	<i>C</i>	27	27	1381	1855	3290	1.64%	0.00E+00
		10888	3'UTR	-	T	A	0	30	349	2321	2700	1.11%	2.87E-07
TX 1171	2003	2964	NS1	T165	T	C	22	16	1520	1005	2563	1.48%	9.89E-13
		3898	NS2A	L125M	C	A	29	23	1648	657	2357	2.21%	0.00E+00
		4736	NS3	V42A	T	C	42	38	1108	897	2085	3.84%	3.33E-16
		7488	NS4B	S191	T	C	15	20	1289	1175	2499	1.40%	2.38E-11
		7796	NS5	A39V	C	T	96	38	1159	1444	2737	4.90%	6.66E-16
		7837	NS5	H53Y	C	T	94	122	1475	1344	3035	7.12%	1.11E-16
		8022	NS5	Q114	A	G	18	13	1153	1168	2352	1.32%	1.03E-09
		10888	3'UTR	-	T	A	0	29	340	1863	2232	1.30%	2.66E-08
TX 1175	2003	8454	NS5	V258	A	G	34	31	1544	1377	2986	2.18%	6.65E-30
		10888	3'UTR	-	T	A	0	36	332	1258	1626	2.21%	2.22E-16
TX 1576	2003	732	pr	R89	C	T	74	76	1649	2249	4048	3.71%	1.11E-16
		2495	NS1	N9S	A	G	914	540	1000	536	2990	48.63%	0.00E+00
		10393	NS5	L905	C	T	21	9	1295	1027	2352	1.28%	3.55E-09

*Boldface, SNVs which encode amino acid substitutions. Italics, both synonymous and non-synonymous SNVs which represent a reversion to the prototype NY99 sequence.

Table 5-6: Single nucleotide variants (SNVs) detection in natural WNV populations, p -value < 0.05 (Bonferroni-corrected)
(Continued)

Strain	Year	Position*	Gene	Residue	Mutation		SNV Count		Reference Count		Depth-of-Coverage	SNV Frequency (%)	p-value
					REF	SNV	+ Strand	- Strand	+ Strand	- Strand			
TX 1576		10408	3'UTR	-	C	T	57	48	1295	837	2237	4.69%	1.11E-16
		10888	3'UTR	-	T	A	0	22	242	1402	1666	1.32%	6.99E-07
v4095	2003	10449	3'UTR	-	C	T	2	10	489	1526	2027	0.59%	0.00E+00
v4369	2003	1194	E	T76	C	T	36	29	1215	1596	2876	2.26%	1.11E-16
		2043	E	N359	C	T	22	8	1783	791	2604	1.15%	3.13E-08
		2428	E	V488F	G	T	24	6	2172	898	3100	0.97%	1.34E-06
		3138	NS1	C222	T	C	13	28	1254	2708	4003	1.02%	7.55E-09
		4870	NS3	H87Y	C	T	20	22	1801	2193	4036	1.04%	3.61E-09
		10393	NS5	L905	T	C	5	1	888	943	1837	0.33%	0.00E+00
M6019	2006	211	C	N39D	A	G	113	48	1306	350	1817	8.86%	0.00E+00
		728	pr	K88R	A	G	176	147	1432	1849	3604	8.96%	0.00E+00
		2390	E	S475N	G	A	122	90	1186	1156	2554	8.30%	0.00E+00
		2605	NS1	I46V	A	G	94	66	2076	1774	4010	3.99%	0.00E+00
		2993	NS1	I175N	T	A	201	73	1437	673	2384	11.49%	0.00E+00
		5629	NS3	I340V	A	G	128	40	2031	748	2947	5.70%	2.22E-16
		7096	NS4B	T61A	A	G	188	74	1270	779	2311	11.34%	1.11E-16
		7431	NS4B	K172	A	G	227	53	1860	879	3019	9.27%	7.84E-289
		7749	NS5	E23	A	G	208	233	1414	1527	3382	13.04%	0.00E+00
		7860	NS5	A60	A	G	127	145	1768	935	2975	9.14%	3.66E-278
		10612	3'UTR	-	C	A	14	34	1292	784	2124	2.26%	2.70E-23
M19433	2007	9943	NS5	D755N	G	A	38	24	1826	1100	2988	2.07%	1.11E-16
		10888	3'UTR	-	T	A	0	39	307	1842	2188	1.78%	1.33E-15
TX 7558	2008	93	5'UTR	-	C	T	175	6	1487	71	1739	10.41%	9.30E-197
		1257	E	V97	G	A	37	40	957	1451	2485	3.10%	1.95E-45
		1847	E	K294R	A	G	8	22	866	1587	2483	1.21%	9.53E-09
		2241	E	A425	T	C	208	251	1569	1248	3276	14.01%	0.00E+00
		2926	NS1	L153	T	C	16	19	1754	1228	3017	1.16%	2.56E-09

*Boldface, SNVs which encode amino acid substitutions. Italics, both synonymous and non-synonymous SNVs which represent a reversion to the prototype NY99 sequence.

Table 5-6: Single nucleotide variants (SNVs) detection in natural WNV populations, p -value < 0.05 (Bonferroni-corrected)
(Continued)

Strain	Year	Position*	Gene	Residue	Mutation		SNV Count		Reference Count		Depth-of-Coverage	SNV Frequency (%)	p-value
					REF	SNV	+ Strand	- Strand	+ Strand	- Strand			
TX 7558	2008	2940	NS1	D157	T	C	51	18	1535	701	2305	2.99%	5.55E-16
		3556	NS2A	L11	C	T	14	2	1492	99	1607	1.00%	0.00E+00
		3633	NS2A	A36	C	T	6	13	1571	874	2464	0.77%	0.00E+00
		4332	NS2B	L38	C	T	113	36	1428	419	1996	7.46%	1.11E-16
		4871	NS3	H87R	A	G	333	523	1523	2702	5081	16.85%	0.00E+00
		5238	NS3	I209	C	T	35	52	1868	3714	5669	1.53%	6.66E-16
		5553	NS3	F314	C	T	14	22	823	1800	2659	1.35%	3.80E-11
		6780	NS4A	I104	C	T	32	6	2223	1189	3450	1.10%	5.96E-09
		7266	NS4B	L117	T	C	12	3	1739	188	1942	0.77%	0.00E+00
		7793	NS5	S38L	C	T	25	26	1999	1823	3873	1.32%	1.07E-14
M20122	2009	8802	NS5	Y374	C	T	23	13	2285	1614	3935	0.91%	5.25E-07
		3891	NS2A	R122	C	T	27	18	1946	283	2274	1.98%	1.11E-16
		8453	NS5	V258A	T	C	16	15	1682	1373	3086	1.00%	3.60E-06
		61	5'UTR	-	G	A	80	0	1492	1	1573	5.09%	0.00E+00
		2565	NS1	Y32	T	C	32	42	517	1012	1603	4.62%	0.00E+00
		2648	NS1	V60A	T	C	22	10	1375	1036	2443	1.31%	1.03E-09
		2784	NS1	T105	T	C	351	254	637	524	1766	34.26%	0.00E+00
		4361	NS2B	K48R	A	G	10	20	791	1495	2316	1.30%	2.22E-09
		4928	NS3	A106V	C	T	32	17	1015	424	1488	3.29%	0.00E+00
		5250	NS3	M213I	G	A	32	36	785	1090	1943	3.50%	0.00E+00
M38488	2009	5316	NS3	A235	A	G	9	18	573	1062	1662	1.62%	3.32E-09
		7146	NS4B	S77	C	T	36	25	712	447	1220	5.00%	0.00E+00
		102	C	S2	T	C	41	5	959	70	1075	4.28%	0.00E+00
		3931	NS2A	L136	C	T	293	212	1462	1144	3111	16.23%	0.00E+00
		3969	NS2A	F148	C	T	235	214	1613	1203	3265	13.75%	0.00E+00
		93	5'UTR	-	C	T	425	26	760	52	1263	35.71%	0.00E+00
		1799	E	T278I	C	T	12	31	1238	1415	2696	1.59%	1.22E-15
		2604	NS1	I45	C	T	32	36	2132	1910	4110	1.65%	4.28E-24

*Boldface, SNVs which encode amino acid substitutions. Italics, both synonymous and non-synonymous SNVs which represent a reversion to the prototype NY99 sequence.

Table 5-6: Single nucleotide variants (SNVs) detection in natural WNV populations, p -value < 0.05 (Bonferroni-corrected) (Continued)

Strain	Year	Position	Gene	Residue	Mutation		SNV Count		Reference Count		Depth-of-Coverage	SNV Frequency (%)	p-value
					REF	SNV	+ Strand	- Strand	+ Strand	- Strand			
TX 8102	2010	3409	NS1	R314G	A	G	30	47	1138	1388	2603	2.96%	1.11E-16
		3435	NS1	R322	C	T	12	14	1045	1527	2598	1.00%	2.73E-06
		3456	NS1	C329	T	C	10	26	1129	2417	3582	1.01%	1.23E-07
		3646	NS2A	L41	C	T	18	22	1466	1683	3189	1.25%	1.65E-11
		3909	NS2A	E128D	G	T	23	8	1894	681	2606	1.19%	1.10E-08
		4251	NS2B	V11	C	T	20	13	1609	747	2389	1.38%	1.03E-10
		7584	NS4B	A223	C	T	19	14	1654	1856	3543	0.93%	8.12E-07
		8232	NS5	L184	C	T	16	15	1572	1385	2988	1.04%	3.32E-07
		9432	NS5	R584	C	T	18	14	1603	1437	3072	1.04%	1.55E-07
		10408	3'UTR	-	C	T	38	29	1018	666	1751	3.83%	0.00E+00
		10435	3'UTR	-	C	T	15	10	863	579	1467	1.70%	4.77E-10
TX 8349	2011	1353	E	T129	C	T	46	30	1473	1725	3274	2.32%	4.74E-37
		1371	E	I135	C	T	14	20	1544	1928	3506	0.97%	2.06E-07
		1588	E	T208A	A	G	21	2	1652	217	1892	1.22%	4.12E-07
		2506	NS1	L13	C	T	20	20	1828	1456	3324	1.20%	3.78E-11
		<i>3537</i>	<i>NS2A</i>	<i>D4</i>	<i>C</i>	<i>T</i>	8	8	1616	540	2172	0.74%	0.00E+00
		4209	NS2A	N228	C	T	16	14	1365	1262	2657	1.13%	3.62E-08
		5841	NS3	D410	C	T	17	22	2167	1630	3836	1.02%	7.91E-09
		5967	NS3	A452	C	T	14	23	2090	1911	4038	0.92%	1.96E-07
		6871	NS4A	V135M	G	A	171	173	1463	1992	3799	9.06%	1.11E-16
		7183	NS4B	D90N	G	A	30	20	1691	1250	2991	1.67%	0.00E+00
		7637	NS4B	T241I	C	T	43	27	1709	2114	3893	1.80%	0.00E+00
		7878	NS5	V66	C	T	119	60	1309	709	2197	8.15%	0.00E+00
		<i>8811</i>	<i>NS5</i>	<i>N377</i>	<i>C</i>	<i>T</i>	8	6	2189	590	2793	0.50%	0.00E+00
		<i>9352</i>	<i>NS5</i>	<i>L558</i>	<i>T</i>	<i>C</i>	24	17	1493	831	2365	1.73%	0.00E+00
		10410	3'UTR	-	A	T	14	11	1439	1067	2531	0.99%	4.46E-06
		10582	3'UTR	-	G	T	31	8	1205	2110	3354	1.16%	3.19E-10
		10888	3'UTR	-	T	A	5	18	238	1612	1873	1.23%	6.74E-07
TX 8375	2011	59	5'UTR	-	T	C	31	0	1753	0	1784	1.74%	6.78E-13

*Boldface, SNVs which encode amino acid substitutions. Italics, both synonymous and non-synonymous SNVs which represent a reversion to the prototype NY99 sequence.

Table 5-6: Single nucleotide variants (SNVs) detection in natural WNV populations, p -value < 0.05 (Bonferroni-corrected) (Continued)

Strain	Year	Position*	Gene	Residue	Mutation		SNV Count		Reference Count		Depth-of-Coverage	SNV Frequency (%)	p-value
					REF	SNV	+ Strand	- Strand	+ Strand	- Strand			
TX 8375	2011	96	5'UTR	-	G	A	62	1	1101	45	1209	5.21%	1.11E-16
		810	M	T23	C	T	25	35	1684	1709	3453	1.74%	6.77E-23
		1243	E	R93G	A	G	20	16	1398	1971	3405	1.06%	1.19E-08
		1374	E	K136	G	A	347	393	1090	1347	3177	23.29%	1.11E-16
		2222	E	A419V	C	T	69	239	1441	1533	3282	9.38%	0.00E+00
		3366	NS1	R299	C	T	80	39	1426	661	2206	5.39%	0.00E+00
		3409	NS1	R314G	A	G	85	126	1134	1354	2699	7.82%	2.22E-16
		3444	NS1	T325	T	C	31	54	1021	1813	2919	2.91%	0.00E+00
		3864	NS2A	F113	C	T	106	141	1570	1303	3120	7.92%	1.70E-228
		4515	NS2B	M99I	G	A	158	167	1442	1613	3380	9.62%	3.33E-16
		5667	NS3	I352	C	T	26	19	1873	1404	3322	1.35%	2.25E-13
		5784	NS3	Y391	C	T	149	112	2010	1339	3610	7.23%	7.39E-241
		6036	NS3	Y475	T	C	88	95	1276	1864	3323	5.51%	0.00E+00
		7516	NS4B	L201	T	C	114	84	1652	1238	3088	6.41%	2.54E-173
		8247	NS5	P189	G	A	58	85	1436	1550	3129	4.57%	0.00E+00
		8913	NS5	L411	G	A	52	66	1606	2525	4249	2.78%	0.00E+00
		10777	3'UTR	-	T	C	22	21	759	1424	2226	1.93%	0.00E+00
TX 8551	2012	93	5'UTR	-	C	T	29	0	1068	57	1154	2.51%	4.44E-16
		123	C	G9	C	T	12	2	279	178	471	2.97%	3.32E-09
		1983	E	P339	T	C	114	110	1408	1276	2908	7.70%	0.00E+00
		6471	NS4A	S1	T	A	197	133	1777	1150	3257	10.13%	3.33E-16
		7259	NS4B	V115A	T	C	196	72	1260	838	2366	11.33%	0.00E+00
		<i>7581</i>	<i>NS4B</i>	<i>T222</i>	<i>C</i>	<i>T</i>	212	252	1412	1363	3239	14.33%	0.00E+00
		7599	NS4B	H228	C	T	30	16	1787	1374	3207	1.43%	4.55E-15
		7644	NS4B	T243	A	G	97	206	1628	1642	3573	8.48%	2.22E-16
		10888	3'UTR	-	T	A	1	27	309	1592	1929	1.45%	1.04E-09
TX 8560	2012	93	5'UTR	-	C	T	348	14	817	32	1211	29.89%	3.33E-16
		371	C	T91I	C	T	41	47	1518	1803	3409	2.58%	1.11E-16
		450	C	I118	C	T	16	18	1793	1198	3025	1.12%	6.26E-09

*Boldface, SNVs which encode amino acid substitutions. Italics, both synonymous and non-synonymous SNVs which represent a reversion to the prototype NY99 sequence.

Table 5-6: Single nucleotide variants (SNVs) detection in natural WNV populations, p -value < 0.05 (Bonferroni-corrected)
(Continued)

Strain	Year	Position	Gene	Residue	Mutation		SNV Count		Reference Count		Depth-of-Coverage	SNV Frequency (%)	p-value
					REF	SNV	+ Strand	- Strand	+ Strand	- Strand			
TX 8560	2012	698	pr	Y78F	A	T	35	29	1656	1917	3637	1.76%	0.00E+00
		732	pr	R89	C	T	23	20	1606	2664	4313	1.00%	5.59E-09
		842	M	S34L	C	T	19	13	1621	1771	3424	0.93%	1.28E-06
		950	M	V70A	T	C	22	29	1495	1280	2826	1.80%	0.00E+00
		1734	E	G256	C	T	9	27	1455	1639	3130	1.15%	2.78E-09
		2069	E	N368S	A	G	105	85	1510	1150	2850	6.67%	5.43E-170
		2439	E	V491	T	C	12	25	1830	1069	2936	1.26%	7.23E-11
		2651	NS1	S61F	C	T	25	78	1877	2110	4090	2.52%	7.85E-53
		3129	NS1	V220	C	T	10	26	1415	2748	4199	0.86%	1.82E-06
		3249	NS1	Y260	C	T	25	47	1304	1175	2551	2.82%	3.72E-37
		3531	NS2A	N2	C	T	6	7	1429	767	2209	0.59%	0.00E+00
		4764	NS3	H51	T	C	21	13	1584	1552	3170	1.07%	6.76E-08
		5925	NS3	G438	A	G	57	54	1947	2411	4469	2.48%	1.11E-16
		6735	NS4A	V89	C	T	25	22	1746	1846	3639	1.29%	1.64E-13
		7593	NS4B	L226	T	C	6	10	1483	1477	2976	0.54%	0.00E+00
		8811	NS5	N377	C	T	56	28	1837	783	2704	3.11%	0.00E+00
		8820	NS5	T380	C	T	80	28	2017	874	2999	3.60%	0.00E+00
		9028	NS5	H450Y	C	T	7	26	699	2290	3022	1.09%	2.46E-08
		10281	NS5	N867	C	T	15	20	1363	1963	3361	1.04%	4.32E-08
		10414	3'UTR	-	T	C	13	14	1060	1005	2092	1.29%	3.68E-08
		10435	3'UTR	-	C	T	25	21	1064	719	1829	2.52%	0.00E+00
		10626	3'UTR	-	C	T	33	14	1299	900	2246	2.09%	1.26E-21
TX 8567	2012	93	5'UTR	-	C	T	255	9	1063	40	1367	19.31%	5.55E-16
		5286	NS3	T225	C	T	37	36	1770	1624	3467	2.11%	0.00E+00
		6631	NS4A	A55T	G	A	110	45	1801	1012	2968	5.22%	4.19E-120
		7672	NS4B	L253	C	T	18	30	1921	1685	3654	1.31%	3.72E-14
TX 8572	2012	93	5'UTR	-	C	T	158	5	1186	56	1405	11.60%	4.44E-16
		<i>10448</i>	<i>3'UTR</i>	-	<i>A</i>	<i>T</i>	145	148	839	488	1620	18.09%	0.00E+00
		10888	3'UTR	-	T	A	0	27	217	1614	1858	1.45%	3.86E-09

*Boldface, SNVs which encode amino acid substitutions. Italics, both synonymous and non-synonymous SNVs which represent a reversion to the prototype NY99 sequence.

Table 5-7: Conserved SNVs in natural WNV populations, related to **Table 5-6**

Strain	Year	Position*	Gene	Residue	Mutation		SNV Count		Reference Count		Depth-of-Coverage	SNV Frequency (%)	p-value
					REF	SNV	+ Strand	- Strand	+ Strand	- Strand			
TX 114	2002	93	5'UTR	-	C	T	67	20	425	134	646	13.47%	2.22E-16
TX 7558	2008	93	5'UTR	-	C	T	175	6	1487	71	1739	10.41%	9.30E-197
TX8102	2010	93	5'UTR	-	C	T	425	26	760	52	1263	35.71%	0.00E+00
TX 8551	2012	93	5'UTR	-	C	T	29	0	1068	57	1154	2.51%	4.44E-16
TX 8560	2012	93	5'UTR	-	C	T	348	14	817	32	1211	29.89%	3.33E-16
TX 8567	2012	93	5'UTR	-	C	T	255	9	1063	40	1367	19.31%	5.55E-16
TX 8572	2012	93	5'UTR	-	C	T	158	5	1186	56	1405	11.60%	4.44E-16
TX 1576	2003	732	pr	R89	C	T	74	76	1649	2249	4048	3.71%	1.11E-16
TX 8560	2012	732	pr	R89	C	T	23	20	1606	2664	4313	1.00%	5.59E-09
TX 8102	2010	3409	NS1	R314G	A	G	30	47	1138	1388	2603	2.96%	1.11E-16
TX 8375	2011	3409	NS1	R314G	A	G	85	126	1134	1354	2699	7.82%	2.22E-16
TX 114	2002	6871	NS4A	V135M	G	A	21	18	1137	1907	3083	1.27%	2.27E-11
TX 8349	2011	6871	NS4A	V135M	G	A	171	173	1463	1992	3799	9.06%	1.11E-16
TX 8349	2011	<i>8811</i>	<i>NS5</i>	<i>N377</i>	<i>C</i>	<i>T</i>	8	6	2189	590	2793	0.50%	0.00E+00
TX 8560	2012	<i>8811</i>	<i>NS5</i>	<i>N377</i>	<i>C</i>	<i>T</i>	56	28	1837	783	2704	3.11%	0.00E+00
TX 114	2002	10408	3'UTR	-	C	T	15	18	1348	1631	3012	1.10%	2.43E-08
TX 1576	2003	10408	3'UTR	-	C	T	57	48	1295	837	2237	4.69%	1.11E-16
TX 8102	2010	10408	3'UTR	-	C	T	38	29	1018	666	1751	3.83%	0.00E+00
TX 8102	2010	10435	3'UTR	-	C	T	15	10	863	579	1467	1.70%	4.77E-10
TX 8560	2012	10435	3'UTR	-	C	T	25	21	1064	719	1829	2.52%	0.00E+00
TX 114	2002	10777	3'UTR	-	T	C	4	10	150	189	353	3.97%	1.22E-10
TX 8375	2011	10777	3'UTR	-	T	C	22	21	759	1424	2226	1.93%	0.00E+00
TX 1153	2003	10888	3'UTR	-	T	A	0	30	349	2321	2700	1.11%	2.87E-07
TX 1171	2003	10888	3'UTR	-	T	A	0	29	340	1863	2232	1.30%	2.66E-08
TX 1175	2003	10888	3'UTR	-	T	A	0	36	332	1258	1626	2.21%	2.22E-16

*Boldface, SNVs which encode amino acid substitutions. Italics, both synonymous and non-synonymous SNVs representing a reversion to the prototype NY99 sequence.

Table 5-7: Conserved SNVs in natural WNV populations, related to **Table 5-6** (Continued)

Strain	Year	Position	Gene	Residue	Mutation		SNV Count		Reference Count		Depth-of-Coverage	SNV Frequency (%)	p-value
					REF	SNV	+ Strand	- Strand	+ Strand	- Strand			
TX 1576	2003	10888	3'UTR	-	T	A	0	22	242	1402	1666	1.32%	6.99E-07
M19433	2007	10888	3'UTR	-	T	A	0	39	307	1842	2188	1.78%	1.33E-15
TX 8349	2011	10888	3'UTR	-	T	A	5	18	238	1612	1873	1.23%	6.74E-07
TX 8551	2012	10888	3'UTR	-	T	A	1	27	309	1592	1929	1.45%	1.04E-09
TX 8572	2012	10888	3'UTR	-	T	A	0	27	217	1614	1858	1.45%	3.86E-09

*Boldface, SNVs which encode amino acid substitutions. Italics, both synonymous and non-synonymous SNVs representing a reversion to the prototype NY99 sequence.

Figure 5-1: Bayesian-inferred coalescent phylogenetic tree using the NGS-indexed FASTA sequences of all applied, regional Houston, Texas isolates. Phylogenetic relationships indicated in the NY99 (blue), NA/WN02 (orange), and SW/WN03 (red) U.S. genotypes and published Harris Co. paradigm groupings (Groups 1-10)^{63,351}. Scale bar indicates divergence time (in years).

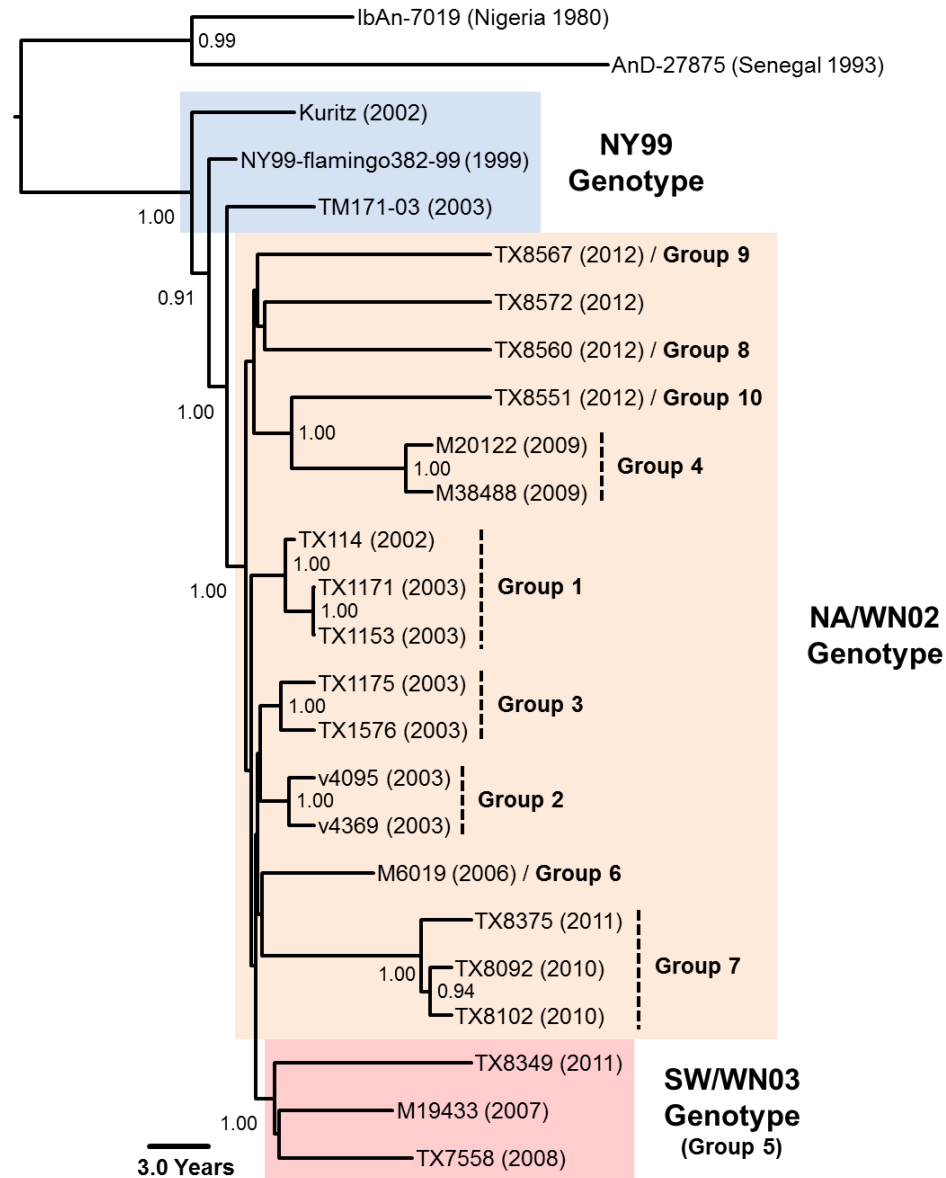


Figure 5-2: West Nile virus incidence in Harris Co., Texas, 2002-2012. Shannon entropies (S_n ; scale $\times 10,000$) indicated for all sequenced bird- (open circles) and mosquito pool-derived (filled circles) WNV isolates per year⁴⁷.

[WNV-positive mosquito pool data acquired from the Harris Co. Public Health and Environmental Services: Mosquito Control Division (via personal communication)]

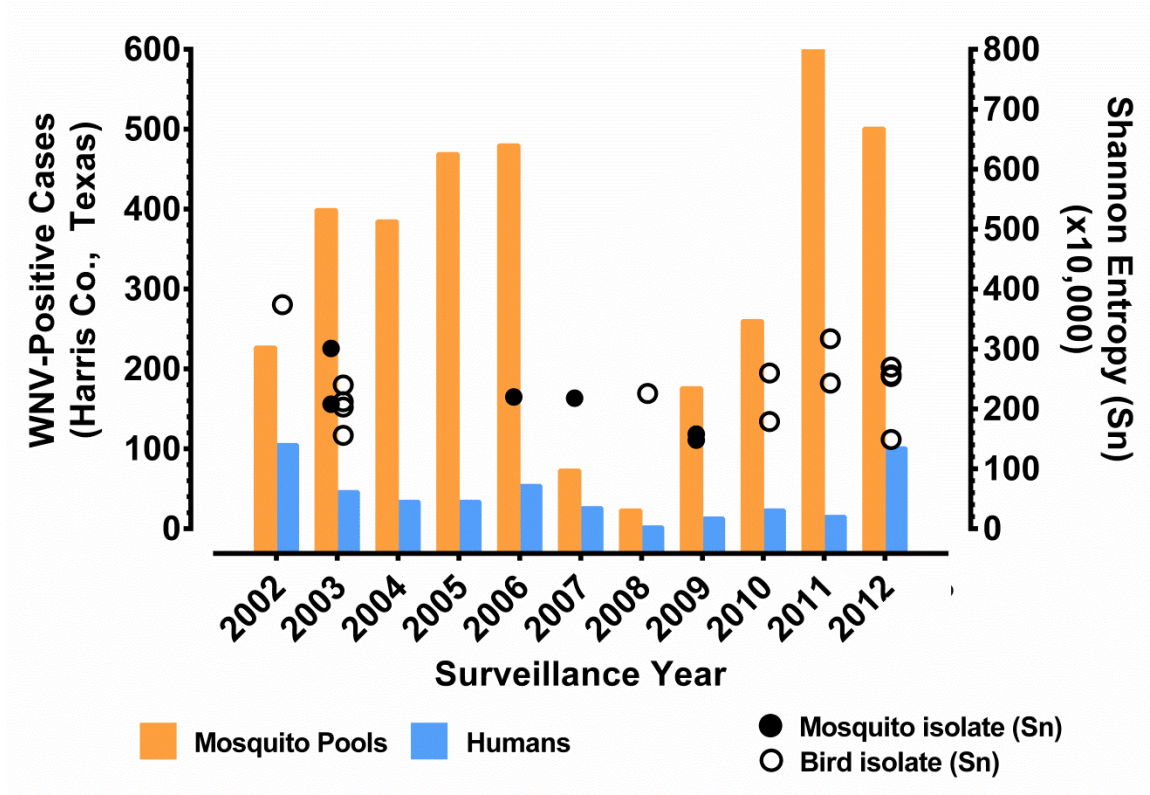


Figure 5-3: Host-specific genetic variation across the WNV genome for both A) bird- and B) mosquito pool-derived isolates per genomic region. Super-imposed values, nucleotide positions exhibiting >0.05 mean S_n (red dashed line) per host-type. Circles represent detected, conserved synonymous (open, black) and non-synonymous (filled red) SNVs in bird-derived isolates relative to their genomic position.

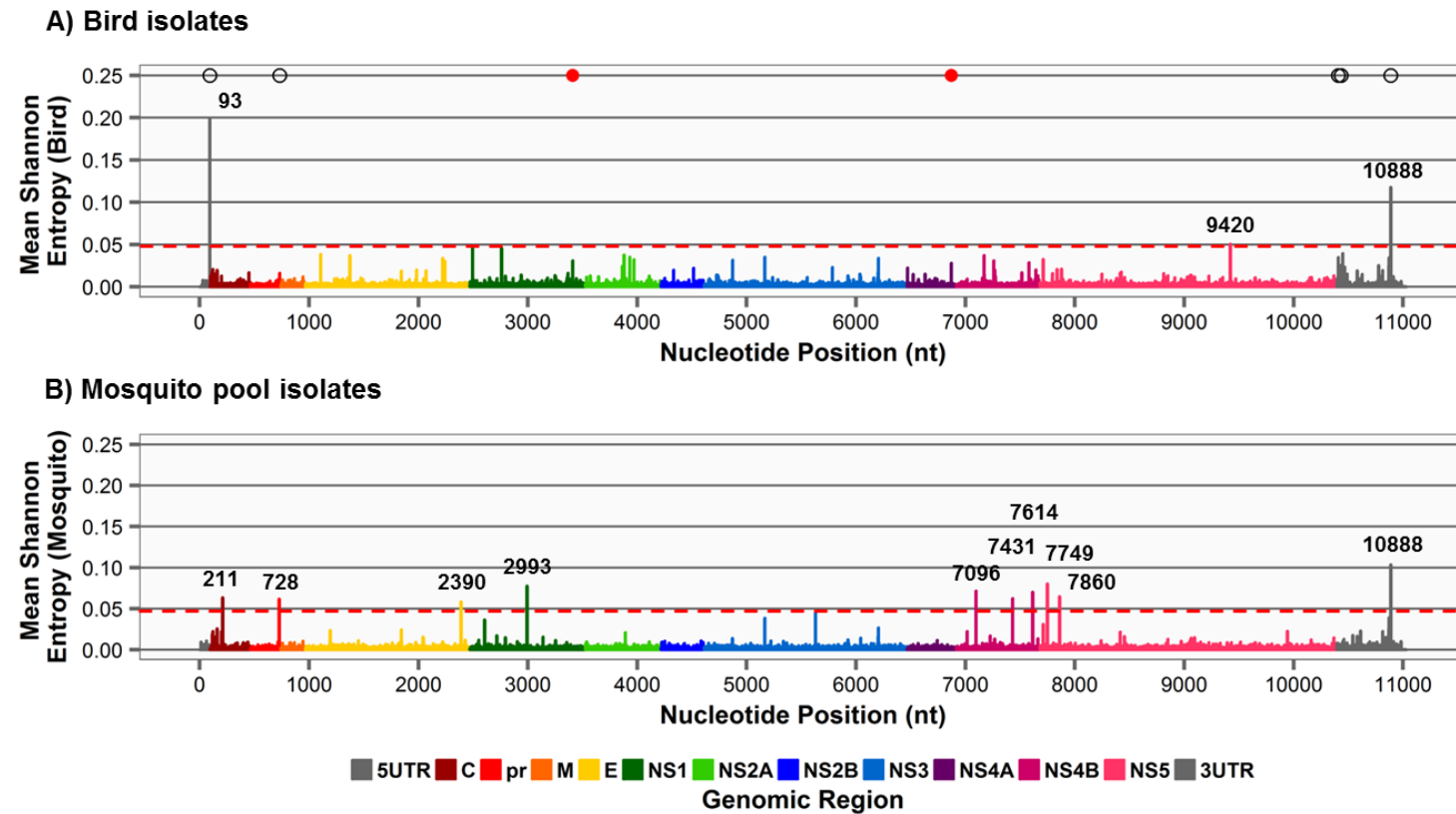


Figure 5-4: Detection of conserved SNVs in natural WNV populations. A) SNVs identified in >1 WNV isolate with both synonymous (black) and non-synonymous (red) variants indicated relative to their genomic. B) Proportion of both bird- and mosquito-derived isolates encoding 4 of the 7 highlighted SNVs conserved in the 5'/3'UTRs at 1.0 to >20% mutational frequencies or in the encoded consensus sequence (black). *Six of the eight detected SNVs at position 10888 were significant ($p < 0.05$) in SNVer alone.

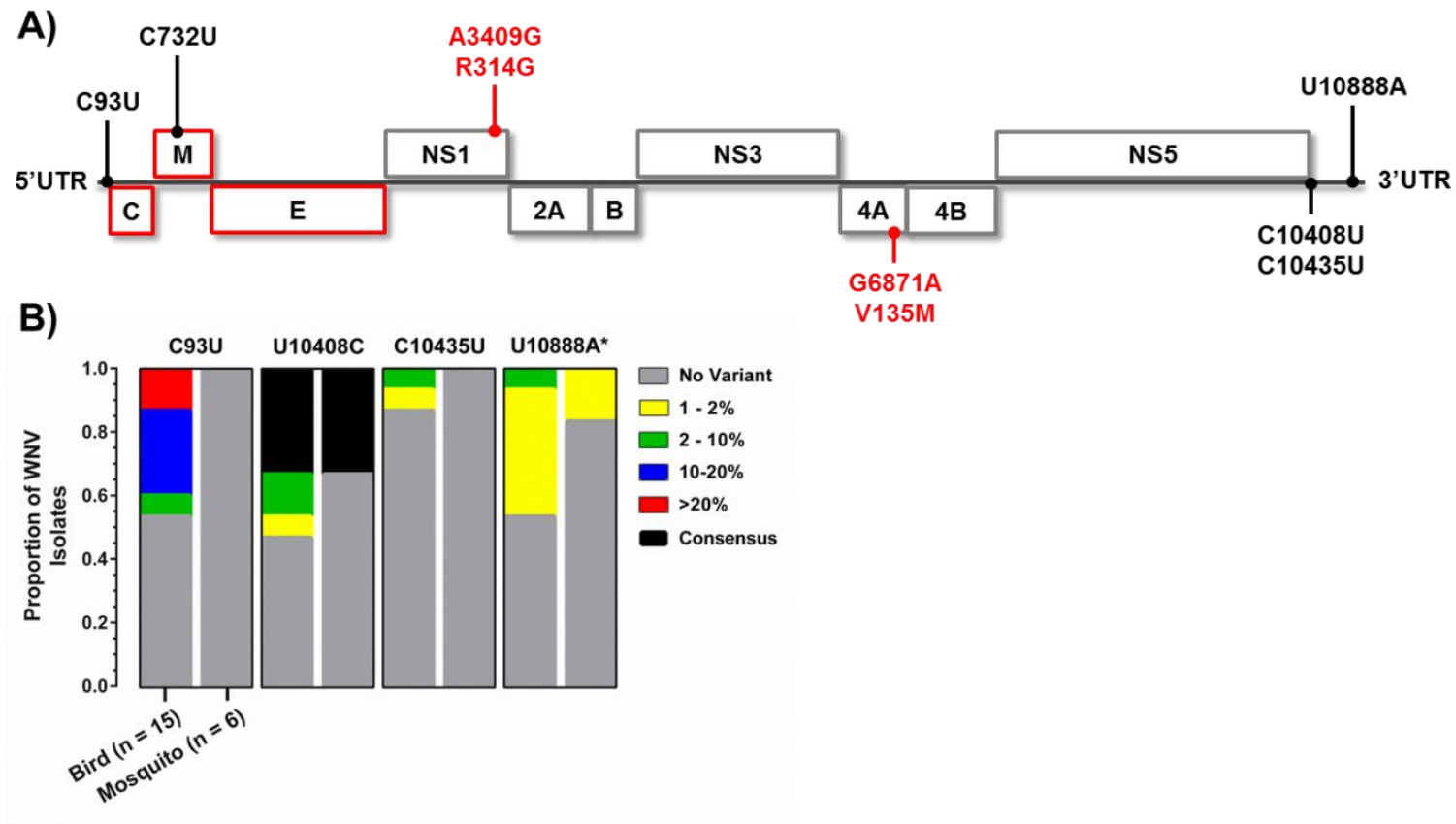


Figure 5-5: West Nile virus virulence is linked to viral RNA population diversity, related to **Fig. 5-6**. Mean calculated Shannon entropy (scale x10,000) is indicated per region of the WNV genome for all applied isolates from *low* (dark blue; 0) to *high* (dark red; ≥ 600) genetic diversity.

[***In vivo* i.p. LD₅₀ values for TX114, TX1576, TX1153, TX1175, and v4369 are derived from **Davis, C.T., et al. (2004) Virology**]⁵⁷

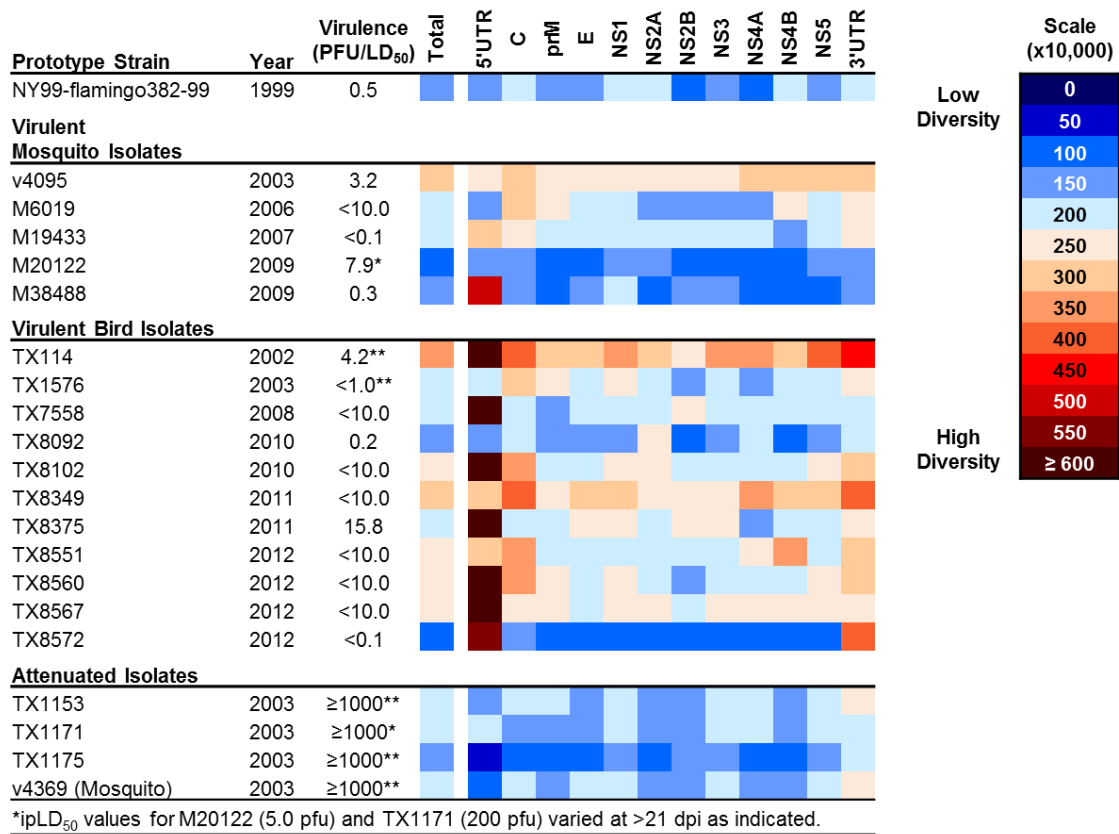
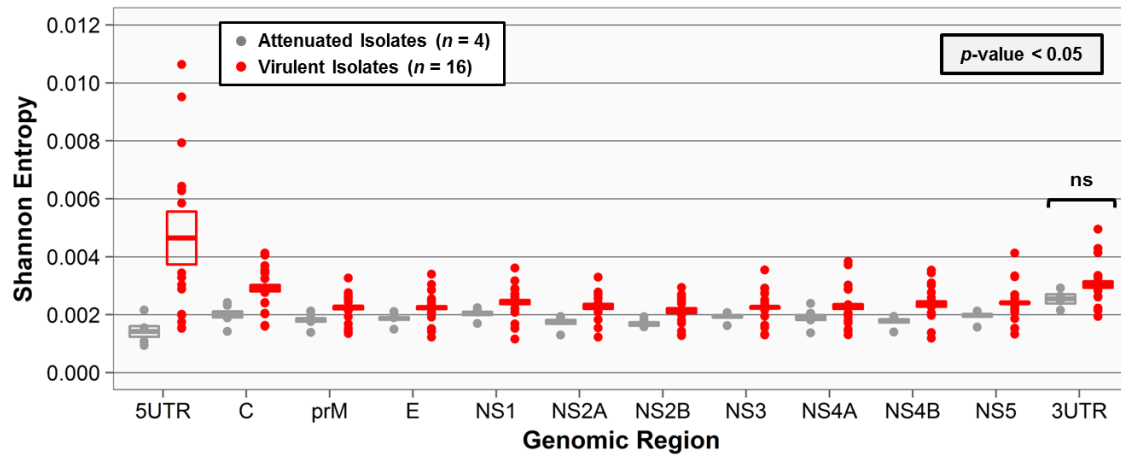


Figure 5-6: Mouse neuroinvasive phenotype is linked to increased WNV genetic diversity, related to **Fig. 5-5**. Mean Shannon entropy (\pm SEM) is indicated for both attenuated (grey) and virulent (red) isolates per region in the WNV genome. Dots, discrete mean values per isolate. Not significant, n.s.



CHAPTER 6

Role of the West Nile virus NS4B protein in modulation of the innate immune response

6.1 INTRODUCTION

Prior chapters identified natural WNV isolates in North America with amino acid substitutions in several virus-encoded genes. The NS4B protein was selected as a platform to investigate the potential role of amino acid substitutions in the phenotype of WNV infection. In-depth studies in multiple flaviviral systems have defined the functional role of the NS4B *N*-terminal domain in the antagonism of host innate Type-I IFN- α/β expression^{175,177}; however, the identification of attenuated, *in vivo* neurotropism linked to NS4B residues P38G and C102S further implicated the potential role of the NS4B protein in the regulation of host innate antiviral responses¹⁷⁸⁻¹⁸⁰.

In this chapter, application of IFN-competent human alveolar A549 cell culture provided an established *in vitro* model to interrogate the role of the NS4B protein in modulation of innate antiviral responses in the presence of intact Type-I IFN- α/β signaling and expression^{198,267,372,373}. Inclusion of well-characterized, NY99ic-engineered NS4B clones ($n = 14$) in this model developed a pipe-line to probe differential gene/protein expression linked to flavivirus-conserved NS4B residues (D35, P38, and Y45), *in vitro* passage-selected mutants (L97M, A100V, T116I, and T241A), cysteine residues (102, 120, 227, and 237), and the natural E249G substitution (TX1153 isolate) encoded in three distinct domains of the NS4B protein^{57,112,168,169,174,179,374} (**Fig. 6-1**). Furthermore, determination of *in vitro* multiplication kinetics and temperature-sensitive (*ts*) *in vitro* phenotype provided an initial benchmark to confirm published trends and link differential gene/protein expression with WNV multiplication kinetics^{300,301}. In addition,

application of both the attenuated C102S and P38G/T116I NS4B mutants relative to their non-attenuated alanine counter-parts: C102A and P38A established a mechanism to bridge NS4B residue-specific *in vivo* attenuation and trends in innate antiviral responses¹⁷⁸⁻¹⁸⁰.

6.2 RESULTS

6.2.1 *In vitro* platform: A549 cell culture

In order to establish base-line innate antiviral responses upon infection of A549 cell culture with either natural WNV isolates or NY99ic-derived mutants, 95% confluent A549 cultures (6.9×10^4 cells/cm²) were inoculated in triplicate at a MOI of 0.1 with NY99-flamingo382-99 (virulent wild-type), TX1153 (attenuated wild-type; encoding an E249G substitution), or the NY99ic-derived NS4B-P38G/T116I (+NS3-N480H mutation) mutant (**Table 6-1**)^{57,174,375,376}. Both the supernatant and cellular monolayer (per replicate) were collected at 6, 12, 24, and 36 hours post-infection (hpi) in order to interrogate differential *in vitro* multiplication kinetics and define the time-course of associated gene/protein expression in this initial A549 *in vitro* platform.

6.2.1.1 MULTIPLICATION KINETICS AND IN VITRO PHENOTYPE

In vitro plaque titration of collected supernatants confirmed similar multiplication kinetics for all natural and NY99ic-derived isolates at all time-points: 6, 12, 24, and 36 hpi in A549 cell culture (**Fig. 6-2A**). However, significant differences (p -value < 0.0001 ; $\alpha = 0.05$) in infectious viral titers were observed among all applied isolates for all time-points via a Two-Way ANOVA and Tukey's *post hoc* test. Exceptions ($p > 0.05$) included NY99 vs. TX1153 (24 hpi) and TX1153 vs. P38G/T116I (36 hpi) at the later time-points. In addition, plaque titration of 36 hpi supernatants at 41°C compared to 37°C

confirmed that none of the isolates exhibited temperature-sensitive (*ts*) *in vitro* phenotypes (**Fig. 6-2B**)^{300,301}. Based on these results, observed trends in gene/protein expression between NY99, TX1153, and P38G in the proposed *in vitro* platform could be due to underlying, differential multiplication kinetics in infectious virus particles.

6.2.1.2 MARKERS OF INNATE IMMUNE MODULATION: GENE AND PROTEIN EXPRESSION

Inclusion of both the attenuated TX1153 and P38G/T116I isolates in the *in vitro* platform provided a model to define the role of both natural and NY99ic-derived attenuated WNV populations in modulation of the host innate immune response. RNA was extracted from collected A549 cell suspensions using the Qiagen QIAshredder and RNeasy Mini kits and processed into cDNA with the BioRad® iScript™ Advanced cDNA Synthesis kit (756 ng RNA/reaction) via standard qRT-PCR technique. Resulting cDNA samples were applied to the BioRad® PrimePCR Flavivirus Infections H96 panel (9.45ng cDNA/well) with quantification of 40 genes using SYBR® Green on the CFX Connect™ real-time PCR detection system. Raw C_T values were normalized to *gapdH* gene expression prior to determination of fold gene expression (relative to both mock and NY99) using the $-\Delta\Delta C_T$ method in the Bio-Rad CFX Manager v3.1 software package^{332,333}.

Based on applied Bonferroni-corrected ($n = 6$ tests; $\alpha/n < p$ -value) Student's *t*-test statistical comparisons, optimal differential gene expression was detected at the 36 hpi time-point. In particular, no significant differences (p -value < 0.05) in fold-expression were detected for all genes in the PrimePCR panel at either 6 or 12 hpi relative to mock infection. Limited expression (≤ 2 -fold) of the *casp8*, *cd14*, *mapk11*, *nfkb1*, *tlr1*, and *tnfrsf1a* genes was confirmed for all isolates at both the 24 and 36 hpi time-points (**Fig. 6-3** and **Appendix A**). Furthermore, despite ≥ 2 -fold increase in *il1b*, *tlr2*, *tlr4*, and *tnf* gene expression at either/both 24 and 36 hpi time-points (relative to mock), no significant

differences were detected (**Appendix A**). In addition, both *cd86* and *nos2* exhibited consistent fold-expression below the limit-of-detection (mean C_T values ≥ 35 cycles) at all time-points supporting exclusion from subsequent analyses.

NS4B-P38G/T116I infection resulted in the significant ≥ 5.28 -fold up-regulation (p -value < 0.05 ; relative to NY99) of select genes involved in Toll-like receptor (TLR)-3 and TLR7 signaling (*irf7* and *tlr3*), RIG-I-like receptor (RLR) signaling (*ddx58* and *ifih1*), RNaseL (*oas3*) activation, and IFN-stimulated gene expression: *ifl44*, *ifl44l*, *ifit1*, *ifit3*, and *isg15* at both 24 and 36 hpi (**Figs. 6-4** and **6-5**). Furthermore, detected induction of all significant gene expression (relative to mock) at 24 hpi was restricted to NS4B-P38G/T116I infection, excluding the *ccl5* gene. In addition, despite marked mean up-regulation (≥ 40 -fold) of *ccl5*, *ifnb1*, *il1rl1*, *il6*, and *il8* in NS4B-P38G/T116I infection at 36hpi, deviation among biological replicates resulted in no significant differences. In comparison, both NY99 and TX1153 demonstrated limited differential gene expression at 24 hpi in the *in vitro* A549 model; however, TX1153 exhibited significant 2.08- to 11.57-fold down-regulation (p -value < 0.05) for 14 genes involved in cytokine/chemokine induction and IFN-stimulated gene expression (relative to NY99 infection) at 36 hpi (**Fig. 6-5**).

In order to corroborate the above RNA results, culture supernatants collected at 6, 12, 24, and 36 hpi were applied to the BioRad® Bio-Plex Pro™ Human Cytokine 27-plex panel using the Bio-Plex® 200 detection system linking downstream cytokine/chemokine protein production with observed trends in gene regulation. Detection of raw protein concentrations (pg/mL) utilized the BioRad® Bio-Plex Manager™ software package. Statistical comparisons were conducted in GraphPad Prism v6.05 using the Kruskal-Wallis One-Way ANOVA and Dunn's *post-hoc* test ($\alpha = 0.05$); however, due to the limited sample size ($n = 3$) and high intra-replicate variance, no significant relationships were detected (**Appendix B**). In addition, production was not detected for nine proteins: eotaxin (CCL11), FGF Basic, GM-CSF, IL-1R α , IL-2, IL-5,

IL-15, MIP-1 α (CCL3), and TNF- α at all applied time-points. Protein expression ≥ 4 -fold (relative to mock) was limited to NS4B-P38G/T116I infection at 36 hpi (excluding RANTES); however, up-regulation of several proteins: IL-6, IL-8, IP-10 (CXCL10), GCS-F, MIP-1 β , and RANTES (CCL5) was detected—all of which are associated with the induction of an innate pro-inflammatory immune response consistent with observed gene expression trends (**Fig. 6-6**).

Based on these initial results (at a MOI of 0.1), optimal differential gene/protein expression occurred at the 36 hpi time-point for both wild-type and NY99ic-derived isolates. In particular, both the attenuated wild-type TX1153 isolate and NY99ic-derived NS4B-P38G/T116I mutant exhibited significant (p -value < 0.05) modulation of the host innate antiviral responses linked to disparate TLR and RLR signaling, IFN-stimulated gene expression, and pro-inflammatory cytokine/chemokine secretion. Selection of gene-targets from this initial *in vitro* model provided an optimized platform to interrogate the role of specific residues and/or domains in the WNV NS4B protein as contributors to host innate immune responses.

6.2.2 Role of the NS4B protein in the WNV innate immune response

Prior phenotypic studies identified the role of 14 natural, conserved residues or *in vitro* passage-selected mutations in the NS4B protein as potential molecular determinants linked to *in vivo* neuroinvasive and/or neurovirulent WNV phenotype in the 3-5 week old NIH Swiss or Swiss Webster mouse model (**Table 6-1**)^{112,174,179,180}. Incorporation of all 14 NS4B mutants in the optimized A549 platform provided an *in vitro* model to probe both residue and domain-dependent modulation of the host innate antiviral responses (**Figs. 6-7** and **6-8**). In addition, inclusion of the virulent P38A and C102A NS4B mutants and their attenuated P38G/T116I (+NS3-N480H) and C102S counter-parts established a mechanism to link differential gene/protein expression to published *in vivo* WNV

phenotype (**Table 6-1**). In brief, (identical to the prior initial platform) 95% confluent A549 cell monolayers were inoculated in quadruplicate at a MOI of 0.1 with all 14 NY99ic-derived NS4B mutants or NY99ic positive control. Both culture supernatants and monolayers were harvested at 36 hpi and applied to the defined *in vitro* multiplication kinetics, quantitative RT-PCR, and Bio-Plex Pro™ methods.

6.2.2.1 MULTIPLICATION KINETICS AND IN VITRO PHENOTYPE

Plaque titration of culture supernatants (incubated at 37°C) collected at 36 hpi confirmed limited deviation in multiplication kinetics $\leq 1 \log_{10}$ pfu/mL relative to NY99ic (**Table 6-2**); however, a temperature-sensitive (*ts*) *in vitro* phenotype was confirmed for the C102S mutant with a $3.32 \log_{10}$ pfu/mL reduction (*p*-value < 0.0001) in viral titer at 41°C consistent with prior published studies¹⁸⁰. Plaque morphologies also varied between clones at 37°C with small (*sp*; <1 mm in diameter) and large plaque (*lp*; >3 mm) phenotypes identified for multiple NS4B clones with no clear link to observed multiplication kinetics (**Table 6-2**) or published *in vivo* phenotype (**Table 6-1**). In conclusion, despite the indicated exceptions, overall multiplication kinetics and *ts* phenotypes remained conserved between NS4B mutants in the *in vitro* A549 expression platform as reported in prior Vero cell culture-based studies^{174,179,180}.

6.2.2.2 DIFFERENTIAL GENE/PROTEIN EXPRESSION IS BOTH RESIDUE- AND DOMAIN-DEPENDENT

Based on detected base-line gene expression in the initial platform, 19 of the 40 targeted genes in the BioRad® PrimePCR™ Flavivirus Infections H96 panel were applied to a custom PrimePCR™ panel focusing on differential gene expression in TLR and RLR signaling, Type I IFN- α/β induction, and hallmark cytokines/chemokines implicated in acute WNV infection (**Fig. 2-3**). In particular, inclusion of *irf7* in this panel linked *myd88* expression with TLR7-dependent modulation of IFN-stimulated gene (ISG)

expression^{276,377}; furthermore, substitution of *trim5* with *trim25* evaluated differential TRIM25 induction as an upstream regulator of RIG-I ubiquitination and subsequent activation^{378,379}. Following application of all processed samples to the custom PrimePCR™ panel—as defined above—raw C_T values (per gene) were normalized to *gapdH* and either mock or NY99ic expression using the $-\Delta\Delta C_T$ method in the BioRad® CFX Manager v3.1. Statistical comparisons applied a Bonferroni-corrected *p*-value generated from individual Student's *t*-tests ($\alpha = 4.17 \times 10^{-4}$; $n = 120$ independent tests).

Detected differential gene expression trends remained conserved for applied NS4B mutants within defined protein domains: N-terminus, transmembrane domain (TMD) 3, and TMD5 relative to both mock and NY99ic infection (**Figs. 6-7** through **6-9**). Exceptions included amino acid substitutions L97M, A100V, and T241A which exhibited limited deviation in gene expression from NY99ic (**Fig. 6-9** and **Appendix C**). Besides these mutations, all engineered substitutions in the NS4B N-terminus (in general) resulted in modulated ISG expression and cytokine/chemokine induction with significant (*p*-value < 0.05) down-regulation of *ifnb1*, *ifit3*, *il6*, and *il8* relative to NY99ic infection. Similar trends were identified for all five NS4B substitutions (except T241A) located in both TMD3 and TMD5 with conserved down-regulation (*p*-value < 0.05) of *cxcl0*, *ifnb1*, *ifih1*, *ifit3*, *il6*, *il8*, and *il12a*; these results also parallel expression profiles for the NS4B-C102A/S mutants in the C-terminal portion of the N-terminus. Furthermore, the NS4B-E249G substitution in the C-terminal region of TMD5 resulted in ≥ 3 -fold down-regulation of all genes in the custom PrimePCR™ panel (relative to NY99ic) excluding *irf3*, *irf7*, *myd88*, *oas3*, *tlr3*, *tnfsf10*, and *trim25*. Several NS4B mutants also exhibited limited or insignificant modulation of several targeted genes: *ddx58*, *irf3*, *irf7*, *myd88*, *oas3*, *tlr3*, and *trim25* involved in TLR and RLR signaling relative to either NY99ic or mock infection (**Fig. 6-9** and **Appendix C**).

Paired application of the BioRad® Bio-Plex Pro™ panel identified domain-dependent trends in cytokine/chemokine production consistent with NS4B-modulated

gene expression. Overall, 16 of the 27 protein targets in the Bio-Plex Pro™ panel exhibited expression below the limit-of-detection; however, ≥ 2 -fold down-regulation of IL-6, IL-8, IP-10 (CXCL10), and RANTES (CCL5) expression (relative to NY99ic) was observed for all NS4B mutants excluding L97M, A100V, and T241A (**Fig. 6-10** and **Appendix D**). In contrast, both IL-1 β and TNF- α were undetectable for all NS4B mutants consistent with detected gene-based expression. Implementation of a Kruskal-Wallis One-Way ANOVA ($\alpha = 0.05$) and Dunn's *post-hoc* test provided limited statistical support for all proteomic trends; however, consistent, base-line induction of IL-12(p70), MCP-1 (CCL2), and VEGF in all applied samples supported identified NS4B domain- and residue-dependent differential protein expression trends.

6.2.2.3 NS4B PROTEIN-LINKED ATTENUATION AND INNATE IMMUNE EXPRESSION

Ex vivo and *in vivo* phenotypic studies in multiple platforms have defined the role of both NY99ic-derived P38G/T116I (+NS3-N480H) and C102S NS4B protein substitutions in viral attenuation^{178-180,295}. Inclusion of paired attenuated and non-attenuated (P38A and C102A) NS4B mutants at these residues in Specific Aim 3 provided a mechanism to probe critical *in vitro* molecular determinants conferring NS4B-linked *in vivo* WNV attenuation. Initial *in vitro* multiplication kinetics confirmed published *in vitro* phenotypes with reduced infectious titers and small plaque (*sp*) morphologies in both P38G and P38A infection—relative to NY99ic; in addition, the NS4B-C102S mutant exhibited temperature-sensitive (*ts*) *in vitro* multiplication kinetics consistent with prior studies despite loss of the *sp* phenotype retained in the NS4B-C102A mutant^{179,180}. Based on detected gene expression profiles, NS4B-P38G/T116I exhibited ≥ 2 -fold up-regulation of all applied genes in the custom PrimePCR™ panel except *irf3*, *myd88*, *nfkb1*, *nfkb2*, and *trim25*; in particular, the NS4B-P38G/T116I mutant demonstrated significant (p -value < 0.05) up-regulation of *ddx58*, *ifit3*, and *oas3* relative

to P38A infection (**Fig. 6-11A**). In contrast, no differences in differential gene expression were detected between the C102A and C102S mutants in the A549 *in vitro* expression platform (**Fig. 6-11B**); limited non-significant differences in protein production further supported these trends.

6.2.3 Role of the NS4B protein in natural WNV isolates

Inclusion of 10 additional WNV isolates encoding natural NS4B substitutions—in addition to a total of 41 other amino acid substitutions in the polyprotein of these 10 isolates (**Table 6-4**)—in the *in vitro* A549 expression model enabled initial affirmation of the identified trends using the NY99ic reverse genetics platform in natural NS4B-mediated modulation of host innate antiviral responses (**Table 6-3**). In addition, the 2003 TX1175 Harris Co. isolate was also included as an attenuated control encoding only Capsid (C) and 3'UTR coding mutations. Identical to the aforementioned NY99ic-derived *in vitro* studies, culture supernatants were collected at 36 hpi to determine differential multiplication kinetics and cytokine/chemokine expression in all applied isolates. Overall, significant (p -value < 0.05) variation in multiplication kinetics was confirmed for multiple isolates at both 37°C and 41°C (relative to NY99) despite no detectable temperature-sensitive (*ts*) phenotypes; in particular, all isolates encoding the NS4B-E249G substitution exhibited consistent reductions in viral titers (**Table 6-5**). In contrast, identified trends in detected pro-inflammatory cytokine/chemokine expression remained variable with limited, conserved domain- (*N*-terminus or TMD5) or position-dependent trends (**Fig. 6-12** and **Appendix E**); furthermore, application of a Kruskal-Wallis One-Way ANOVA and Dunn's *post-hoc* test detected no significant (p -value < 0.05) differences. Despite these limitations, isolates encoding both *N*-terminal S14I and N26T mutations resulted in the potent ≥ 30 -fold induction of IFN- γ , IL-6, IP-10, MIP-1 β , and RANTES relative to mock infection; however, these trends varied with additional

compensatory NS4B substitutions (*e.g.*, S14I + I176T versus S14I + I240M). Similar *in vitro* trends were identified in the C-terminal region of TMD5—*e.g.*, E249G substitution. In contrast, the attenuated TX1175 isolate control with no NS4B substitutions exhibited similar cytokine/chemokine expression to NY99 except for limited, non-significant (<2-fold) down-regulation of both IL-6 and RANTES demonstrating the significance of substitutions in the NS4B on cytokine/chemokine activity.

In conclusion, based on the consolidated results from the applied *in vitro* A549 platform, substitution of several residues in the NS4B protein *N*-terminal, TMD3, and TMD5 domains resulted in significant (p -value < 0.05), region-conserved impairment of ISG expression (**Fig. 6-9**), cytokine/chemokine induction (**Fig. 6-10**), and T cell recruitment/signaling. These results reaffirm the published role of the NS4B protein in Type-I IFN- α/β modulation and identified differential IL-6, IL-8, IP-10 (CXCL10), and RANTES (CCL5) expression as novel residue- and domain-dependent molecular markers for NS4B-linked innate antiviral regulation during acute WNV infection. Despite these trends, initial application of natural isolates (encoding either *N*-terminal or TMD5 substitutions) in this *in vitro* platform provided limited insight into the role of natural NS4B substitutions without additional phenotypic studies—due (in part) to additional substitutions in the encoded polyprotein sequence (**Table 6-4**).

6.3 DISCUSSION

The human A549 cell-line provided an established IFN-competent *in vitro* model to interrogate the role of both flavivirus-conserved residues and passage-adapted substitutions in the WNV NS4B protein as novel molecular determinants driving modulation of innate antiviral responses and published *in vivo* phenotype prior to directed *in vivo* studies^{198,267,380}. *In vitro* application of both wild-type and NY99ic-derived isolates (with defined *in vivo* phenotype) including the published, live-attenuated NS4B-

P38G/T116I (+NS3-N480H) mutant^{178,179,295} in an initial, pilot platform provided a proof-of-concept defining *in vitro* benchmarks for the applied NS4B studies.

Based on this optimized *in vitro* platform, several gene- and protein-targets involved in Toll-like and RIG-I-like receptor signaling, ISG responses, and pro-inflammatory cytokine/chemokine induction were linked to attenuated *in vivo* phenotype. In this regard, despite shared *in vivo* attenuation, detected antiviral responses to both the wild-type TX1153 isolate and NS4B-P38G/T116I mutant exhibited significant (p -value < 0.05) , opposing trends (**Figs. 6-4** and **6-5**). *In vivo* attenuation is not a unilateral but multi-factorial immune state; thus, discrepancies are not unexpected. However, perspective matters. Prior *in vivo* studies in multiple flaviviral systems affirmed the role of the NS4B N-terminus in Type-I IFN- α/β antagonism through inhibition of Jak-STAT signal transduction^{175,177,270}. *In vitro* expression of *inflb1* and ISGs with the NS4B-P38G/T116I mutant—in particular—either agreed (wild-type NY99; **Figs. 6-3 to 6-6**) or disagreed (NY99ic; **Figs. 6-9** and **6-10**) with these published trends depending on the applied positive control (as indicated)^{178,295}. Notably, the NY99ic reverse-genetics system (derived from the wild-type NY99 strain) is stabilized by two additional substitutions in the NS2A (H119Y) and NS3 (K104R) proteins (**Table 5-3**)¹¹². In this regard, the NS2A protein—in particular—has also been linked to IFN- β antagonism which may contribute to the higher base-line antiviral responses observed in all applied NY99ic-derived mutants (*e.g.*, NS4B-P38G/T116I clone) compared to wild-type isolates^{198,199}.

From this perspective, *in vivo* attenuation is linked to the potent down-regulation of IFN- β and pro-inflammatory cytokine expression in natural TX1153 infection; in contrast, hyper-activation of general innate antiviral responses is implicated in development of the attenuated NS4B-P38G/T116I *in vivo* phenotype. However, the presence of additional mutations in the natural TX1153 isolate limits direct attribution of detected trends to the encoded NS4B-E249G substitution alone^{57,174}. Furthermore, strain-dependent variation in viral replication kinetics or intra-cellular RNA loads could

confound cross-strain/mutant comparisons of innate antiviral responses in this *in vitro* model^{178,381}. However, despite limited discrepancies in observed *in vitro* phenotypes, these results affirmed published immune phenotypes described in multiple, WNV-permissive *in vitro* and *ex vivo* cell-lines (**Fig. 6-1** and **Table 6-2**)^{57,174,178,180}.

Subsequent application of established NY99ic-derived NS4B mutants—in several structural domains (**Figs. 6-7** and **6-8**)^{168,169}—provided an *in vitro* model to interrogate the role of distinct molecular determinants in NS4B-induced innate antiviral responses^{174,178-180}. Overall, integration of NY99ic-engineered NS4B mutants identified NS4B domain-conserved innate immune trends (**Figs. 6-9** and **6-10**)^{112,174,179,180}. In particular, substitutions in the *N*-terminal domain at residues D35, P38, Y45, and C102 resulted in the down-regulation of *ifnb1* and ISG expression consistent with proposed residue-specific IFN- α/β antagonism in the NS4B protein *N*-terminus¹⁷⁵. Furthermore, applied NS4B mutants in either TMD3 or TMD5 resulted in the significant (*p*-value < 0.05), conserved down-regulation of *ifnb1*, *ifih1* (MDA-5), *ifit3*, *il6*, *il8*, *cxcl10* (IP-10), and *il12a* (among others) compared to NY99ic infection. Exceptions included residues L97, A100, and T241 which elicited similar *in vitro* expression profiles to NY99ic infection; based on these trends, incorporation of conserved substitutions (based on the Gonnet-PAM250 matrix) or undefined structural limitations at these residues could confer the observed null *in vitro* phenotype^{168,169}. In contrast, mutation of residue E249 in the NS4B *C*-terminus (in TMD5) resulted in the potent down-regulation of all applied gene targets and encoded cytokines/chemokines: IL-6, IL-8, IP-10 (*cxcl10*), and RANTES (*ccl5*) with potential implications in modulated *in vitro* virus replication^{375,376}. Furthermore, initial *in vitro* studies using natural isolates implicated the complicated role of multiple amino acid substitutions in the polyprotein on the dysregulation of innate immunity (**Fig. 6-12**); in effect, this observation limits the direct application of natural WNV isolates to identify underlying determinants driving NS4B-induced innate antiviral responses without additional confirmation utilizing NY99ic-derived viruses encoding

selected natural NS4B substitution combinations plus non-NS4B substitutions (**Table 6-5** and **Appendix E**). However, the limited dysregulation of innate antiviral responses elicited via TX1175 infection (which does not encode an NS4B substitution) supports the critical role of encoded NS4B protein substitutions in the observed innate immune trends.

From this perspective, *in vitro* application of the NY99ic reverse genetics system confirmed the immunological role of the NS4B protein linked to IFN- β and ISG expression, pro-inflammatory cytokine/chemokine induction, and T cell recruitment (**Figs. 6-9** and **6-10**). In particular, *ifnb1*, *il6* (IL-6), *il8* (IL-8), *ccl5* (RANTES), and *cxcl10* (IP-10) all served as significant, molecular markers of NS4B-linked NY99ic *in vitro* innate immune phenotype similar to other *in vitro*, *ex vivo*, and *in vivo* models^{178,268,382-385}; however, limited detection of both IL-1 β and TNF- α expression (among other discrepancies) could be the result of cell- or tissue-specific antiviral responses^{242,288,291,295,385,386}. In addition, the absence of differential *irf3* and *myd88* gene expression relative to both NY99ic and mock infection limits the *in vitro* role of the NS4B protein in either IRF-3- or Myd88-dependent TLR3 and TLR7 regulation of downstream IFN- β and ISG expression^{373,387-389}. Interestingly, the NS4B C-terminal region (*i.e.*, E249G mutant) is linked to the potent regulation of all targeted genes—including *ifih1* (MDA-5), *ddx58* (RIG-I), *nfkb1*, and *nfkb2*. In effect, these data support the potential NS4B-triggered activation of redundant, non-canonical or upstream signaling cascades in IFN- α/β , or pro-inflammatory cytokine expression—such as IPS-1 (MAVS), IRF-5, and IRF-7^{274,276,278-281,373,377,390}. Further research is needed to map intrinsic structure-function relationships in the WNV NS4B protein (and others) to critical antiviral molecular targets and associated *in vivo* phenotypes.

Inclusion of the attenuated NS4B-P38G/T116I and C102S mutants and paired, virulent counterparts: P38A and C102A provided a means to evaluate the tentative link(s) between (1) genomic position; (2) published *in vivo* phenotype; and (3) detected innate antiviral responses¹⁷⁸⁻¹⁸⁰. Based on these data, the relationship between the NS4B protein

and *in vivo* attenuation is multi-factorial (**Fig. 6-11**). For example, *in vitro* NS4B-P38G/T116I infection exhibited ≥ 2 -fold up-regulation of all applied gene-targets except *irf3*, *myd88*, *nfbk1*, *nfbk2*, and *trim25* with significant (p -value < 0.05) over-expression of *ddx58*, *ifit3*, and *oas3* genes relative to the non-attenuated P38A mutant. These trends implicate potential IRF-3- and Myd88-independent induction of Type-I IFN- α/β -stimulated antiviral responses via RIG-I- and IRF-9/STAT2-dependent mechanisms in published *ex vivo* and *in vivo* P38G attenuation^{178,278,280,295,372,373}. In contrast, *in vitro* comparison of both C102A- and C102S-stimulated innate antiviral responses identified no significant, differential gene or protein expression; however, the NS4B-C102S mutant alone exhibited an attenuated, temperature-sensitive (*ts*) *in vitro* phenotype (**Table 6-2**)¹⁸⁰. In effect, residue-specific *in vivo* attenuation of the C102S mutant could stem from NS4B-linked antagonism of the WNV replication kinetics and *not* direct modulation of innate antiviral responses.

In conclusion, application of an NS4B reverse-genetics platform in the A549 *in vitro* expression model identified novel insights into the role of NS4B residue- and domain-dependent regulation of innate antiviral responses during acute WNV infection. However, additional research is needed in both applied *ex vivo* and *in vivo* (e.g., IFNAR^{-/-} mouse model) expression of this *in vitro* NS4B platform in order to confirm these NS4B protein-linked trends in a natural, multi-faceted immune environment.

Table 6-1: Natural and NY99ic-derived isolates applied in the *in vitro* A549 platform with documented *in vitro* and *in vivo* phenotypes

Strain	Temperature-sensitive (ts)*	i.p.LD ₅₀ (pfu)*	Passage history†	GenBank accession no.
NY99-flamingo382-99 (382-99; NY99)	-	< 1.0	V	AF196835
TX1153	ts	> 10,000	V	AY712945
NY99ic-derived NS4B clones‡				
NY99ic	-	< 1.0	V	-
D35E	-	0.4	V	-
P38A	-	7.0	V (p2)	-
P38G/T116I +NS3-N480H	-	> 10,000	V	-
Y45F	-	< 0.1	V	-
L97M	-	0.4	V	-
A100V	-	0.7	V	-
C102A	-	0.7	V	-
C102S	ts	> 10,000	V	-
T116I	-	0.7	V	-
C120S	-	0.7	V	-
C227S	-	2.0	V	-
C237S	-	6.0	V	-
T241A	n.d.	n.d.	Transfection	-
E249G	-	1.2	V	-

*Indicated temperature-sensitive (ts) *in vitro* phenotypes and intraperitoneal (i.p.) LD₅₀ information extracted from published studies^{174,179,180}. Not determined, n.d.

†Natural isolates or NY99ic-derived clones were passaged in Vero (V) cell culture with successive passages as indicated. Transfection, initial transfection stock in Vero cells.

‡Individual NY99ic-derived NS4B clones encode each indicated NS4B mutation in the NY99-flamingo382-99 infectious clone (NY99ic) backbone.

Table 6-2: NY99ic-derived NS4B mutant *in vitro* multiplication kinetics and temperature-sensitive (*ts*) phenotype in Vero cell culture

Clone	Plaque size†	Mean Infectious Titer (log ₁₀ PFU/mL)*					p-value‡
		37°C‡	SEM	41°C‡	SEM	Δ‡	
NY99ic	<i>mp</i>	7.67	0.07	7.37	0.09	-0.30	-
D35E	<i>mp</i>	7.75	0.03	7.47	0.08	-0.28	-
P38A	<i>sp</i>	7.70	0.08	7.49	0.08	-0.21	-
P38G/T116I§	<i>sp</i>	6.40	0.11	5.44	0.07	-0.96	-
Y45F	<i>mp</i>	7.66	0.06	7.54	0.01	-0.11	-
L97M	<i>mp</i>	7.49	0.16	7.42	0.03	-0.07	-
A100V	<i>mp</i>	7.75	0.04	7.57	0.01	-0.18	-
C102A	<i>sp</i>	7.17	0.04	6.69	0.05	-0.48	-
C102S	<i>mp</i>	6.87	0.01	3.55	0.43	-3.32	<0.0001
T116I	<i>mp</i>	6.94	0.03	6.69	0.02	-0.26	-
C120S	<i>lp</i>	7.42	0.02	6.94	0.03	-0.48	-
C227S	<i>mp</i>	7.23	0.10	6.92	0.11	-0.24	-
C237S	<i>lp</i>	7.69	0.07	7.63	0.01	-0.24	-
T241A	<i>lp</i>	7.53	0.02	7.24	0.13	-0.27	-
E249G	<i>lp</i>	7.23	0.00	6.77	0.05	-0.46	-

*Temperature-sensitive (*ts*) phenotype determined 72 hours post-infection and defined as a mean reduction (Δ) ≥2.0 log₁₀ PFU/mL in virus titer at 41°C.

†Plaque size determined at 37°C and designated as small (*sp*) < 1.0mm, medium (*mp*) 1-3mm, or large (*lp*) > 3.0mm.

‡One-Way ANOVA with Tukey's *post-hoc* test (α = 0.05) conducted to test for significant differences (*p*-value < 0.05) in mean infectious titers at both 37°C and 41°C and Δ values relative to NY99ic (indicated in boldface red).

§P38G/T116I data is derived from **Fig. 6-2**.

Table 6-3: Natural Texas WNV isolates encoding NS4B substitutions, related to **Table 6-4**.

Strain*	Substitution(s)†	Location	Isolation host‡	Year-of-isolation	Passage history§	GenBank accession no.
NY99 (382-99)	-	New York, NY	Chilean flamingo	1999	V	AF196835
TX1175	-	Harris Co., TX	Blue jay	2003	V	GU828000
TX8567	S14I	Harris Co., TX	Blue jay	2012	V	KC333381
TX AR12-3564	S14I + I176T	Ellis Co., TX	<i>Cx. quinquefasciatus</i>	2012	V	-
TX AR12-1648	S14I + I240M	Collin Co., TX	<i>Cx. quinquefasciatus</i>	2012	V	KC711058
TX8551	K24R	Harris Co., TX	Blue jay	2012	V	KC333377
TX8604	N26T	Harris Co., TX	House sparrow	2012	V	KC333387
TX AR12-9793	I240M	Collin Co., TX	<i>Cx. quinquefasciatus</i>	2012	V	-
TX AR12-10674	I202T + I240M	Collin Co., TX	<i>Cx. restuans</i>	2012	V	KC711059
TX AR12-1486	L244V	Denton Co., TX	<i>Cx. quinquefasciatus</i>	2012	V	KC711057
TX1153	E249G	Harris Co., TX	Mourning dove	2003	V	AY712945
TX1171	E249G	Harris Co., TX	Blue jay	2003	V	AY712946
TX AR10-6572	E249G	El Paso, TX	<i>Cx. tarsalis</i>	2010	V (p2)	JX015523

*NY99-flamingo382-99 (NY99; AF196835) and TX1175 included as a positive controls with no encoded NS4B substitution.

†Indicated substitutions encoded in the WNV NS4B protein *N*-terminus or transmembrane domain 5 (TMD5).

‡Mosquito pool isolates collected from *Culex* (Cx.) spp. pools.

§Strains were passaged from initial isolation stocks in Vero (V) cell culture with serial passages (p) as indicated.

Table 6-4: Encoded non-NS4B amino acid substitutions in applied, natural Texas WNV isolates, related to **Table 6-3**.

Strain†	NS4B‡ Substitution(s)	C*				prM			E					NS1					NS2A							
		32	36	104	119	55	140	156	51	89	123	126	159	179	35	41	51	98	236	308	326	46	52	58	92	95
NY99	-	R	S	K	A	V	V	V	A	A	T	I	V	K	Y	G	K	K	I	I	D	F	T	V	A	L
TX1175	-	S	A
TX8567	S14I	.	.	.	S	N	.	A	.	.	A	.	.	.	V	.	.	I	.	.	F
TX AR12-3564	S14I + I176T	.	.	.	S	V	N	.	A	V	.	.	I	.	.	F
TX AR12-1648	S14I + I240M	.	.	.	S	N	.	A	V	.	.	I	.	.	F
TX8551	K24R	.	.	R	A	V
TX8604	N26T	T	.	.	.	A	.	H	.	R	.	.	.	G
TX AR12-9793	I240M	M	A	R	I	.	.	.
TX AR12-10674	I202T + I240M	.	N	A	I	.	.	.
TX AR12-1486	L244V	T	A	R	T	.	.
TX1153	E249G	I	A
TX1171	E249G	I	A
TX AR10-6572	E249G	A	A	L

* C, capsid; prM, pre-membrane; E, envelope; NS, nonstructural: NS1 and NS2A.

† Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. Dots indicate no change from the NY99 isolate. TX1175 does not encode an NS4B substitution and is included as an attenuated non-NS4B negative control. Shaded regions indicate substitutions conserved in >1 isolate.

‡ Indicated substitutions encoded in the WNV NS4B protein *N*-terminus or transmembrane domain 5 (TMD5).

Table 6-4: Encoded non-NS4B amino acid substitutions in applied, natural Texas WNV isolates, related to **Table 6-3** (Continued).

Strain†	NS4B Substitution‡	NS2A*			NS2B	NS3			NS4A	NS4B							NS5									
		126	188	190	120	253	334	549	85	14	24	26	176	202	240	244	249	33	42	44	49	199	314	687	804	860
NY99	-	L	R	K	V	N	S	Y	A	S	K	N	I	I	I	L	E	I	H	R	V	R	K	A	A	A
TX1175	-
TX8567	S14I	T	H	.	I
TX AR12-3564	S14I + I176T	T	.	.	I	.	.	T
TX AR12-1648	S14I + I240M	T	.	.	I	M
TX8551	K24R	.	K	.	I	R	I
TX8604	N26T	T	T	Y
TX AR12-9793	I240M	.	K	M
TX AR12-10674	I202T + I240M	.	K	.	.	S	T	M	R	.	.	.
TX AR12-1486	L244V	F	V
TX1153	E249G	G	V	.
TX1171	E249G	G	L	.	D	V	.
TX AR10-6572	E249G	.	.	R	T	G	.	.	K	.	.	R	.	.	T

* NS, nonstructural: NS2A, NS2B, NS3, NS4A, NS4B, and NS5.

† Indicated amino acid changes are relative to the prototype NY99-flamingo382-99 strain [NY99; AF196836]. TX1175 does not encode an NS4B substitution and is included as an attenuated non-NS4B negative control. Dots indicate no change from the NY99 isolate.

Shaded regions indicate substitutions conserved in >1 isolate.

‡ Indicated substitutions encoded in the WNV NS4B protein *N*-terminus or transmembrane domain 5 (TMD5).

Table 6-5: Natural NS4B mutant *in vitro* multiplication kinetics and temperature-sensitive (*ts*) phenotype

Strain	Substitution(s)	Plaque size†	Mean Infectious Titer (Log ₁₀ PFU/mL)*				Δ‡	p-value‡
			37°C‡	SEM	41°C‡	SEM		
NY99 (382-99)	-	<i>lp</i>	7.80	0.03	7.30	0.03	-0.51	-
TX1175	-	<i>sp</i>	7.00	0.06	6.80	0.04	-0.19	-
TX8567	S14I	<i>lp</i>	7.84	0.05	7.46	0.02	-0.38	-
TX AR12-3564	S14I + I176T	<i>lp</i>	7.27	0.05	6.88	0.03	-0.39	-
TX AR12-1648	S14I + I240M	<i>lp</i>	7.64	0.05	7.40	0.04	-0.25	-
TX8551	K24R	<i>lp</i>	7.73	0.03	7.62	0.02	-0.11	-
TX8604	N26T	<i>lp</i>	7.29	0.08	6.82	0.31	-0.47	-
TX AR12-9793	I240M	<i>lp</i>	7.20	0.06	6.68	0.10	-0.52	-
TX AR12-10674	I202T + I240M	<i>mp</i>	7.13	0.10	6.90	0.04	-0.23	-
TX AR12-1486	L244V	<i>lp</i>	7.44	0.03	7.06	0.04	-0.38	-
TX1153	E249G	<i>sp</i>	6.98	0.01	6.03	0.03	-0.95	-
TX1171	E249G	<i>sp</i>	7.25	0.09	5.70	0.04	-1.55	<0.0001
TX AR10-6572	E249G	<i>mp</i>	6.91	0.06	6.66	0.03	-0.25	-

*Temperature-sensitive (*ts*) phenotype determined 72 hours post-infection and defined as a mean reduction (Δ) ≥2.0 log₁₀ PFU/mL in virus titer at 41°C.

†Plaque size determined at 37°C and designated as small (*sp*) < 1.0mm, medium (*mp*) 1-3mm, or large (*lp*) > 3.0mm.

‡One-Way ANOVA with Tukey's *post-hoc* test (α = 0.05) conducted to test for significant differences (p-value < 0.05) in mean infectious titers at both 37°C and 41°C and Δ values relative to NY99 (382-99) (indicated in boldface red).



			1	10	20	30	40	50
NY99 382-99	(1)	NEMGWLDKTKSDISSLFGQRIEVENFS---												
KUNV FLSDX	(1)	NEMGWLDKTKSDISGLFGQRIETKENFS---IGEFLLDLRPATAWSLYAVTTAVLTPLI												
SLEV Kern217	(1)	NEMGLEKTKSDIAKLFGSQPGSVFATRTTPWDISLDIKPATAWALYAAATVMVTPLI												
JEV Nakayama	(1)	NEYGMLEXTKADLKSMFMGGKTQASGLTG----LPSMALDLRPAATAWALYGSTVVLTPLL												
MVEV 51	(1)	NEYGMLERTKTDIRNLFGKSLIEENEVHIPPFDFFFTLDLKPATAWALYGSTVVLTPLI												
YFV Asibi	(1)	NELGMLEKTKEDLFGKK--NLIIP--SSASPWSWPDLDLKPAAAWTVYVGIVTMLS PML												
DENV1 Thai	(1)	NEMGLELETTKKDLGIGH-----VA AENHHHATMLDVLDRPASAWTLYAVATTITPMM												
DENV2 NGC	(1)	NEMGFLEKTKKDLGLGS-----ITTQ-QPENILDI DRPASAWTLYAVATTFVT PML												
DENV3 H87	(1)	NEMGLELETTKRDLGMSK-----EPGV-VSP TSYLDVDLHPASAWTLYAVATTVITPML												
DENV4 H241	(1)	NEMGLIEKTKTDFGFYQ-----VK---TETTILDVDLRPASAWTLYAVATTILT PML												
<i>TBEV Neudoerfl</i>	(1)	NEMGFLEKTKADLS TALWSERE E P----RP WSEWTNVDIQPARSWGTYVLVVS LFTPYI												
<i>POWV LB</i>	(1)	NELGYLEQT KTDISGLFRREDQGG---MVWD AWTNIDIQP ARSWGTYVLIVS LFTPYI												
<i>OHFV Bog</i>	(1)	NEMGFLEKTKADLS AVLWSERE E P----RVW SEWTNIDIQAKSWGT YVLVVS LFTPYI												
		** * : ** *												
		:**:::* *												
		* ..*												

				60	70	80	90	100	110
NY99	382-99	(57)		KHLITS	SDYINT	SLTSIN	VQASAL	PTLARG	FPFVDV
KUNV	FLSDX	(57)		KHLITS	SDYINT	SLTSIN	VQASAL	PTLARG	FPFVDV
SLEV	Kern217	(60)		KHLITTY	VNFSLT	AIASQA	GVLLGL	TNGMPT	AMDVSL
JEV	Nakayama	(57)		KHLITSE	YVTTSL	ASINSQ	AGSLVF	LPRGVP	FTDLDL
MVEV	51	(60)		KHLVTS	QYVTTT	SLASIN	AQAGSL	FTLPKG	IPFTDF
YFV	Asibi	(57)		HHWIKV	EYGNLS	LSGIAQ	SASVLS	FMMDKG	IPFMKM
DENV1	Thai	(54)		RHTIENT	TANISLT	AIANQA	AILMGL	DKGWPI	SKMDIG
DENV2	NGC	(53)		RHSIENS	SVVNST	LAIANQA	AVTLMG	LKGWPL	SKMDIG
DENV3	H87	(53)		RHTIEN	STANVS	LAIANQA	AVTLMG	LKGWPL	SKMDIG
DENV4	H241	(50)		RHTIENT	SANLSL	AIAINQA	AVTLMG	LKGWPL	HRMDLG
TBEV	Neudoerfl	(56)		IHQQLQ	TKIQQL	VNSAVA	SGAQAM	RDGGGA	PFPGV
POWV	LB	(56)		IHQQLQ	TKIQRL	VNSSVA	AGTQAM	RDGGGT	PFPGV
OHFV	Bog	(56)		IHQQLQ	TRIQQ	VNSAVA	SGAQAM	RDGGGT	PFPGV

		···120·····130·····140·····150·····160·····170··
NY99 382-99	(116)	TLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNAVVDGIVATDVPELERTTPIMQKKVG
KUNV FLSDX	(116)	TLLFCHYAYMVPGWQAEAMRSAQRRTAAGIMKNAVVDGIVATDVPELERTTPIMQKKVG
SLEV Kern217	(119)	MLLAHYAFMIPGWQAEAMRAAQRRTAAGIMKNAVVDGIVATDIPDLSPATPMTEKKMG
JEV Nakayama	(116)	VLATLHYGYMLPGWQAEALRAAQRRTAAGIMKNAVVDGIVATDVPELERTTPLMQKKVG
MVEV 51	(119)	ILVTLHYGYLLPGWQAEALRAAQRRTAAGIMKNAVVDGIVATDVPELERTTPQMQRRLG
YFV Asibi	(116)	GCAMLHWSLILPGIKAQQSKLAQRRVVFHGVAKNPVVDGNPTVDIEEAPEMPALYEEKLA
DENV1 Thai	(113)	LMLVAHYAIIIGPGLQAKATREAQRRTAAGIMKNPTVDGIVATDIDLPVY-YDAKFEKQLG
DENV2 NGC	(112)	FLVAHYAIIIGPGLQAKATREAQRRAAGIMKNPTVDGIVATDIDLPVY-YDPKFEKQLG
DENV3 H87	(112)	LLLVTHYAIIGPGLQAKATREAQRRTAAGIMKNPTVDGIMTIDIDLPVI-YDSKFEKQLG
DENV4 H241	(109)	VMLLVHYAIIIGPGLQAKATREAQRRTAAGIMKNPTVDGIVATIDLEPIS-YDPKFEKQLG
TBEV Neudoerfl	(115)	GLAALHLAIVVSGLEAELTQRAHKVFFSAMVRNPMVDGDVINPFGEGEAKPALYERKMS
POWV LB	(115)	ALAALHLAVVTSGLEAELTQRAHRAFFSAMVRNPMVDGEIINPDPGDPKALYERKMS
OHFV Bog	(115)	GLAAFHLAIVVSGLEAELTQRAHKVFFSAMVRNPMVDGDVINPFGDGEVVKPALYERKMS
		* . : * : * : * : * : * : * : *

NS4B Structure:	Loop 1 (Cytosol)
<p> </p>	<p> </p>

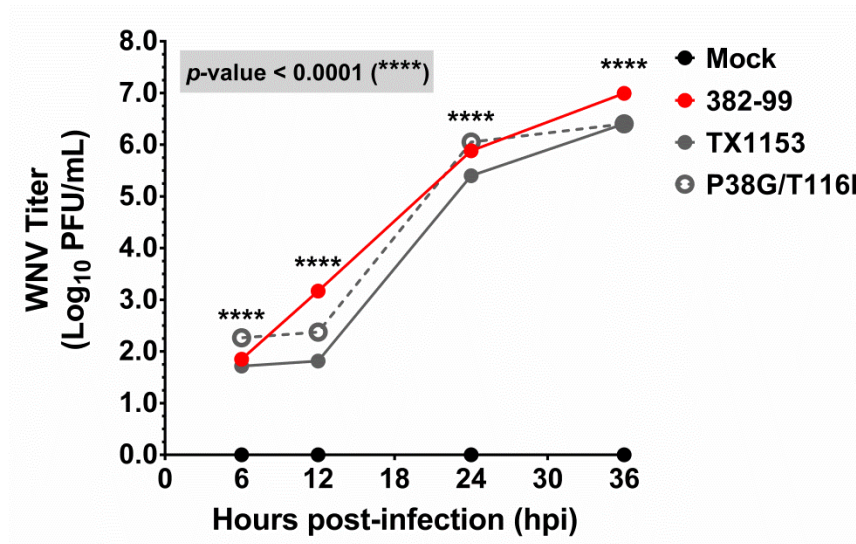
Figure 6-1: Flavivirus NS4B Protein Alignment (Continued)

		180.....190.....200.....210.....220.....230..
NY99 382-99	(175)		QIMLILVSLAAVVVNPSVKTVREAGILITAAAVTLWENGASSVWNATTAIGLCHIMRGG
KUNV FLSDX	(175)		QVMLILVSLAALVVVNPSVKTVREAGILITAAAVTLWENGASSVWNATTAIGLCHIMRGG
SLEV Kern217	(178)		QILLIAAAVLAVLVRPGICSIKEFGVLGSAALVTLIEGTAGVWNCCTTAVGLCNLMRGG
JEV Nakayama	(175)		QVLLIGVSVA AFLVNPVNTTVREAGVLVTAATLTLWDNGASAVWNSTTATGLCHVMRGS
MVEV 51	(178)		QILLVLASVA AVCVNPRITTIREAGILCTAAALTLWDNNSAAWNSTTATGLCHVMRGS
YFV Asibi	(173)		LYLLLALSLASVAMCRTPFSLAE GIVLASAALGPLIEGNTSLWNGPMAVSMTGVMRGN
DENV1 Thai	(171)		QIMLLILCTSQILLMRTTWALCESITLATGPLTTLWEGSPGREFWNTTIAVSMANIFRGS
DENV2 NGC	(170)		QVMLLVLCVTQVLMRRTTWALCEALTATGPITSLWEGNPGREFWNTTIAVSMANIFRGS
DENV3 H87	(170)		QVMLLVLCVAVQLLLMRTTWSALCEVLTATGPITSLWEGSPGKFWNTTIAVSMANIFRGS
DENV4 H241	(167)		QVMLLVLCAGQLLMRTTWAFCEVLTATGPVLTSLWEGNPGREFWNTTIAVSTANIFRGS
<i>TBEV Neudoerfl</i>	(174)		LVLATVLC LMSVVMNRTVASITEASAVGLAAAGQLLRPEADTLWTMPVACGMSGVVGRGS
<i>POWV LB</i>	(174)		LFLAIGLCIAAVALNRTAAAMTEAGAVAVAALGQLLRPEESWWTMPMACGMAGLVGRGS
<i>OHFV Bog</i>	(174)		LILAMILCFMSVVLNRTVPAVTEASAVGLAAAGQLLRPEADTLWTMPVACGLSGVVGRGS
			: . . : : * : . * * . : . **
NS4B Structure:			
		240.....250....
NY99 382-99	(234)		WLSCLSI TWTLIKNMEKPGLKR
KUNV FLSDX	(234)		WLSCLSI TWTLVKNMEKPGLKR
SLEV Kern217	(237)		WLAGMSITWTVYKNVDKPKGKR
JEV Nakayama	(234)		YLAGGSIAWTLIKNADKPSLKR
MVEV 51	(237)		WIAGASIAWTLIKNAEKPAFKR
YFV Asibi	(232)		YYAFVGV MYNLWKMTG---RR
DENV1 Thai	(230)		YLAGAGLAFSLMKSLGGG---RR
DENV2 NGC	(229)		YLAGAGLLFSIMKNNTNT---RR
DENV3 H87	(229)		YLAGAGLALSIMKSVGTG---KR
DENV4 H241	(226)		YLAGAGLAFSLIKNAQTP---RR
<i>TBEV Neudoerfl</i>	(233)		LWGFLPLGHRLWLRASGG---RR
<i>POWV LB</i>	(233)		LWGFLPVLHRIWLRTQGA---RR
<i>OHFV Bog</i>	(233)		LWGFLPLGHRLWLRTSGT---RR
			. : :
NS4B Structure:			
(Post-Cleavage)			... --COO ⁻ (Cytosol) --COO ⁻ (ER-Lumen)

NS4B protein sequences for several prototype mosquito-borne (**boldface**) and tick-borne (*italics*) flaviviruses were aligned using the ClustalW multiple sequence alignment algorithm (<http://www.ebi.ac.uk/Tools/msa/clustalo/>) on the EMBL-EBI server^{310,311}. Sequences are annotated relative to the prototype NY99-flamingo382-99 (NY99 382-99; AF196835) U.S. reference WNV strain with conserved (*) residues and amino acid positions with strong (:) and weak (.) homology indicated according to the Gonnet Pam250 matrix. Domain-specific WNV NS4B protein structure (**Figs. 6-7 and 6-8**) is provided with subcellular localization of the putative transmembrane domains (pTMDs) and loops in either the ER-lumen or cytosol^{168,169}.

Figure 6-2: *In vitro* A) multiplication kinetics and B) temperature-sensitive (*ts*) phenotype of applied isolates in the *in vitro* A549 expression platform

A) *In vitro* multiplication kinetics



B) Temperature-sensitive (*ts*) phenotype at 41°C vs. 37°C

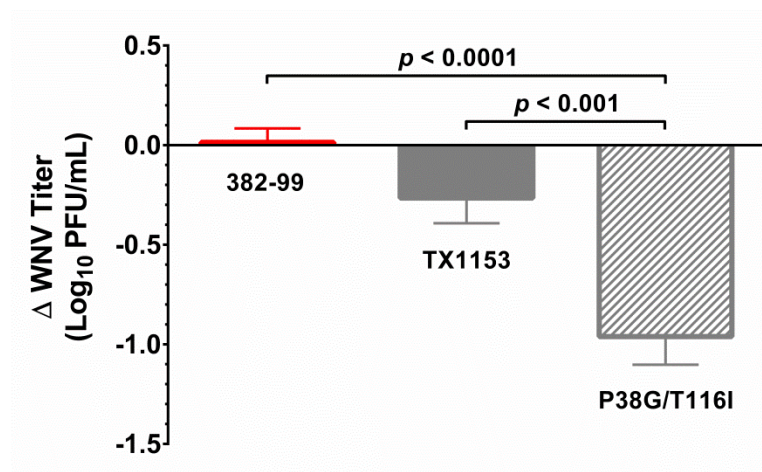


Figure 6-3: Expression profiles of targeted genes (clustered by function) in the *in vitro* human A549 expression platform at both 24 and 36 hpi. Fold up/down-regulation indicated relative to mock infection

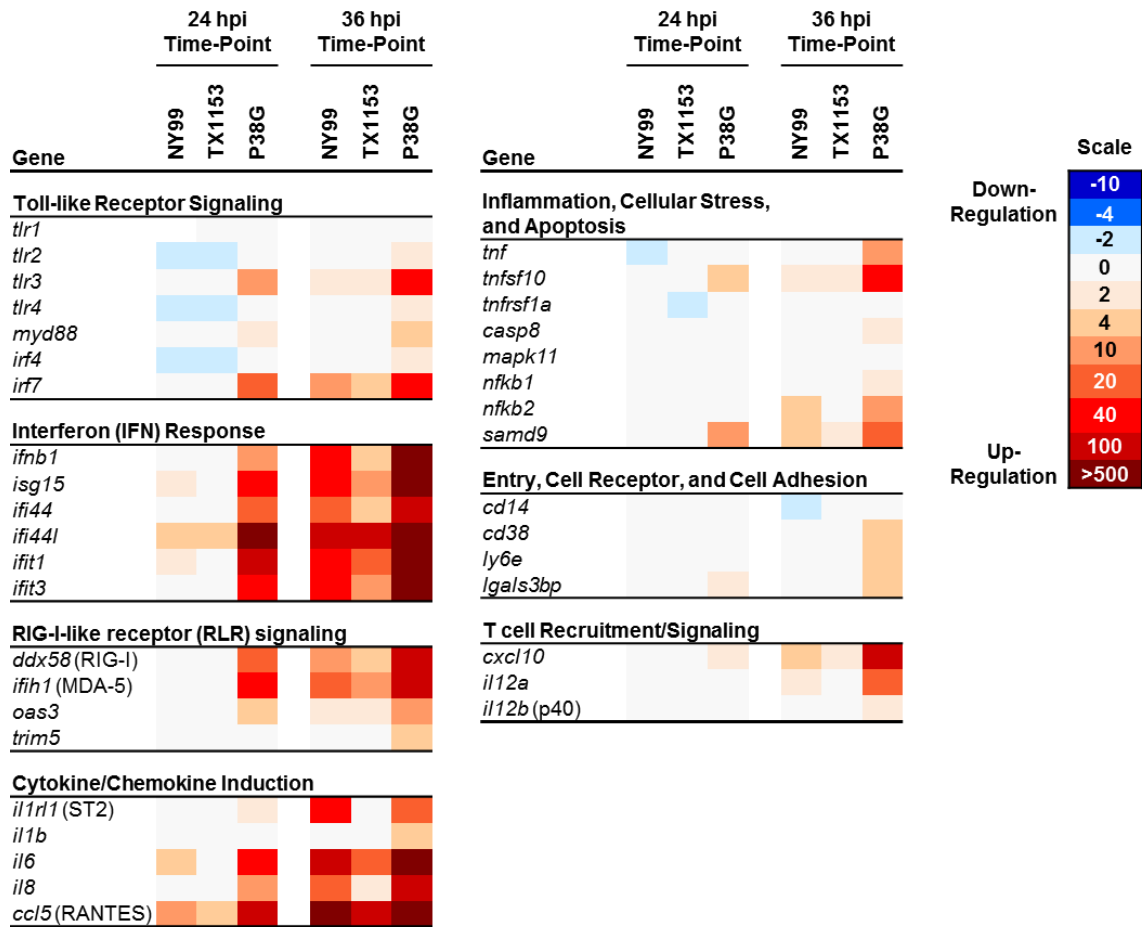


Figure 6-4: Fold-expression of targeted genes at 24 hpi in the PrimePCR™ Flavivirus Infections H96 panel (relative to NY99). Bold red, significant up/down-regulated expression ($p < 0.05$; $\alpha = 0.05$)

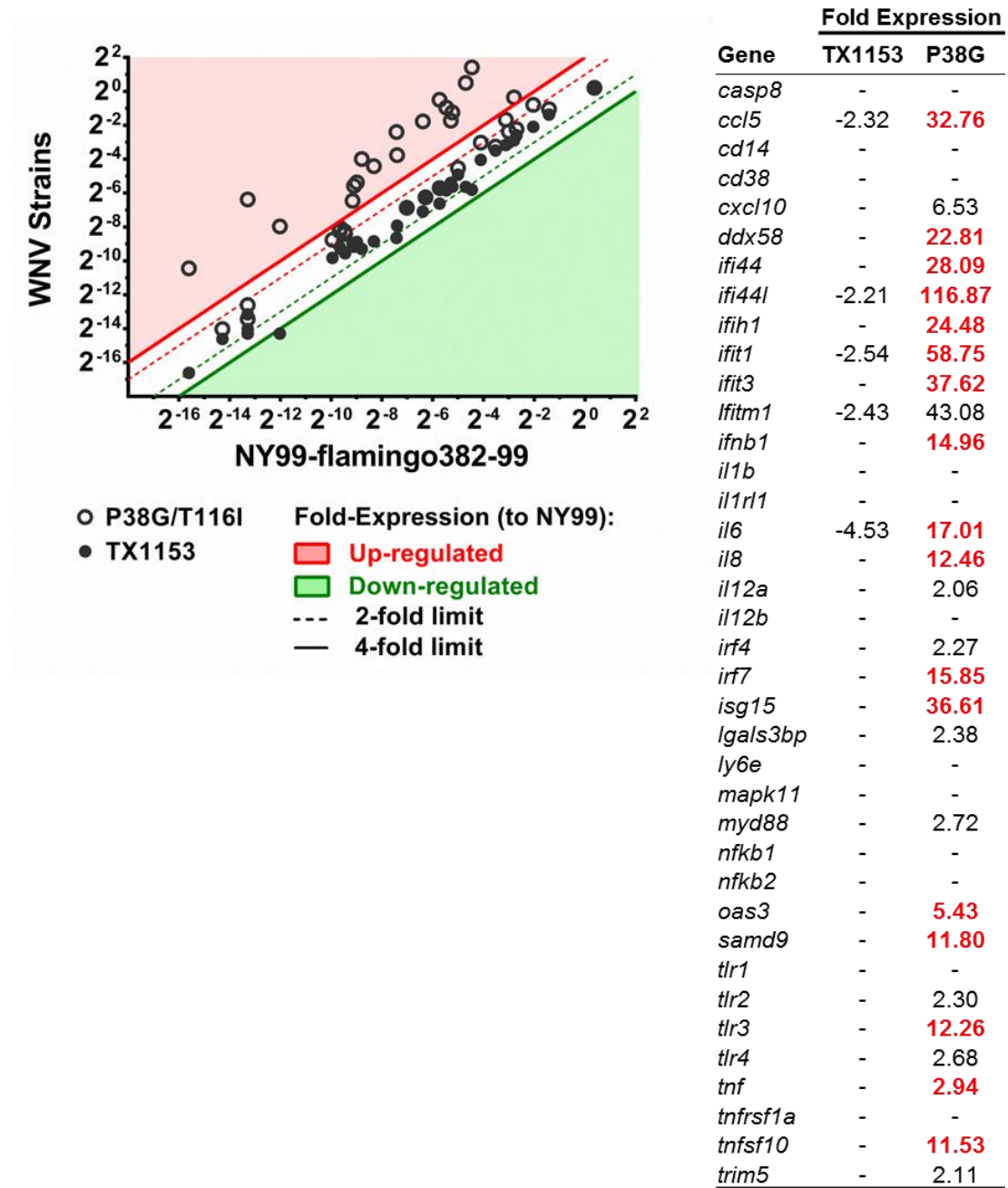


Figure 6-5: Fold-expression of targeted genes at 36 hpi in the PrimePCR™ Flavivirus Infections H96 panel (relative to NY99). Bold red, significant up/down-regulated expression ($p < 0.05$; $\alpha = 0.05$)

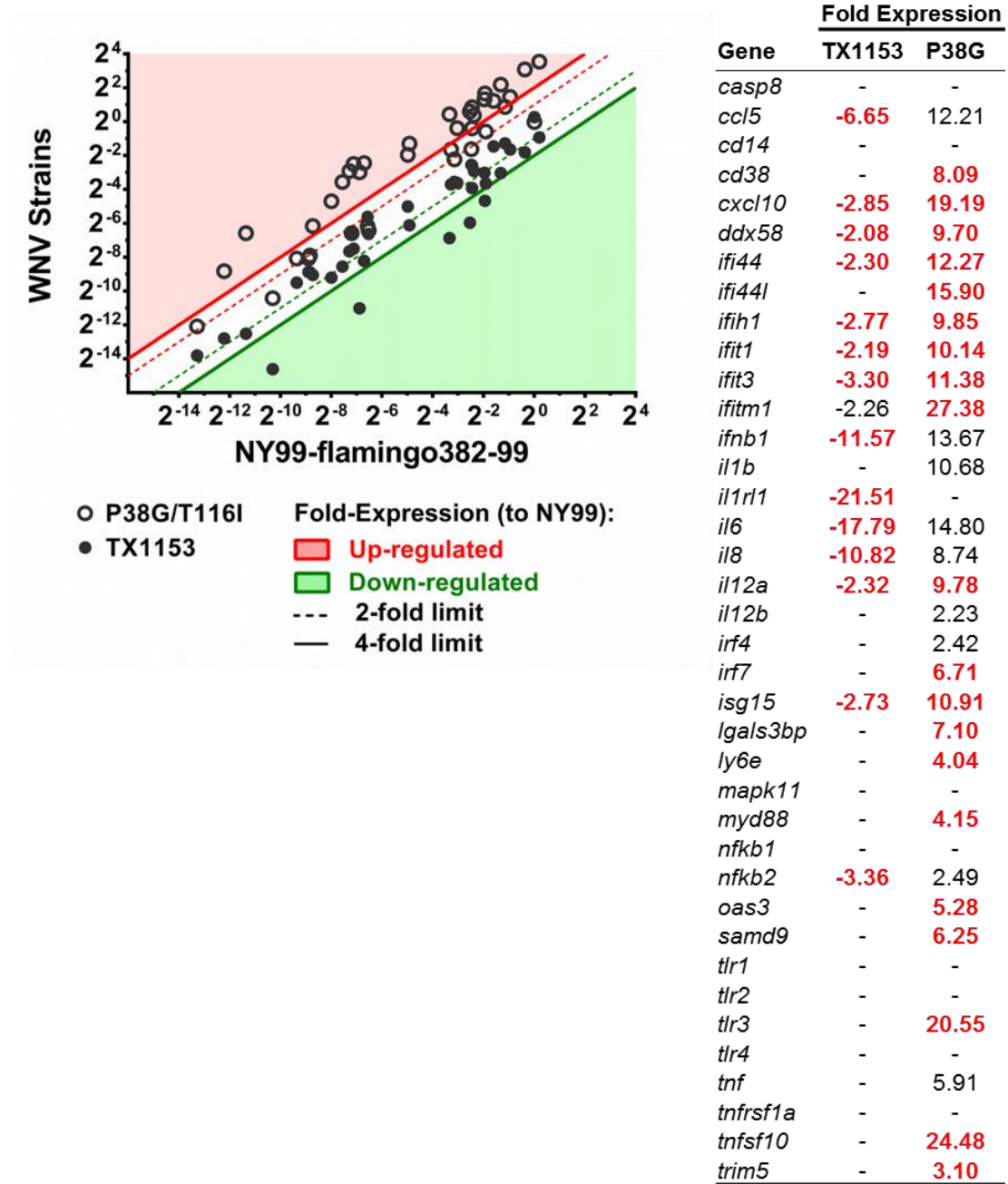


Figure 6-6: Human cytokine/chemokine protein expression in the *in vitro* A549 platform at both 24 and 36 hpi (relative to NY99 infection)

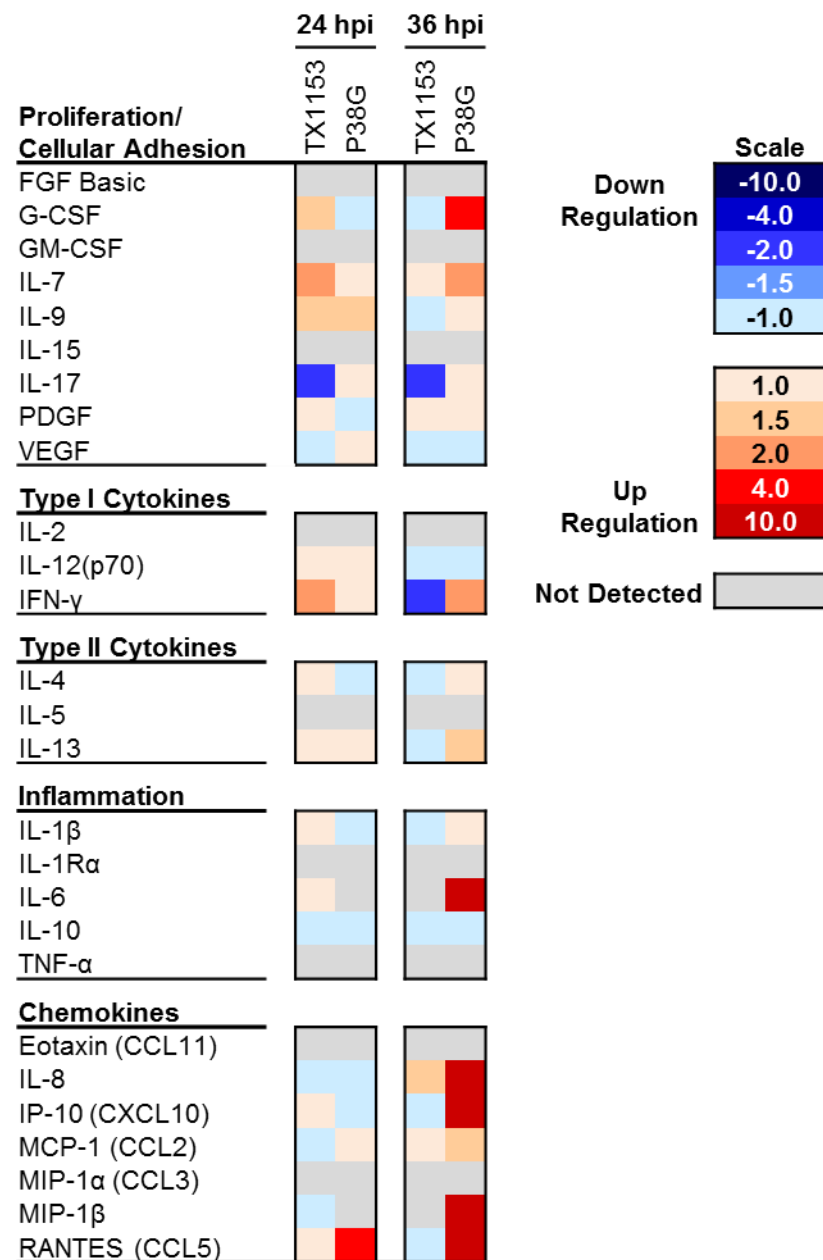


Figure 6-7: Predicted WNV NS4B protein membrane topology *pre-cleavage* based on the DENV-2 NS4B topology. Putative transmembrane domains (pTMD), transmembrane domains (TMD), and loops indicated.

[Modified from: **Miller, S., et al** (2006) *J Biol Chem* and **Zou, J., et al.** (2014) *J Virol*]

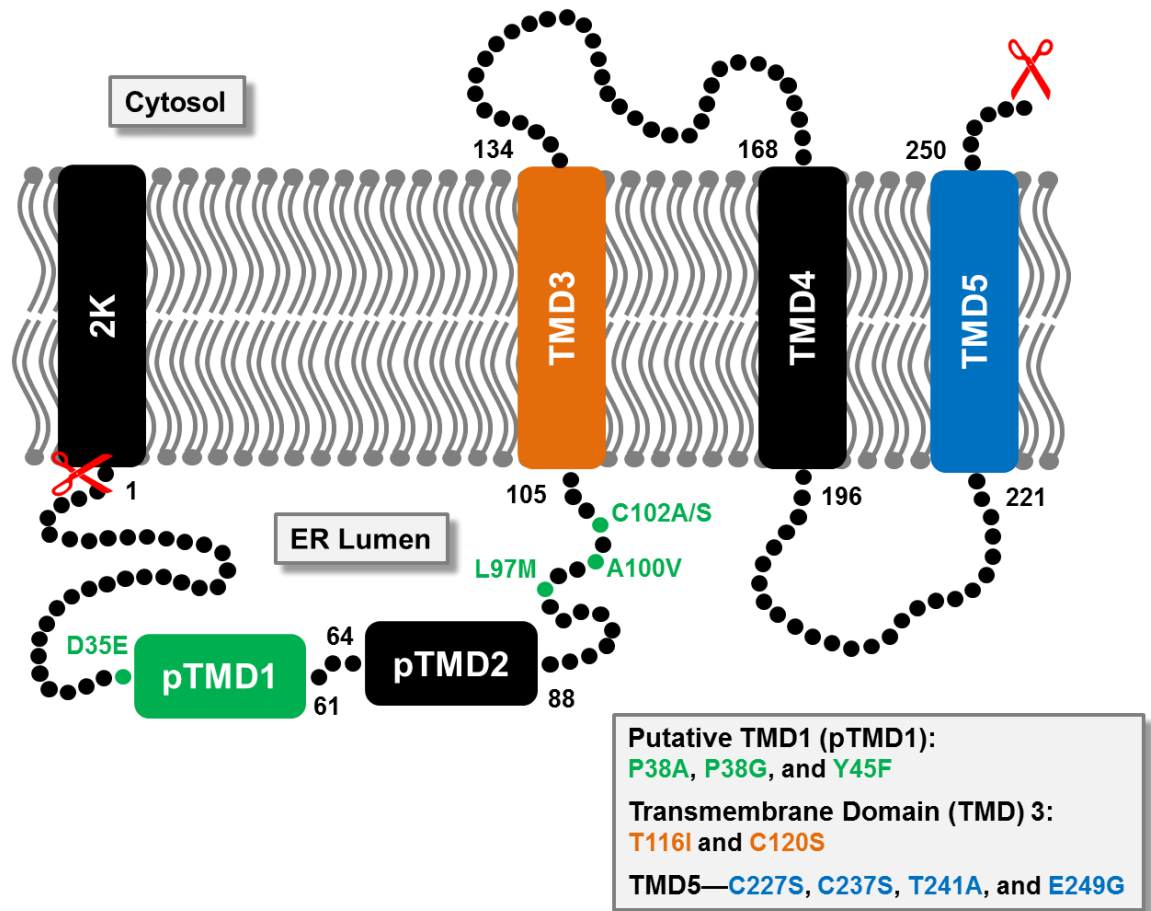


Figure 6-8: Predicted WNV NS4B protein membrane topology *post-cleavage* based on the DENV-2 NS4B topology, related to **Fig. 6-7**.

[Modified from: **Miller, S., et al** (2006) *J Biol Chem* and **Zou, J., et al.** (2014) *J Virol*]

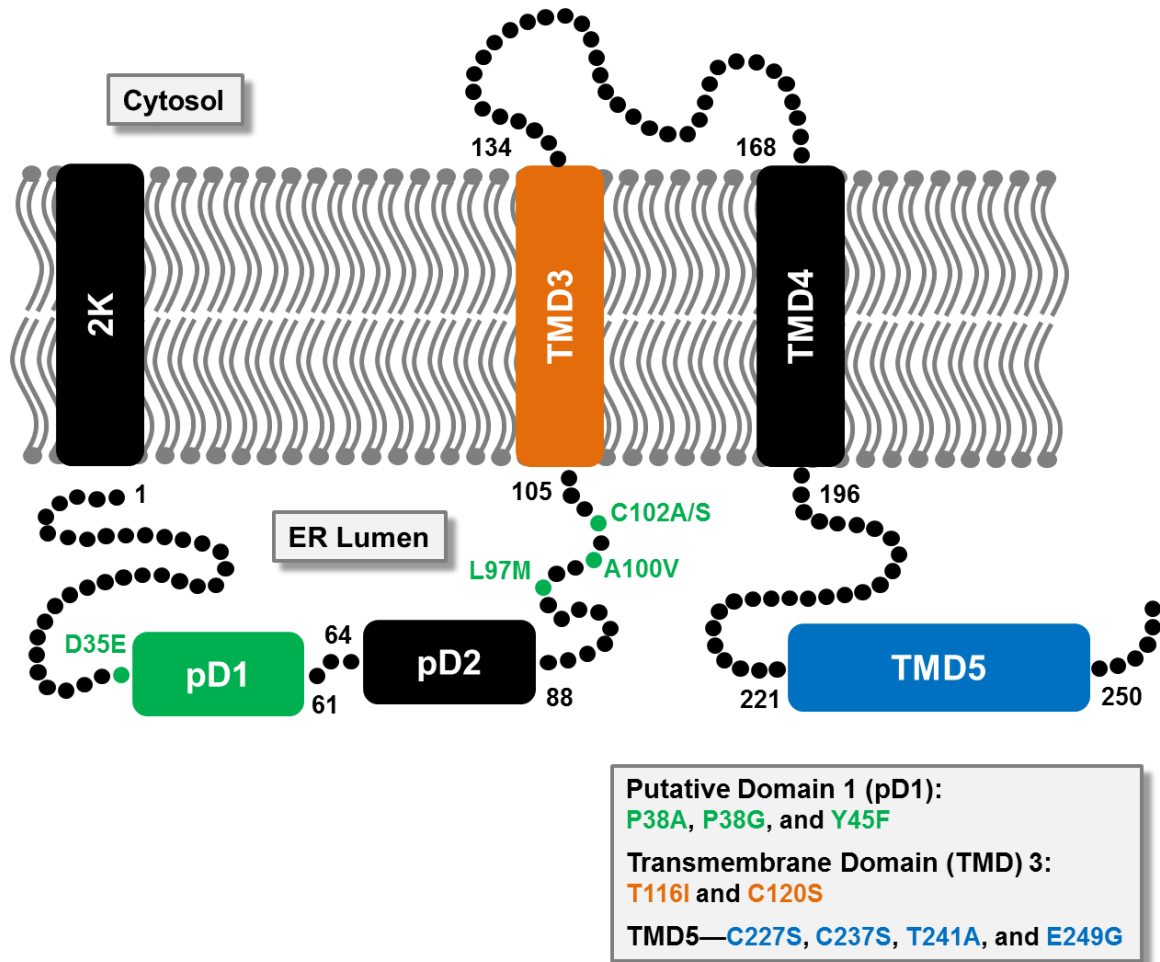


Figure 6-9: Residue- and domain-dependent gene expression (relative to NY99ic) for WNV NS4B mutants in the *in vitro* A549 platform

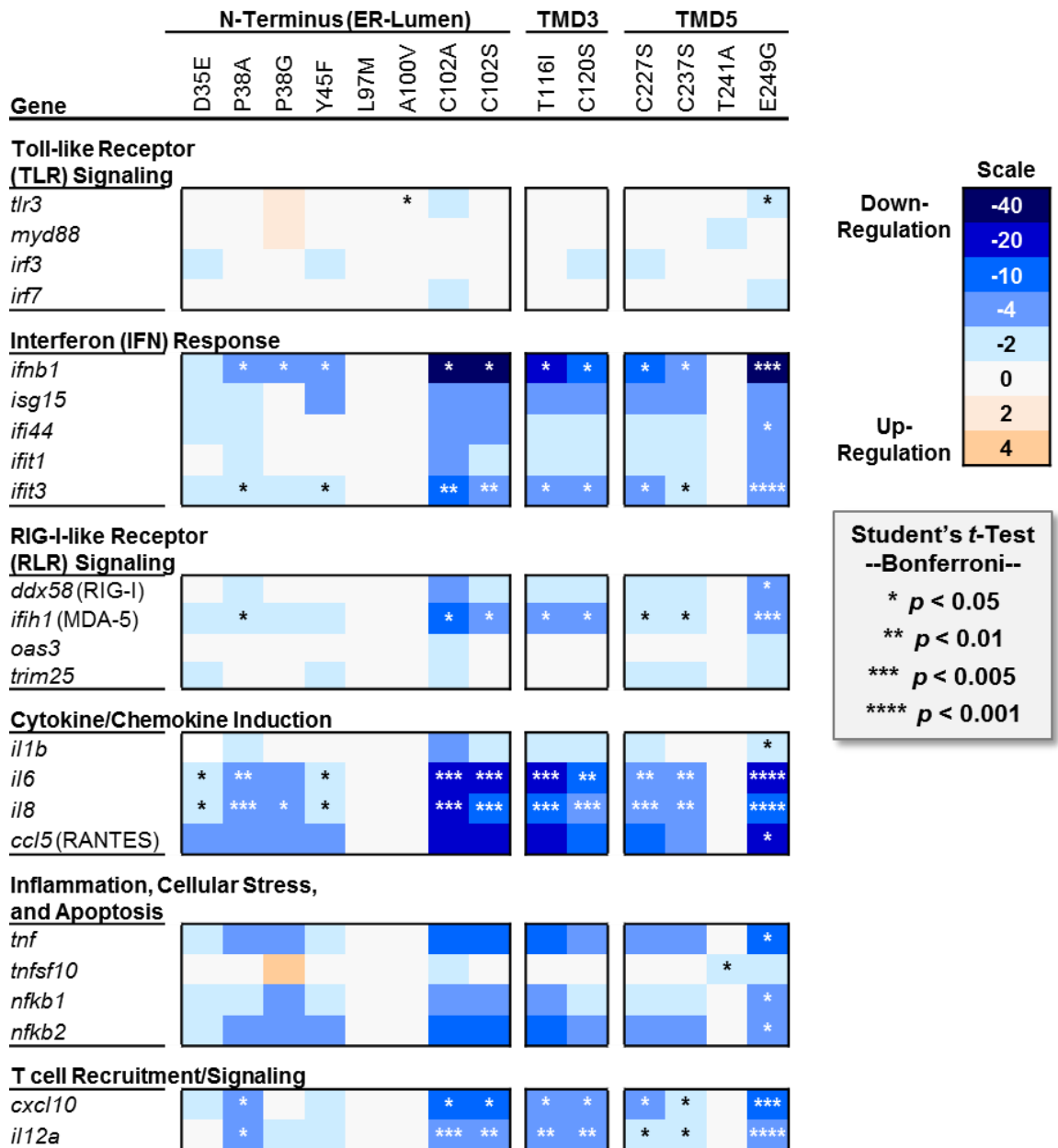


Figure 6-10: Fold-protein induction in NS4B mutant infection (relative to NY99ic)

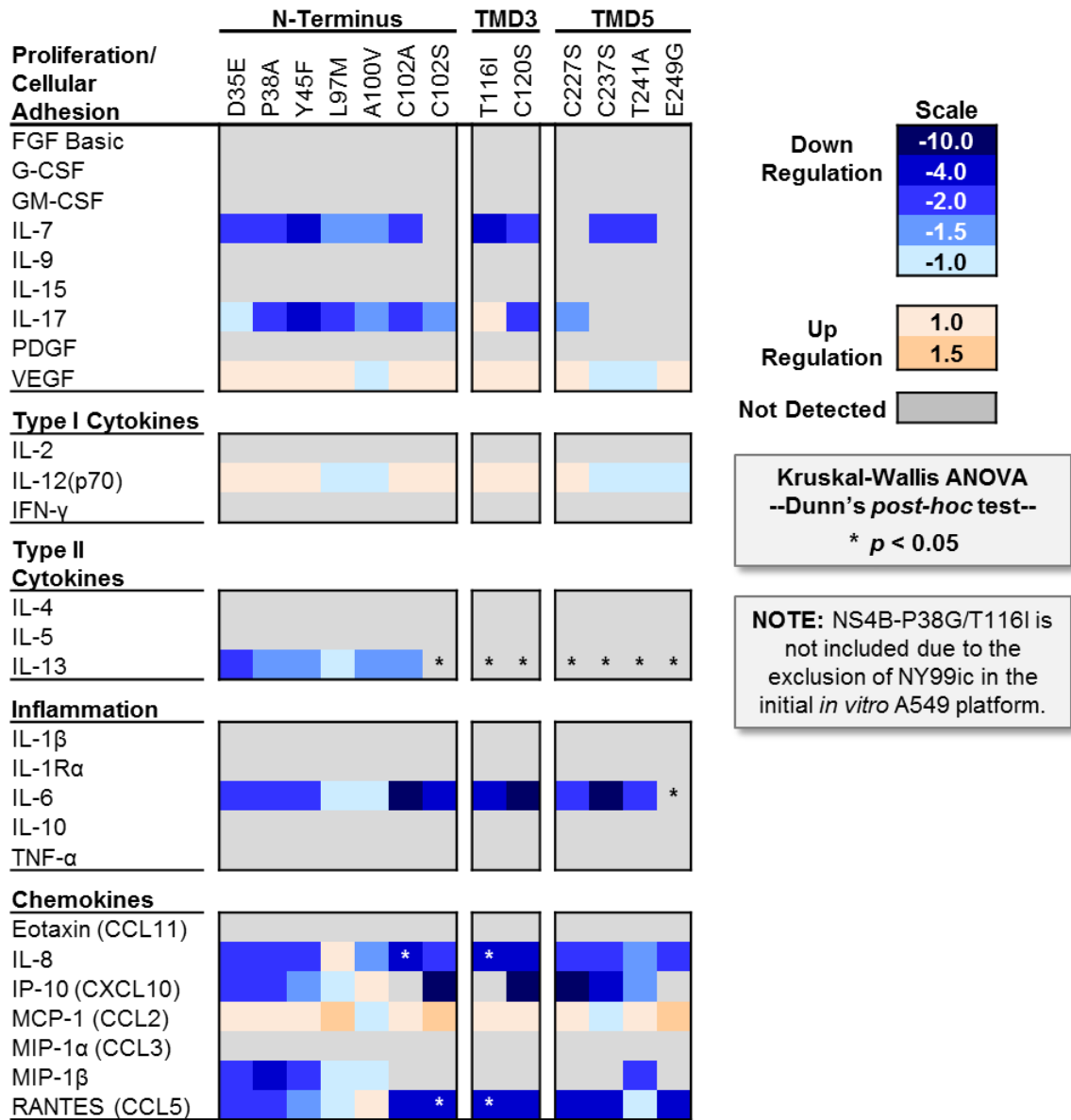


Figure 6-11: Differential gene expression between NS4B mutants: **A)** P38A vs. P38G (attenuated) and **B)** C102A vs. C102S (attenuated). **C)** Fold-expression ≥ 2 -fold is indicated per comparison with significant fold-changes (p -value < 0.05 ; $\alpha = 0.000417$) designated in boldface red.

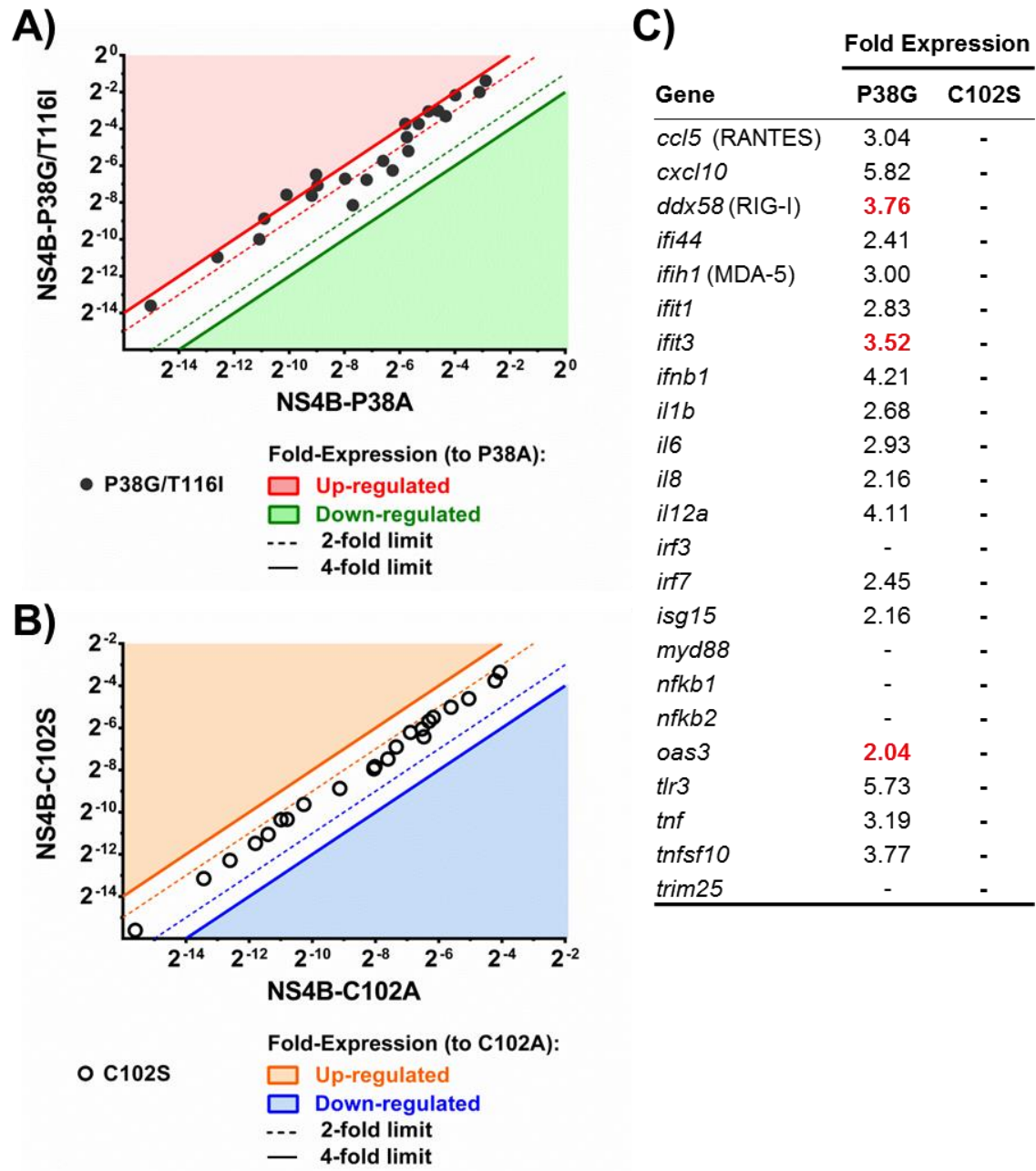
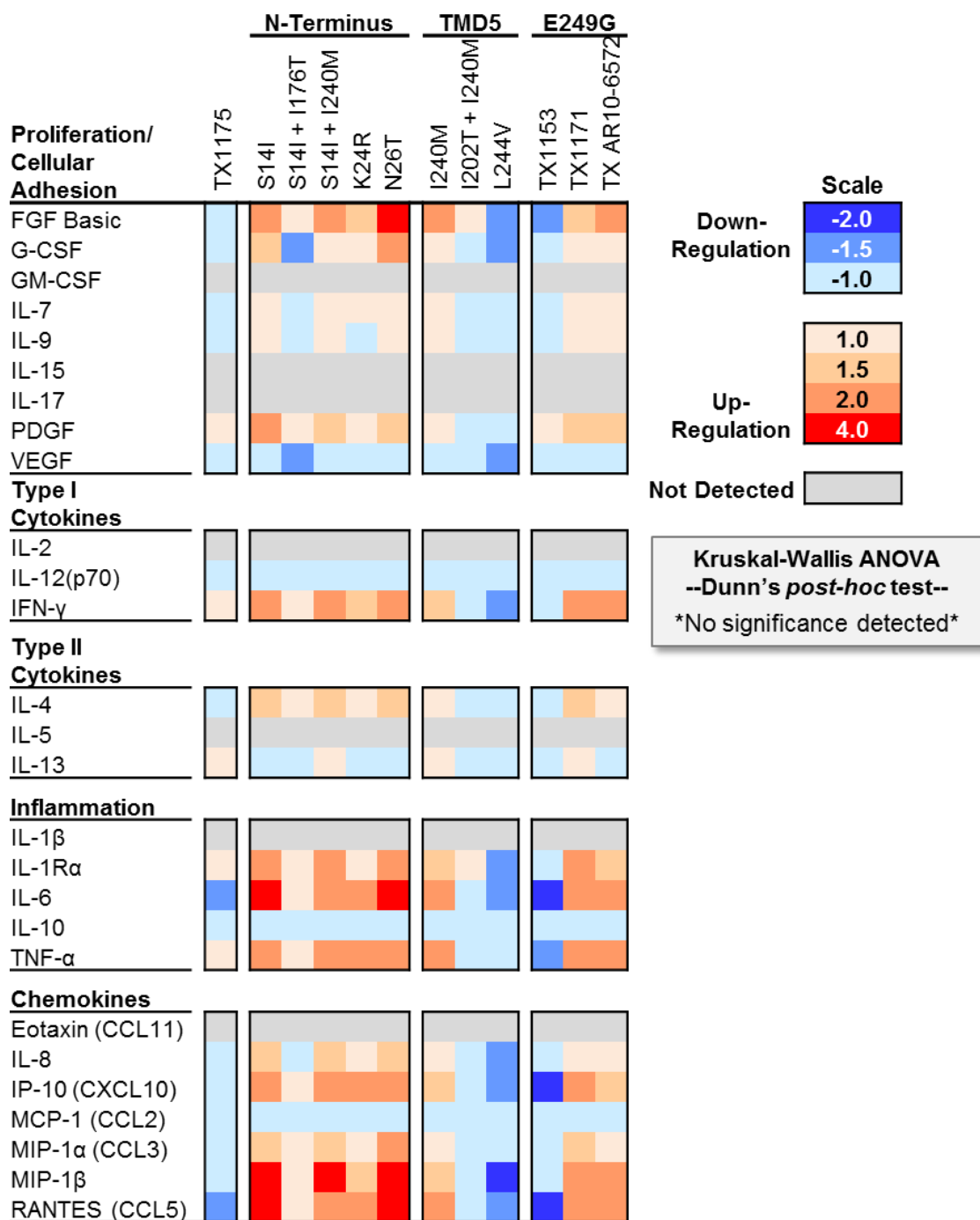


Figure 6-12: Fold-protein induction in natural WNV isolates encoding NS4B substitutions (relative to NY99-flamingo382-99)



CHAPTER 7

Discussion

West Nile virus continues to evolve in the United States. Emergence, extinction, and co-circulation of multiple U.S. genotypes^{42,57-59,63,64,391} coupled with sustained detection in annual surveillance programs and periodic enzootic outbreaks^{48,349} all support the establishment of WNV in North America as a now endemic, public health burden on the U.S. In response to recurring local, regional, and national spikes in clinical human disease, the same questions arise:

- 1) Is this a new strain?
- 2) Is the virus more virulent?
- 3) What is *different* about the virus? (among others)...

Efforts to answer these questions through in-depth ecological, epidemiological, phylogenetic, and phenotypic studies in multiple *in silico*, *in vitro*, and *in vivo* platforms have linked several, conserved substitutions in the WNV genome with altered genotypic distributions, host selection/adaptation, and virulence^{41,57-59,63,66,108,174}. However, integration of these findings into the broader, inter-linked contexts of molecular epidemiology/ecology (environment), quasispecies structure (virus), and viral pathogenesis (host) as intrinsic drivers of WNV evolution has remained challenging.

This dissertation defined critical genetic and molecular determinants involved in the continued evolution of WNV in the U.S. and confirmed the role of the NS4B protein in modulation of host innate antiviral responses. In **Specific aim 1**, *in silico* studies of 2005-2010 El Paso/Ciudad Juarez isolates (**Subaim 1a**) implicated the role of bi-

directional gene-flow on the U.S.-Mexican border as a novel epidemiological stimulus in the continued evolution of WNV in the southwestern U.S. However, limited viral movement in northern Mexico restricts application of these geospatial phylogenetic trends as an *in silico* model of natural, enzootic transmission in this region without additional host/vector surveillance efforts. In this regard, in-depth *in silico* comparisons using 2002-2012 Harris Co., Texas isolates in **Subaim 1b** continued to validate the Harris Co. paradigm as surrogate model for WNV evolution on a national scale; furthermore, application of regional 2010-2012 isolates in this paradigm correlated the introduction of novel genetic and genotypic determinants into Texas with the recent 2012 WNV outbreak in Dallas.

Further application of this paradigm in detailed NGS *in silico* analyses (**Specific aim 2**) identified both host- and genotype-dependent trends in WNV quasispecies structure further linking attenuated neuroinvasive *in vivo* phenotype to low diversity virus populations. Based on these *in silico* trends, directed *in vitro* studies in **Specific aim 3** using the established NY99ic reverse-genetic platform implicated the role of the *N*-terminal and transmembrane domains in the encoded NS4B protein in the dysregulation of host Type-I IFN-mediated innate antiviral responses. Consolidation of these novel *in silico* and *in vitro* results not only addresses the above questions but provides deeper insight into the implications of these underlying molecular mechanisms driving the selection of natural domestic WNV populations throughout its extended tenure in the U.S.—as defined in the following sections.

7.1 HARRIS COUNTY (CO.) PARADIGM

In its transit across the continental U.S., the emergence, extinction, and co-circulation of several, distinct U.S. genotypes has been attributed to unique selection pressures in multiple, disparate WNV-competent hosts—both mosquitoes and birds^{66,116}.

However, not all hosts are equally susceptible with evidence of diverse host-specific competencies^{70,72-76}. Furthermore, distribution of potential hosts varies across the U.S.; in particular, *Culex* spp. mosquito populations provide an excellent example with regional dominance of several species: *Cx. pipiens pipiens* (north)⁴⁵, *Cx. restuans* (north-east)⁸⁶, *Cx. pipiens quinquesfasciatus* (south), and *Cx. tarsalis* (west) implicated in the westward expansion and sustained transmission of WNV throughout the U.S.^{71,88-90}. Birds are no exception—with differing viral fitness between susceptible (stronger) and more resistant (weaker) species contributing distinct selective pressures in the emergence of new viral variants^{44,70,99-101,343}. Furthermore, long-distance migration events in multiple susceptible populations both aids and complicates applied *in silico*, epidemiological, or ecological WNV studies in domestic avian species through the annual selection and subsequent spread of regional WNV populations^{100,392,393}.

In this regard, Texas is unique. Due to geographic convergence of both the Central and Mississippi flyways, Texas serves as a roosting site for >629 resident and migrant wild bird species with >333 (53%) migrant species constituting 98.5% of all known Nearctic-Neotropical migrants in North America (north of Mexico)³⁴⁶. Not all birds are competent reservoir hosts^{44,70,101,102,343}; however, this ecological phenomenon provides an unparalleled, biological context to sample viral gene-flow with the emergence, extinction, and co-circulation of dominant, endemic strains among resident and regional wild bird populations from all over the U.S.; in effect, the Harris Co. paradigm also presents a natural model to test for *in vivo* determinants impacting disparate host-competence and/or viral selective pressures that remains untapped.

From an epidemiological perspective, one question remains: does the Harris Co. paradigm reflect viral gene-flow and evolution for the *entire* U.S.? Not all wild birds in the U.S. are migrant species; furthermore, regional vector-surveillance limits sampling to resident *Culex* species: *Cx. quinquesfasciatus* and *Cx. tarsalis*^{71,88,89}. In effect, the Harris Co. paradigm is biased with excluded *in silico* sampling of adaptive mutations or

dominant strains restricted to regional host populations (outside of Harris Co., Texas) due to various host-dependent selection pressures^{66,67,70,76,86,93,100,101,394}. That being said, the results from **Specific aim 1** (and prior studies) confirmed the significant (p -value < 0.05) *in silico* phylogenetic mixing of Harris Co. isolates with other 1999-2011 U.S. isolates^{58,63,350,351}; however, sampling bias is a known constraint of applied *in silico* viral studies with >78.0% of all included sequences in **Chapter 3-5** phylogenetic comparisons representing surveillance efforts limited to four U.S. states: CT (31.5%), NY (22.0%), TX (16.4%), and IL (8.1%)^{55,57-59,63,68,69,347,348,350,351}. Supplemented surveillance efforts and *in silico* comparisons of novel U.S. isolates from under-represented states in the central, Midwestern, and southeastern U.S. regions would provide critical insight into the proposed impact of the Harris Co. paradigm as a surrogate model for WNV evolution in the U.S as a whole.

From an applied public health perspective, the Harris Co. paradigm offers a unique growing repertoire of natural U.S. isolates with established *in vitro* and *in vivo* phenotypes for applied immunological studies and antiviral/candidate-vaccine development—for which none currently exist. In this regard, current testing models are limited through the use of extinct prototype U.S. strains (*e.g.*, NY99-flamingo382-99 and WN-TX2002) that no longer reflect the distribution of circulating U.S. isolates^{55,58,63,351}. *In silico* clustering of all sequenced isolates from the 2012 WNV outbreak with 2006-2010 U.S. strains—described in **Chapter 4**—reinforces this limitation³⁵¹. *How* additional conserved mutations in current circulating U.S. genotypes impact host innate/adaptive antiviral responses or clinical outcome (compared to prototype strains) as a predictive determinant of annual WNV incidence is still unclear. In turn, the results from **Subaim 1b** do not offer conclusive *in silico* or *in vivo* evidence linking *in situ* genotypic selection the emergence a more “virulent” epidemic strain^{351,395}. Further research is needed to link the multi-factorial roles of the numerous biological events—including (but not limited to) viral mutation, virus-host interactions, integrated vector management, and ecological

mental conditions—which culminate in clinical disease and epidemic scenarios. Future applications of the Harris Co. paradigm will be instrumental in this regard as a well-established surveillance model providing a growing repertoire of novel U.S. strains for applied *in silico*, *in vitro*, and *in vivo* studies to elucidate the elusive mechanisms driving WNV evolution and clinical disease in the western hemisphere.

7.2 DETERMINANTS OF WNV POPULATION STRUCTURE

Initial pursuits into viral evolution, adaptation, and selection applied traditional virological techniques that distilled viruses to a single, static consensus—convenient albeit prone to mutation. In fact, viruses exist as a population, or “cloud” of individual virus RNAs (termed a viral quasispecies), encoding both advantageous and detrimental mutations cooperating to elicit a particular *in vitro* or *in vivo* viral phenotype^{114,369,396,397}. Next generation sequencing (NGS) provides a novel tool to delve below the level of consensus sequencing to record a “deep” census of quasispecies structure resulting from host-selection pressures and/or natural genetic drift^{331,355}. Based on these principles, dual application of the proposed Harris Co. paradigm³⁵¹ in the Illumina HiSeq1000 NGS platform (**Chapter 5**) exploited an established model to define both host- and genotype-dependent determinants driving intrinsic RNA population structure, SNV selection, and *in vivo* phenotype in the natural evolution of WNV in the U.S. However, it is important to note that it is impossible to link detected mutations between assembled reads or differentiate whether identified variants are equally distributed in the population or limited to a minor subset (*i.e.*, one or a few viruses).

Furthermore, in this departure from traditional, consensus-based approaches, NGS-derived methodologies struggle with one intrinsic, double-edged complication: data management. In practical terms, *how* does one enumerate and/or compare sequence depths climbing into the millions of bases per nucleotide position? Furthermore, *how* can

detected trends be correlated to differences in virus population structure? Implementation of a statistical analysis, a modified Shannon entropy index, provides a population-based metric to enumerate and compare the genetic flux (i.e. diversity) of both dominant and “rare” nucleotides: A, C, G, U, and gaps (deletions) between applied isolates at a specified nucleotide position^{330,398}. Limitations in this approach include the need for uniform depth-of-coverage in all applied NGS datasets resulting in the potential loss of critical sequence variants despite optimized efforts in random down-sampling³⁹⁸. Integration of SNV detection algorithms provides a more straight-forward approach with identification of significant (p -value < 0.05) alternate alleles present in <50% (down to ≤1%) of the total viral population at a particular genomic position^{327,328,398}; however, SNV detection can be biased based on the applied algorithm—incorporation of multiple models limits this risk³⁹⁸.

Based on these principles, all defined genotype- and host-dependent trends in **Specific aim 2 (Chapter 5)** reflect a natural cross-section of WNV population depth in applied Harris Co. isolates but offer no direct insight on the identified impact of differential WNV quasispecies structure *in vivo*. In effect, follow-up *in vivo* phenotyping of all applied isolates in the Harris Co. paradigm defined the novel link between *low* genetic diversity and *in vivo* attenuation in *natural* WNV populations (**Fig. 5-5**); however, this link raises further questions: is *low* genetic diversity a result of purifying selection stemming from attenuating, dominant mutations, *vice -versa*, or alternative, cooperative mechanisms? From this perspective, directed *in vitro* studies to test this question are restricted in that it is theoretically impossible to produce viral samples without intrinsic quasispecies structure. Implementation of serial *in vitro* plaque selection or high-fidelity replication would limit viral diversity; however, these techniques do not eliminate quasispecies structure and the resulting viral populations no longer reflect a natural viral system. In this regard, application of the Harris Co. paradigm provides a

defined repertoire of *natural* isolates with defined quasispecies structure—despite potential selection through limited passaging—for further *in vitro* and *in vivo* testing

From a public health perspective, application of both live-attenuated and chimeric vaccine-candidates^{155,399-402}—*e.g.*, NS4B-P38G/T116I + NS3-N480H mutant (**Specific aim 3**)^{178,295}—offers an alternative platform to elucidate the link between inherent viral diversity (*low* or *high*) and the instability or reversion of identified, attenuating mutations driving the *in vivo* phenotype in these platforms. Furthermore, these studies are not limited to WNV but could offer additional, novel insight into the successes and failures of other flaviviral live vaccine platforms^{356,403-406}.

In addition to these aforementioned applications, directed inquiries into *how* quasispecies structure is maintained, purified, or selected during host pathogenesis would offer critical insight into the role of intra-host diversity and selection in the modulation of host innate/adaptive immune responses and the development of clinical signs, symptoms, or disease. From this standpoint, *how* are the individual WNV populations that colonize the peripheral organs (*e.g.*, spleen and kidney) or CNS “different”, and (if so) *how* are these *differences* linked to the development of neuroinvasive, neurovirulent, or persistent WNV infection? Prior studies have established several disease models in *ex vivo* cell-lines, birds, hamsters, and mice that—in concert with described Harris Co. isolates—provide ideal, applied platforms for the future pursuit of these questions.

7.3 MOLECULAR DETERMINANTS OF INNATE IMMUNITY AND ATTENUATION

In silico studies serve an instrumental role in the detection of both conserved and variable molecular elements as an initial benchmark in the conception and rational design of antiviral therapeutics and live-attenuated viral vaccine candidates. Understanding the underlying molecular mechanisms driving the natural selection of identified *in silico* targets and virus-host interactions contributed by each viral gene/protein to observed *in*

vitro and *in vivo* phenotype will provide critical insight in dissecting *how* and *why* WNV causes disease.

Prior applications of the Harris Co. paradigm identified several natural, conserved molecular determinants implicated in altered host-competence or attenuated viral phenotype^{57,58,63,112,174,351,391}; furthermore, additional *in silico* comparisons detected pan-flavivirus conserved and *in vitro* passage-selected residues in the encoded NS4B protein^{179,180}. Based on these studies, directed application of the NY99ic reverse-genetics system to the IFN-competent human A549 cell culture model (**Chapter 6**) provided a novel *in vitro* platform to interrogate the role of specific, engineered molecular determinants in the NS4B protein in innate antiviral expression. However, the human alveolar A549 cell-line does not mimic a *natural* model of WNV infection; despite this limitation, published application of this *in vitro* platform has provided an established proof-of-concept defining critical immune benchmarks prior to directed *ex vivo* and *in vivo* studies delineating the role of Type-I IFN- α/β antagonism in WNV-induced innate immune phenotypes and *in vivo* attenuation^{198,267,372,373,407}.

From this perspective, *in vitro* application of the NY99ic reverse genetics system confirmed the immunological role of the NS4B protein with significant (p -value < 0.05) dysregulation of IFN- β expression and pro-inflammatory immune cascades through innate antiviral mechanisms both consistent and inconsistent with established *in vitro* and *in vivo* WNV and flaviviral phenotypes^{175,177,178,267,280,295,373,382,408}. Dissection of these results reinforces the potential impact of interpretation bias: *perspective matters*. *In vitro* innate antiviral responses either agreed (wild-type NY99) or disagreed (NY99ic) with published trends depending on the applied positive control (as indicated). Furthermore, these *in vitro* trends (**Specific aim 3**) highlight the multi-genic dysregulation of innate antiviral responses despite the critical role of the NS4B protein in WNV infection.

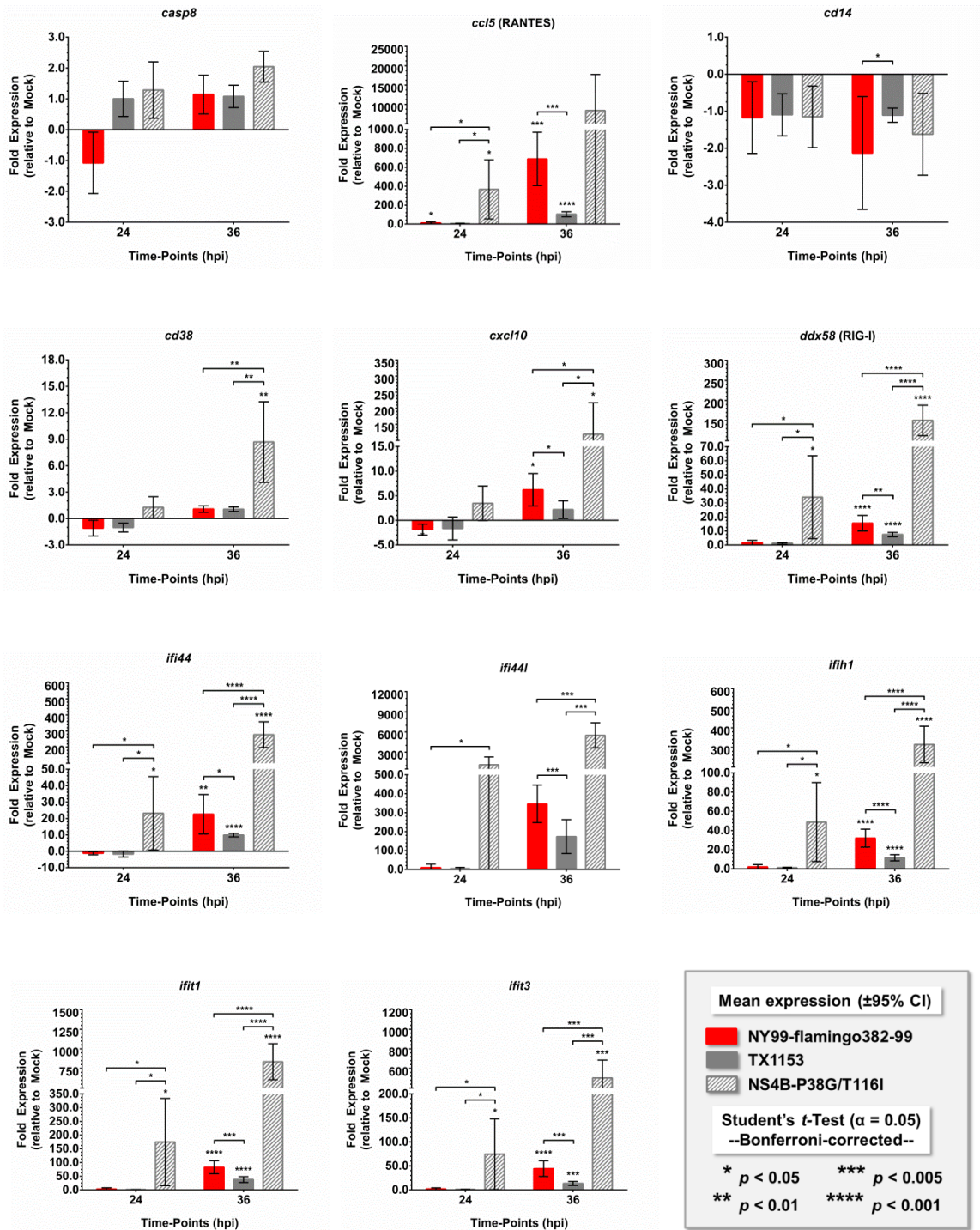
In particular, these identified inconsistencies and exceptions prompt several points for consideration. **First**, are the identified antiviral responses substitution- or position-

dependent—*e.g.*, E249G versus E249X (where X is *any* encoded amino acid)? **Second**, intrinsic differences in innate antiviral responses in the cell-types used for the applied *in vitro* comparisons could polarize observed NS4B-mediated results. In this regard, future application of the NY99ic-derived NS4B *in vitro* model in immune-relevant *in vitro* and *ex vivo* cell-lines: macrophages^{383,386,409-411}, dendritic cells^{178,276,288,292,295,409}, and T lymphocytes⁴¹²⁻⁴¹⁶ (among others) will validate the novel, identified role of the NS4B protein in modulation of the host innate antiviral response. **Third**, what is/are the selective advantage(s) driving NS4B-mediated dysregulation of innate antiviral responses? For example, both NY99ic and the E249G mutant exhibit a virulent *in vivo* neuroinvasive phenotype^{112,174}; however, the E249G mutant exhibited significant (*p*-value < 0.05) comparative antagonism of the host immune response. From this perspective, extension of the NY99ic reverse-genetics *in vitro* platform to additional flaviviral genes/proteins linked to Type-I IFN- α/β antagonism—*e.g.*, NS2A and NS5 proteins (in particular)—would interrogate the potential cooperative, redundant, or independent roles of specific virus-encoded molecular determinants in WNV-mediated innate immune expression. Paired application of a less sensitive antiviral model (*e.g.*, C57BL/6 inbred mouse model⁴¹⁷) may better define the differential link between *in vivo* virulence and innate antiviral responses. In addition, based on the results from **Chapter 5**, intrinsic differences in natural or *in vitro* quasispecies structure offer a novel viral determinant impacting host innate/adaptive antiviral responses although the mechanism driving virulence in diverse virus populations is yet to be defined.

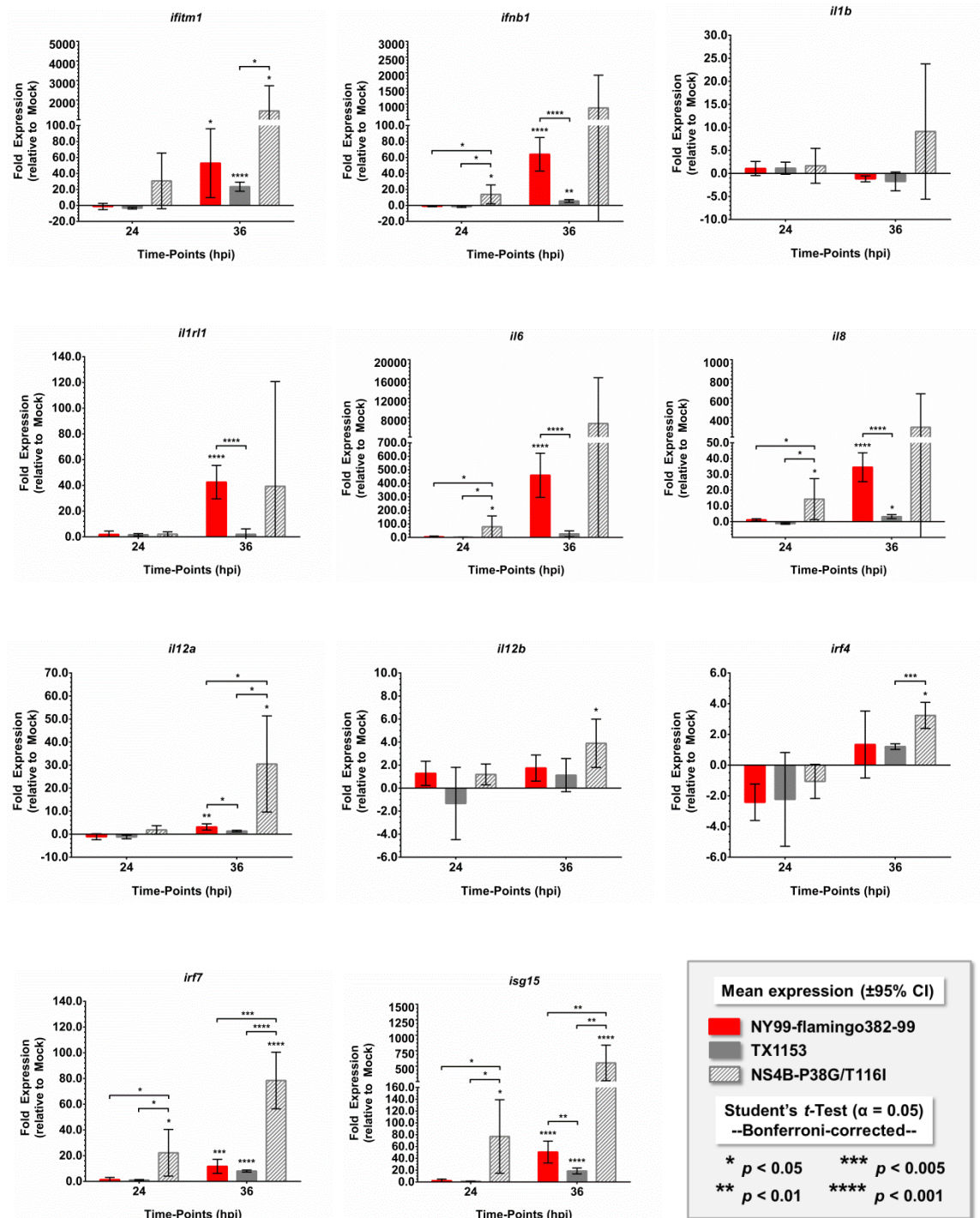
7.4 CONCLUSIONS

West Nile virus is still evolving in the U.S. and will continue to do so. Outbreaks have occurred in the past and will recur in the future. In this regard, the Harris Co. paradigm will provide an established *in silico* platform to elucidate underlying genetic determinants driving the continued emergence of novel U.S. genotypes and strains. Furthermore, future applications of NGS technologies, reverse-genetics tools, and *in vitro* expression platforms will link specific molecular determinants to innate/adaptive antiviral responses and *in vivo* viral phenotypes in the development of novel antiviral therapeutics and live-attenuated vaccine-candidates. *How* natural quasispecies structure fits into this bigger picture of viral ecology, adaptation/selection, and virus-host interactions is an additional exciting and novel puzzle with few linked pieces that's open for exploration.

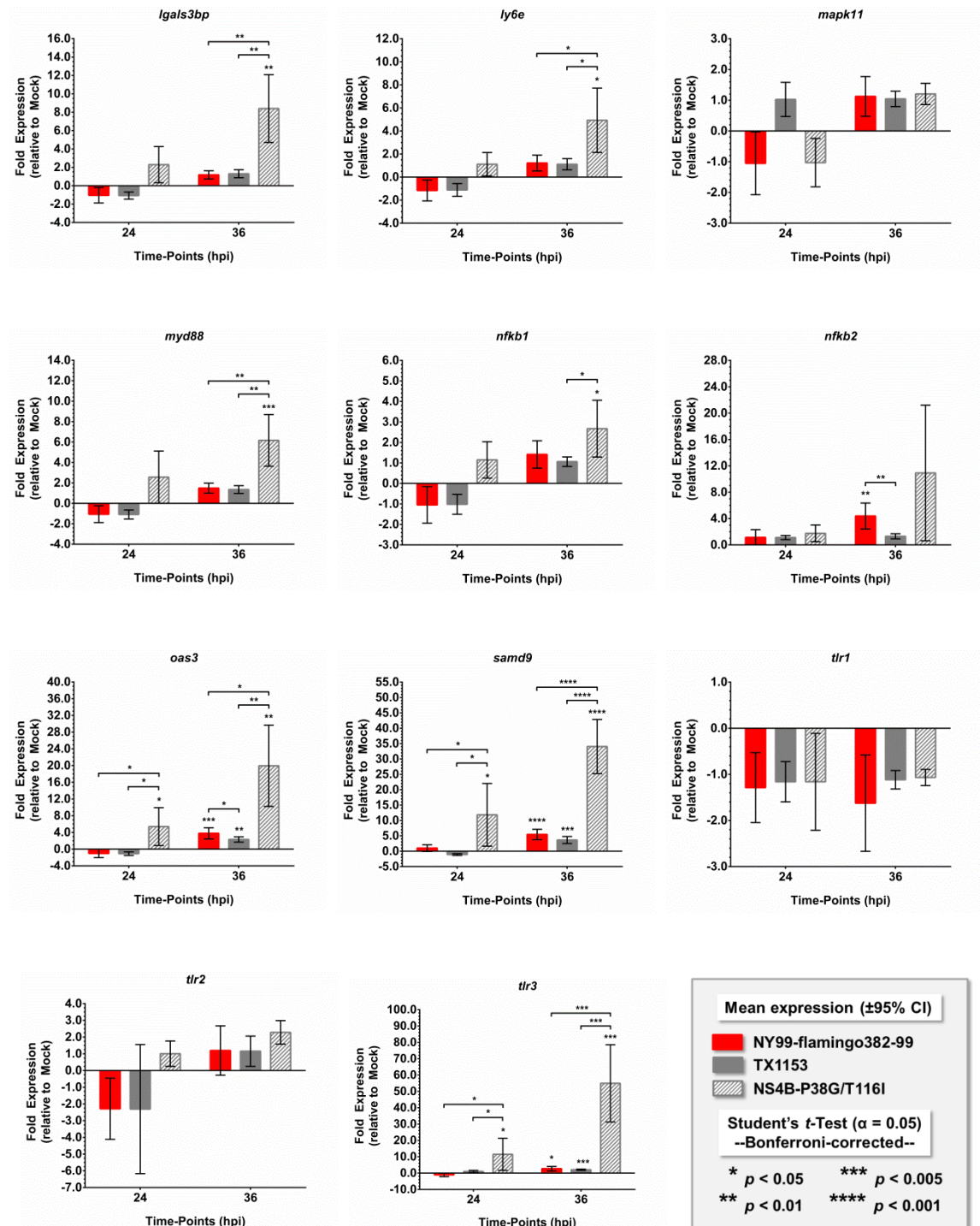
Appendix A: PrimePCR™ Flavivirus Infections H96 Panel



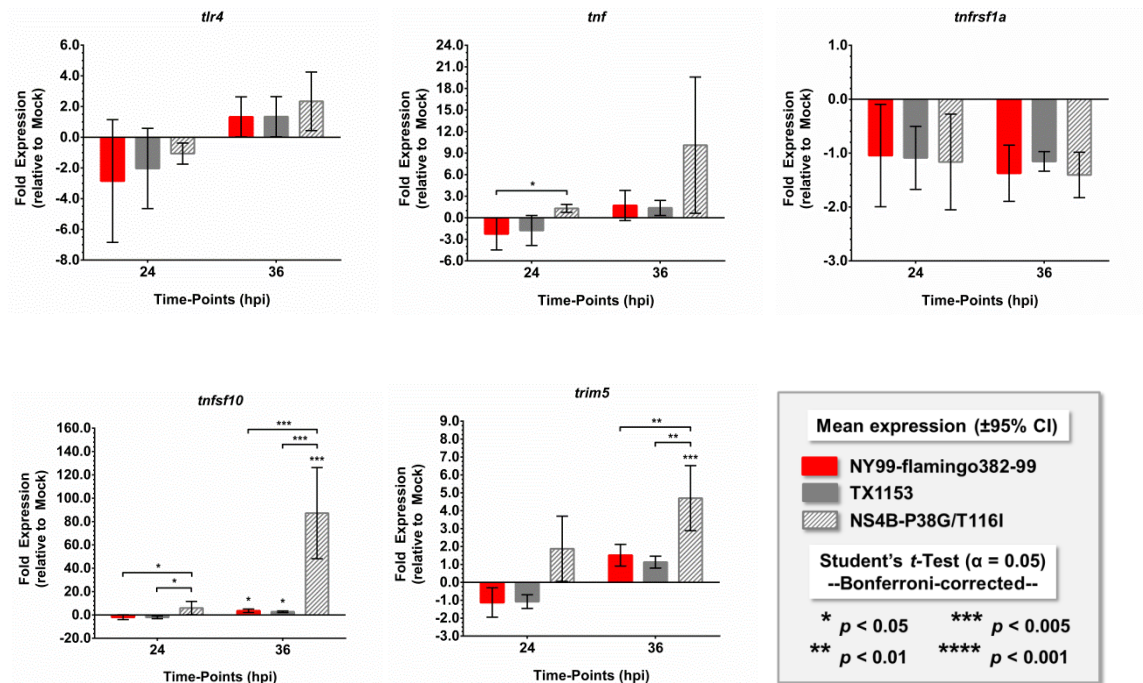
Appendix A: PrimePCR™ Flavivirus Infections H96 Panel (Continued)



Appendix A: PrimePCR™ Flavivirus Infections H96 Panel (Continued)

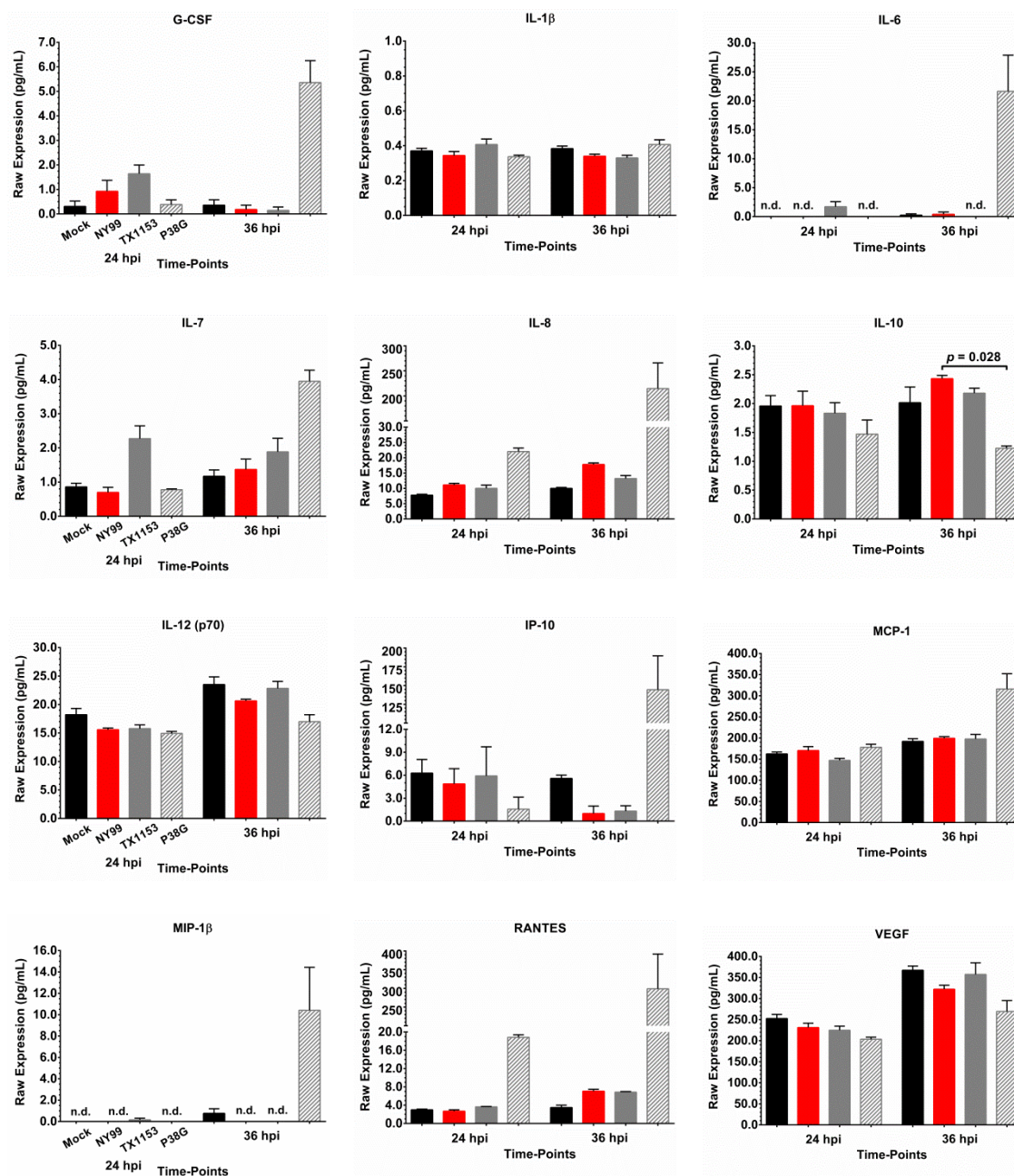


Appendix A: PrimePCR™ Flavivirus Infections H96 Panel (Continued)



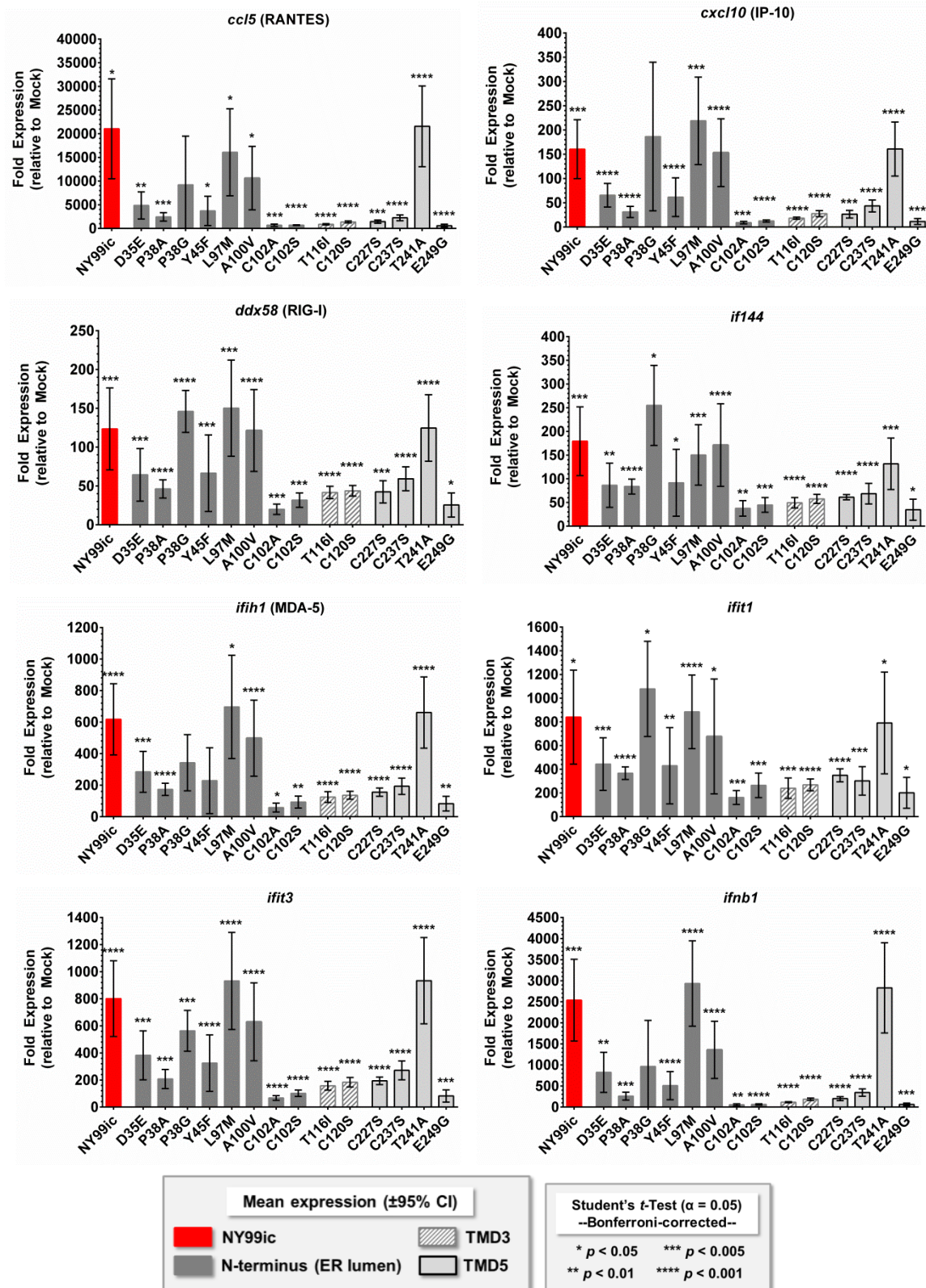
Evaluated target genes *cd86* and *nos2* in the BioRad® PrimePCR™ Flavivirus Infections H96 panel exhibited fold-expression below the limit-of-detection (mean C_T values ≥ 35 cycles) at both the 24 and 36 hpi time-points and are thus not included in the above appendix.

Appendix B: Protein Expression in the *in vitro* A549 Platform

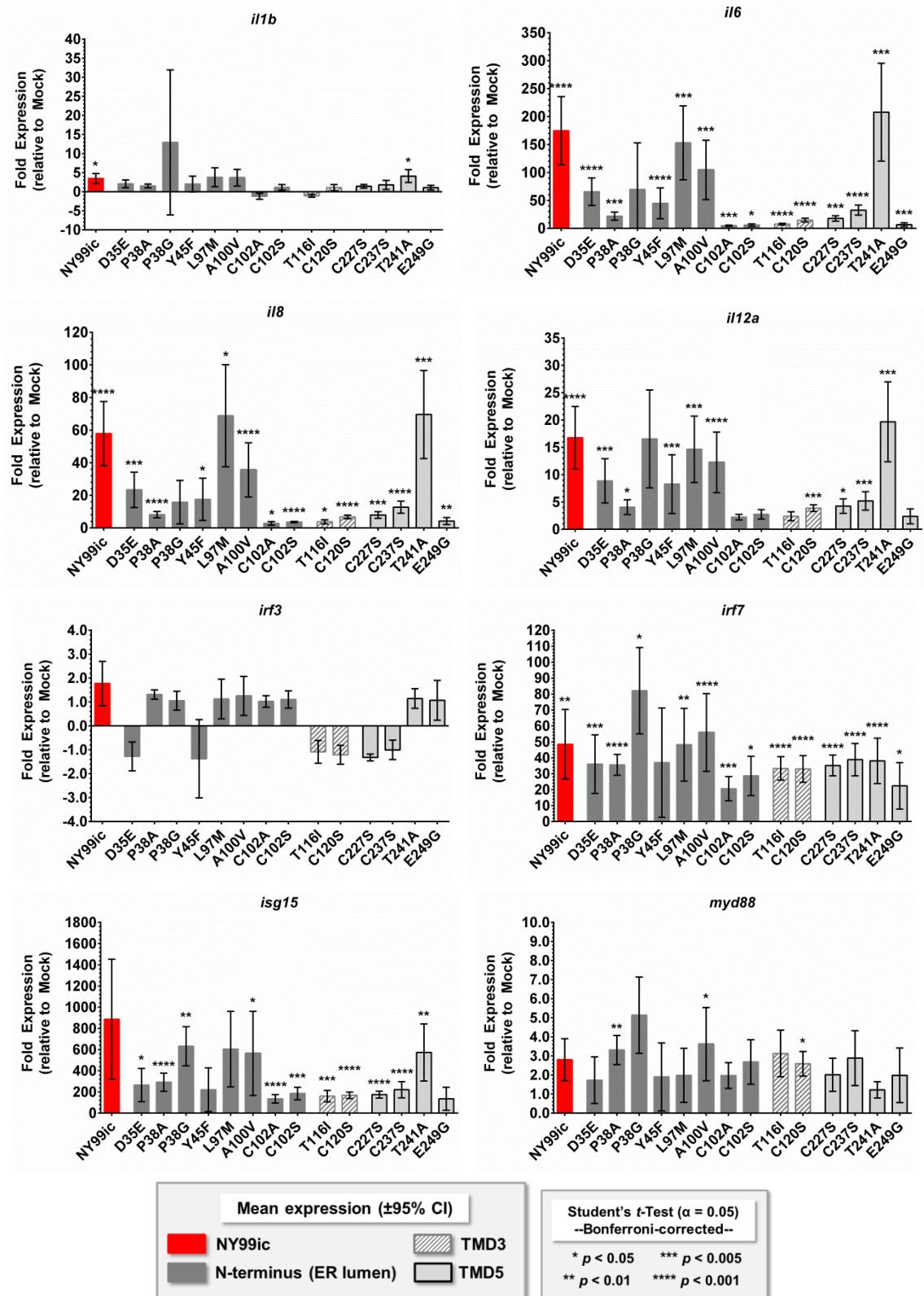


Raw expression (pg/mL) for 16 of the 27 target proteins in the BioRad® Bio-Plex Pro™ 27-plex Group I Cytokine panel was below the limit-of-detection (data not shown). Results are representative of two panels.

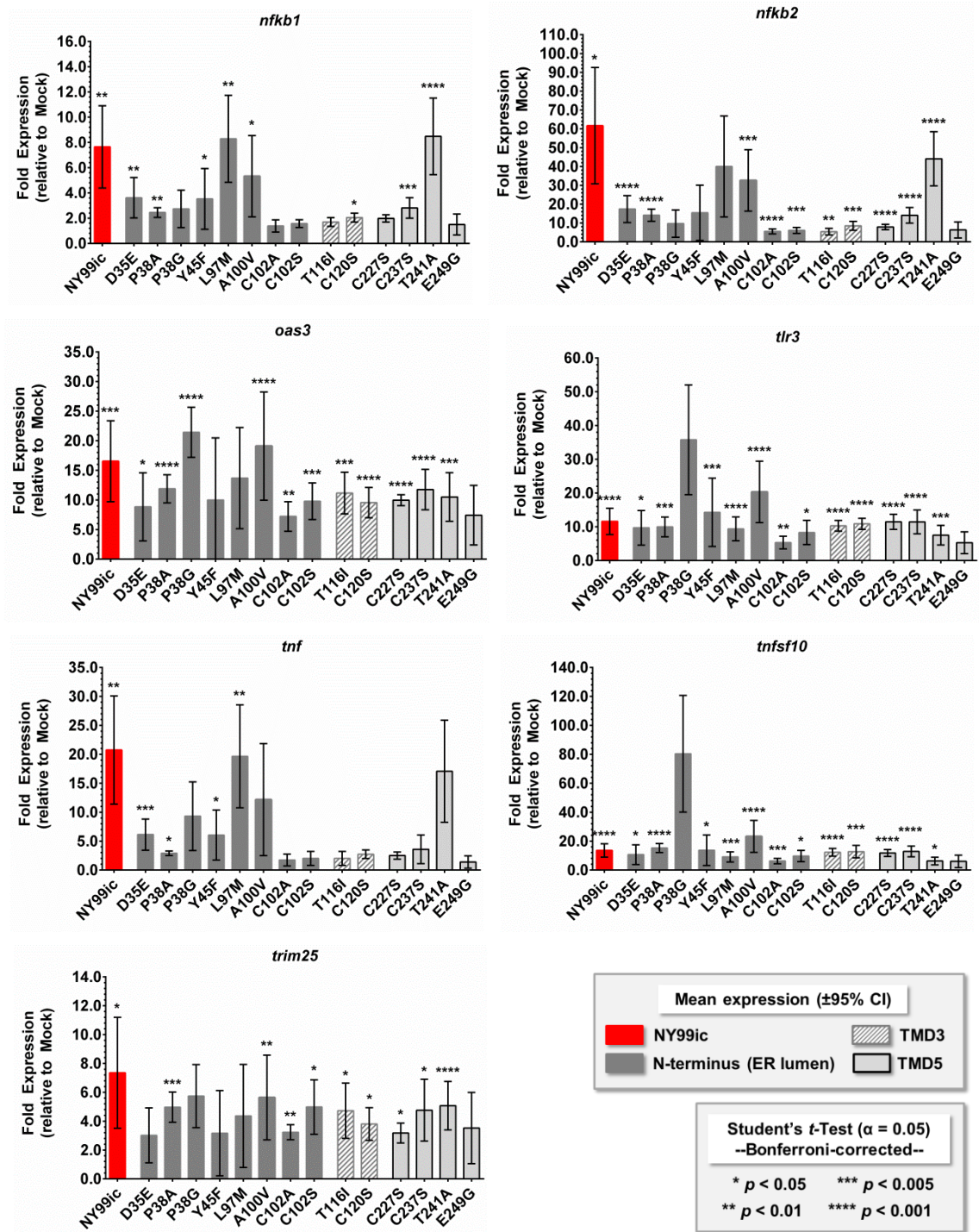
Appendix C: NY99ic NS4B Mutant Gene Expression



Appendix C: NY99ic NS4B Mutant Gene Expression (Continued)

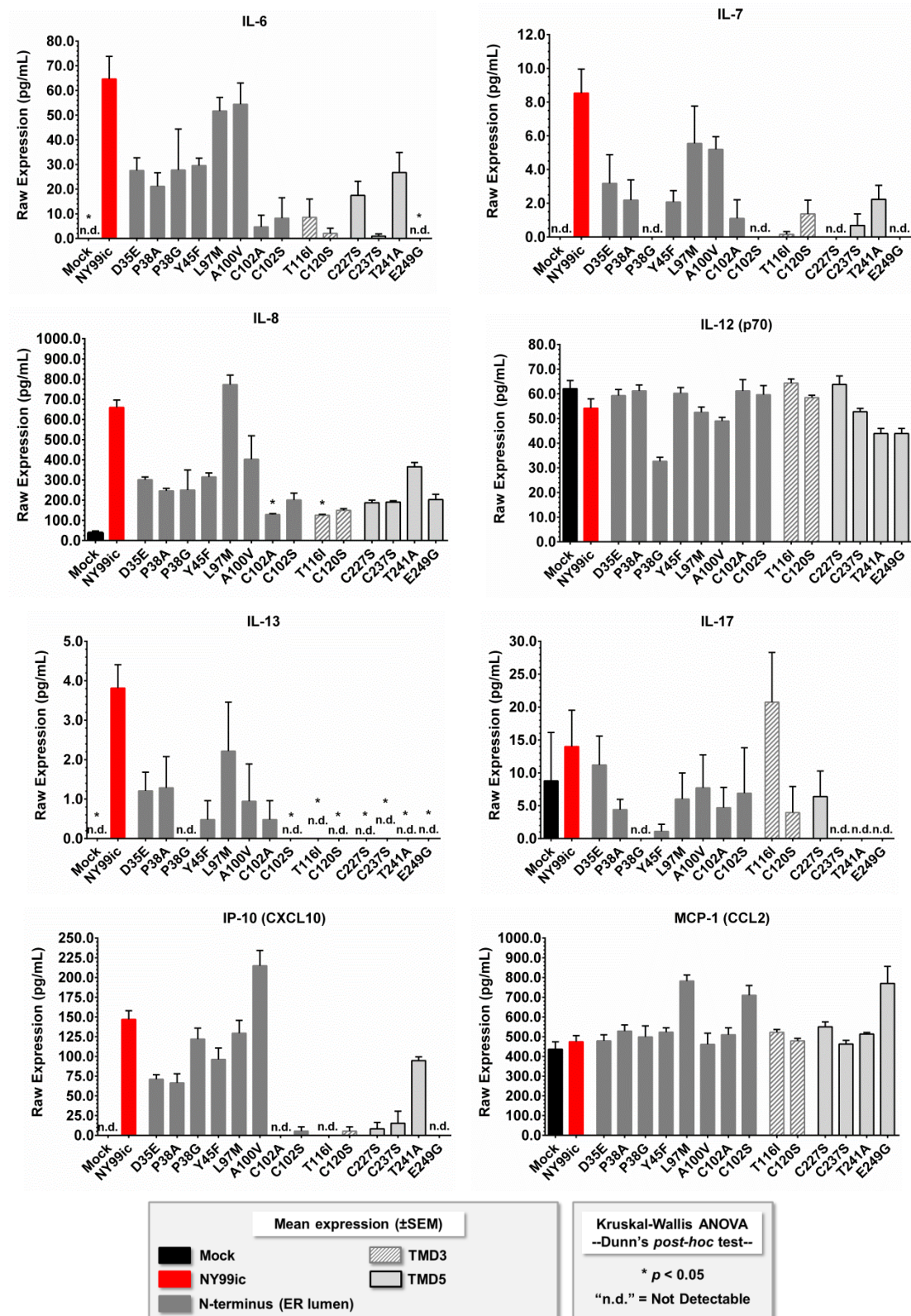


Appendix C: NY99ic NS4B Mutant Gene Expression (Continued)

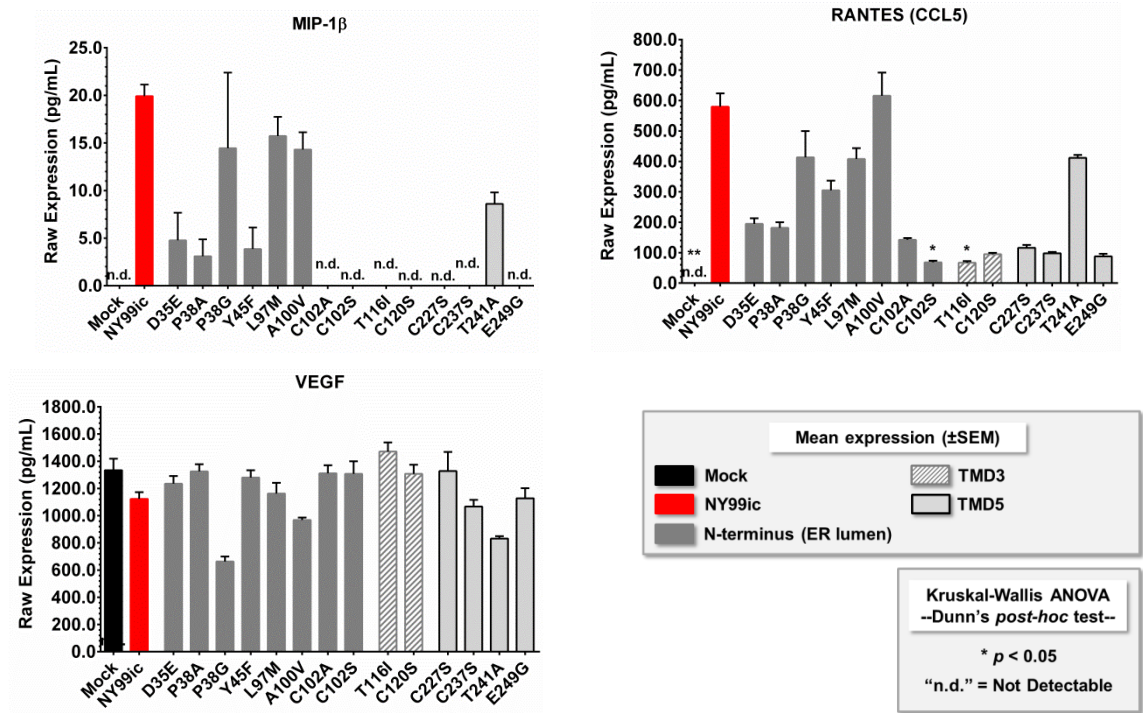


NOTE: Indicated fold-expression for P38G is derived from mock, P38G, and NY99ic samples collected in the initial *in vitro* A549 expression platform.

Appendix D: NY99ic NS4B Mutant Protein Expression

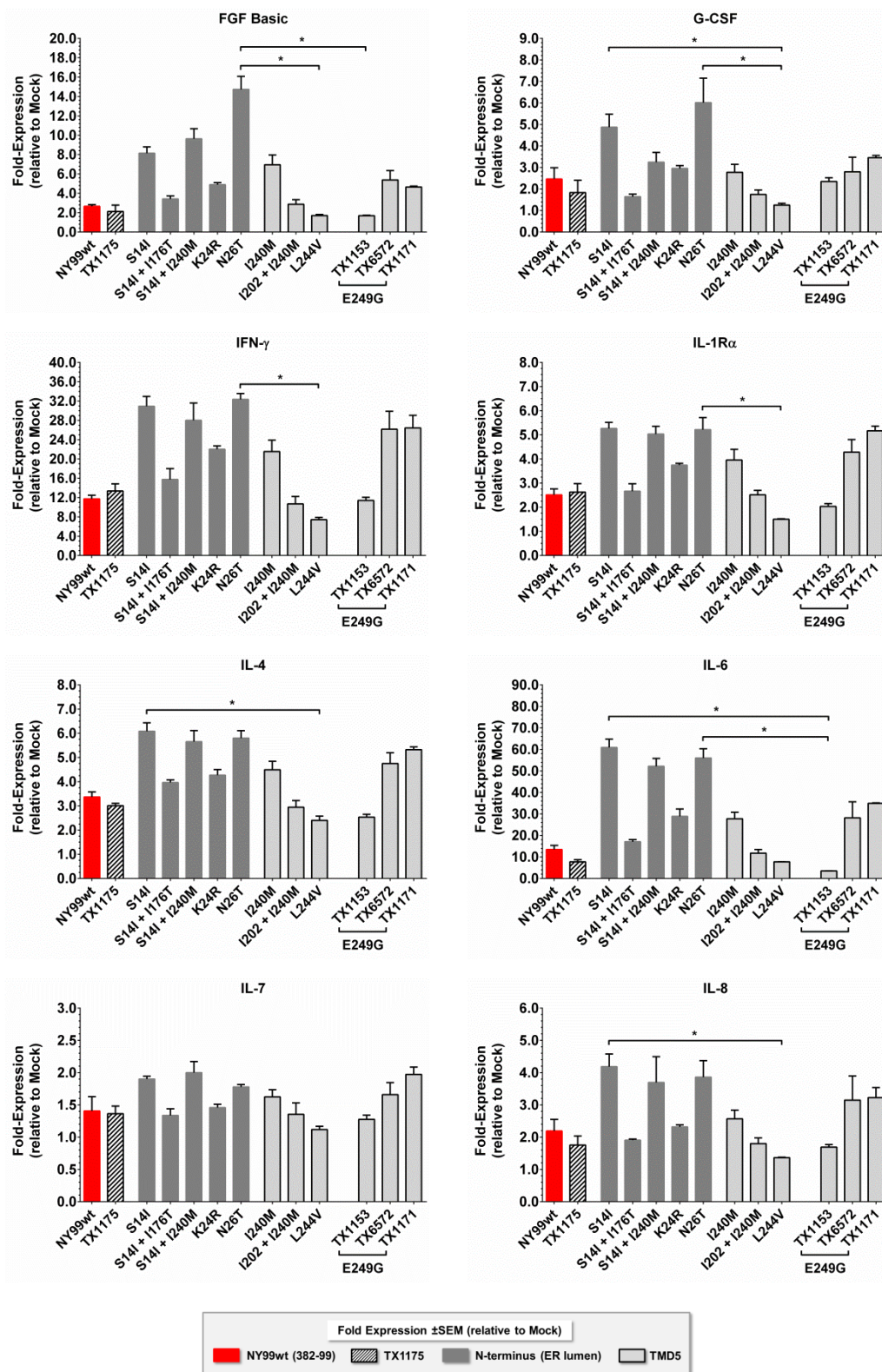


Appendix D: NY99ic NS4B Mutant Protein Expression (Continued)

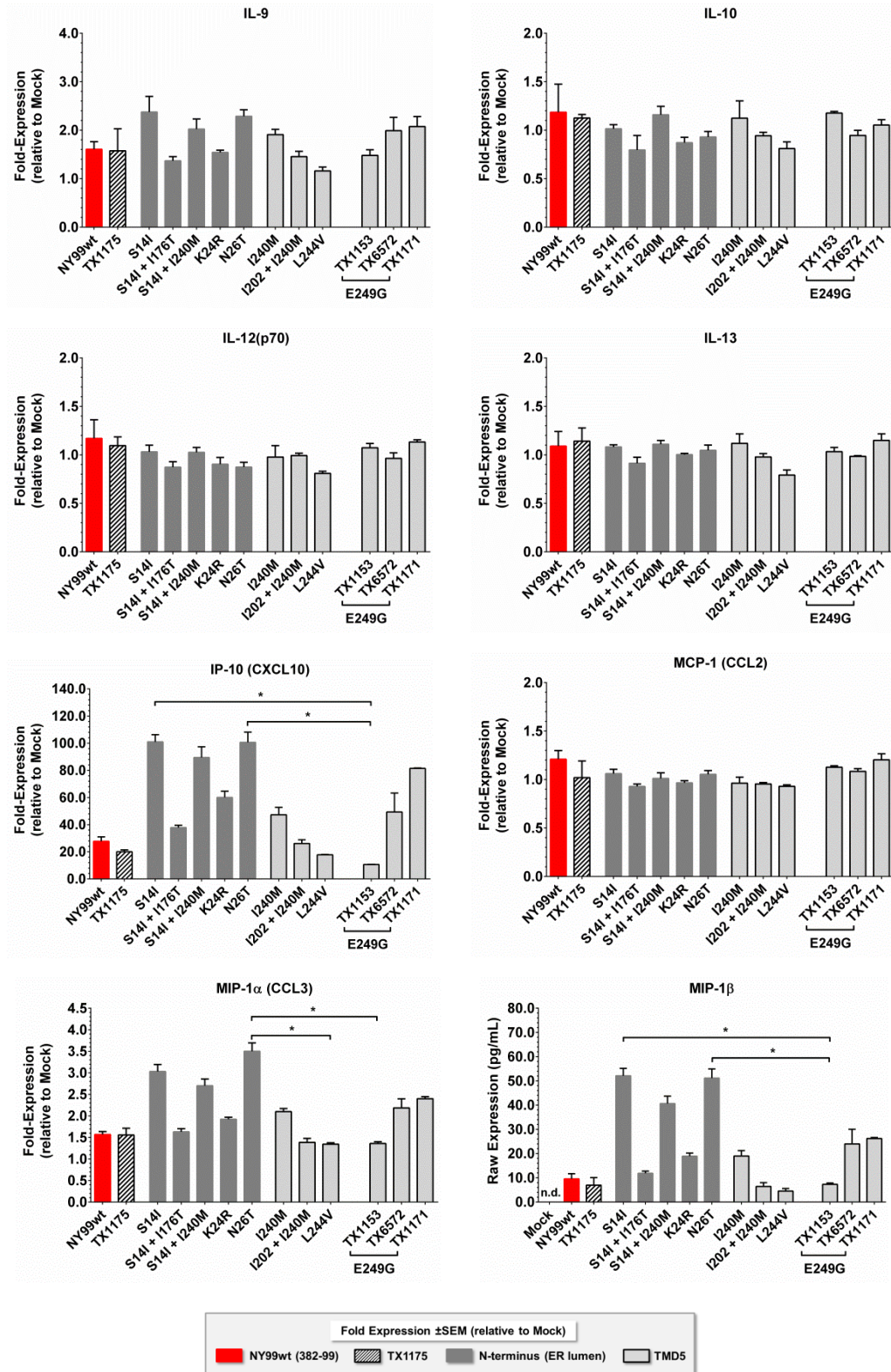


Raw expression (pg/mL) for 16 of the 27 target proteins in the BioRad® Bio-Plex Pro™ 27-plex Group I Cytokine panel was below the limit-of-detection (data not shown). Indicated expression for NS4B-P38G/T116I (+NS3-N480H) is representative of two individual Bio-Plex Pro™ panels.

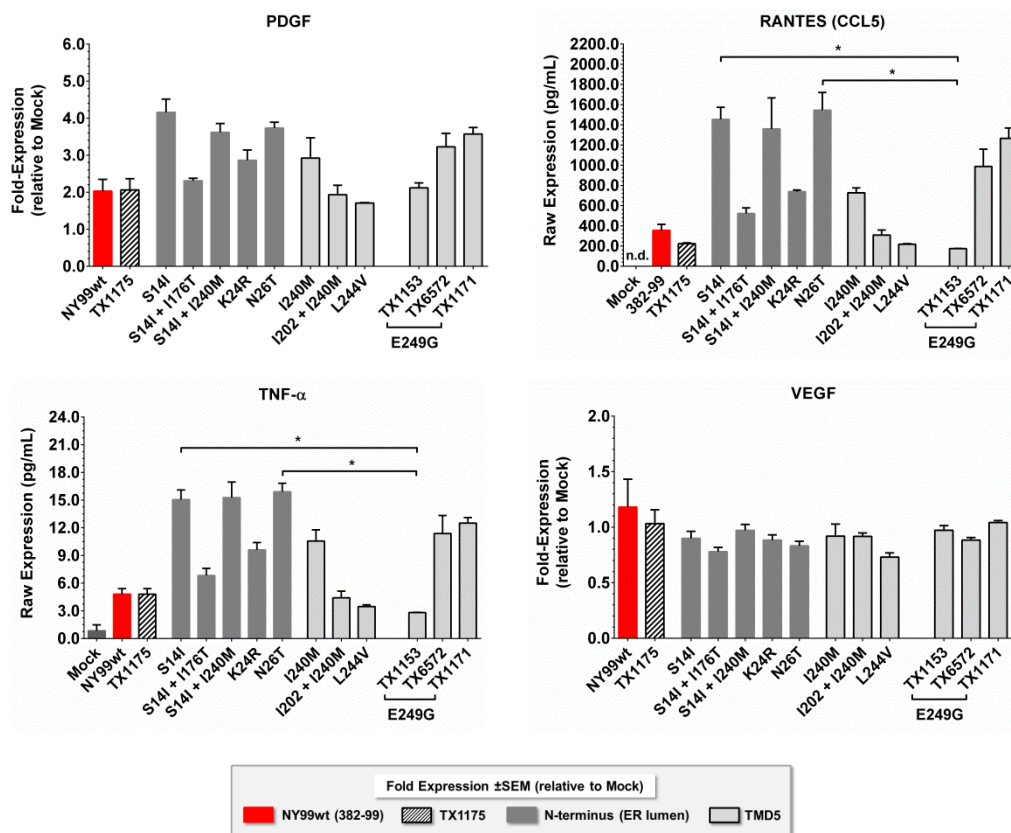
Appendix E: Natural NS4B Mutant Protein Expression



Appendix E: Natural NS4B Mutant Protein Expression (Continued)



Appendix E: Natural NS4B Mutant Protein Expression (Continued)



Fold (and/or raw) protein expression for seven of the 27 targets in the BioRad® Bio-Plex Pro™ 27-plex Group I Cytokine panel were below the standard curve effective range or limit-of-detection: eotaxin, GM-CSF, IL-1 β , IL-2, IL-5, IL-15, and IL-17 (data not shown). Furthermore, raw protein expression (pg/mL) is indicated for the MIP-1 β , RANTES (CCL5), and TNF- α protein targets in all applied natural NS4B mutants due to undetectable levels in the mock (PBS) negative control. No statistical significance was confirmed (relative to NY99wt) via a Kruskal-Wallis One-Way ANOVA and Dunn's *post-hoc* test ($\alpha = 0.05$); however, both TX1153 (E249G) and TX AR12-1486 (L244V) exhibited significant (p -value < 0.05) down-regulation in cytokine/chemokine expression (as indicated) relative to either or both TX8567 (S14I) and TX8604 (N26T).

References

1. Smithburn, K., Hughes, T., Burke, A. & Paul, J. A neurotropic virus isolated from the blood of a native of Uganda. *American Journal of Tropical Medicine* **20**, 471-2 (1940).
2. Schlesinger, R.W. *The Togaviruses : biology structure, replication*, xv, 687 p. (Academic Press, New York, 1980).
3. Smithburn, K.C. & Jacobs, H.R. Neutralization-tests against neurotropic viruses with sera collected in Central Africa. *J Immunol* **44**, 25-31 (1942).
4. Casals, J. & Brown, L.V. Hemagglutination with arthropod-borne viruses. *J Exp Med* **99**, 429-49 (1954).
5. De Madrid, A.T. & Porterfield, J.S. The flaviviruses (group B arboviruses): a cross-neutralization study. *J Gen Virol* **23**, 91-6 (1974).
6. Porterfield, J. Antigenic characteristics and classification of Togaviridae. *The Togaviruses*, 13-46 (1980).
7. Westaway, E. *et al.* Flaviviridae. *Intervirology* **24**, 183-192 (1985).
8. Calisher, C.H. *et al.* Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. *J Gen Virol* **70** (Pt 1), 37-43 (1989).
9. Poidinger, M., Hall, R.A. & Mackenzie, J.S. Molecular characterization of the Japanese encephalitis serocomplex of the flavivirus genus. *Virology* **218**, 417-21 (1996).
10. Murgue, B., Murri, S., Triki, H., Deubel, V. & Zeller, H.G. West Nile in the Mediterranean basin: 1950-2000. *Ann N Y Acad Sci* **951**, 117-26 (2001).
11. Mackenzie, J.S., Gubler, D.J. & Petersen, L.R. Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. *Nat Med* **10**, S98-109 (2004).
12. Mackenzie, J.S. & Williams, D.T. The zoonotic flaviviruses of southern, south-eastern and eastern Asia, and Australasia: the potential for emergent viruses. *Zoonoses Public Health* **56**, 338-56 (2009).

13. May, F.J., Davis, C.T., Tesh, R.B. & Barrett, A.D. Phylogeography of West Nile virus: from the cradle of evolution in Africa to Eurasia, Australia, and the Americas. *J Virol* **85**, 2964-74 (2011).
14. Mann, B.R., McMullen, A.R., Swetnam, D.M. & Barrett, A.D. Molecular epidemiology and evolution of West Nile virus in North America. *Int J Environ Res Public Health* **10**, 5111-29 (2013).
15. Hall, R.A., Scherret, J.H. & Mackenzie, J.S. Kunjin virus: an Australian variant of West Nile? *Ann N Y Acad Sci* **951**, 153-60 (2001).
16. Scherret, J.H. *et al.* The relationships between West Nile and Kunjin viruses. *Emerg Infect Dis* **7**, 697-705 (2001).
17. Bagnarelli, P. *et al.* Human case of autochthonous West Nile virus lineage 2 infection in Italy, September 2011. *Euro Surveill* **16**(2011).
18. Papa, A. *et al.* Genetic characterization of West Nile virus lineage 2, Greece, 2010. *Emerg Infect Dis* **17**, 920-2 (2011).
19. Papa, A., Xanthopoulou, K., Gewehr, S. & Mourelatos, S. Detection of West Nile virus lineage 2 in mosquitoes during a human outbreak in Greece. *Clin Microbiol Infect* **17**, 1176-80 (2011).
20. Sirbu, A. *et al.* Outbreak of West Nile virus infection in humans, Romania, July to October 2010. *Euro Surveill* **16**(2011).
21. Zaayman, D. & Venter, M. West Nile virus neurologic disease in humans, South Africa, September 2008-may 2009. *Emerg Infect Dis* **18**, 2051-4 (2012).
22. Bakonyi, T. *et al.* Explosive spread of a neuroinvasive lineage 2 West Nile virus in Central Europe, 2008/2009. *Vet Microbiol* **165**, 61-70 (2013).
23. Ciccozzi, M. *et al.* Epidemiological history and phylogeography of West Nile virus lineage 2. *Infect Genet Evol* **17**, 46-50 (2013).
24. McMullen, A.R. *et al.* Molecular evolution of lineage 2 West Nile virus. *J Gen Virol* **94**, 318-25 (2013).
25. Rudolf, I. *et al.* West Nile virus lineage 2 isolated from *Culex modestus* mosquitoes in the Czech Republic, 2013: expansion of the European WNV endemic area to the North? *Euro Surveill* **19**, 2-5 (2014).

26. Barzon, L. *et al.* Whole genome sequencing and phylogenetic analysis of West Nile virus lineage 1 and lineage 2 from human cases of infection, Italy, August 2013. *Euro Surveill* **18**(2013).
27. Barzon, L. *et al.* Genome sequencing of West Nile Virus from human cases in Greece, 2012. *Viruses* **5**, 2311-9 (2013).
28. Beasley, D.W. *et al.* Genome sequence and attenuating mutations in West Nile virus isolate from Mexico. *Emerg Infect Dis* **10**, 2221-4 (2004).
29. Beasley, D.W., Li, L., Suderman, M.T. & Barrett, A.D. Mouse neuroinvasive phenotype of West Nile virus strains varies depending upon virus genotype. *Virology* **296**, 17-23 (2002).
30. Bakonyi, T., Hubalek, Z., Rudolf, I. & Nowotny, N. Novel flavivirus or new lineage of West Nile virus, central Europe. *Emerg Infect Dis* **11**, 225-31 (2005).
31. Aliota, M.T. *et al.* Characterization of Rabensburg virus, a flavivirus closely related to West Nile virus of the Japanese encephalitis antigenic group. *PLoS One* **7**, e39387 (2012).
32. Lvov, D.K. *et al.* West Nile virus and other zoonotic viruses in Russia: examples of emerging-reemerging situations. *Arch Virol Suppl*, 85-96 (2004).
33. Bondre, V.P., Jadi, R.S., Mishra, A.C., Yergolkar, P.N. & Arankalle, V.A. West Nile virus isolates from India: evidence for a distinct genetic lineage. *J Gen Virol* **88**, 875-84 (2007).
34. Sotelo, E. *et al.* Phylogenetic relationships of Western Mediterranean West Nile virus strains (1996-2010) using full-length genome sequences: single or multiple introductions? *J Gen Virol* **92**, 2512-22 (2011).
35. Zehender, G. *et al.* Phylogeography and epidemiological history of West Nile virus genotype 1a in Europe and the Mediterranean basin. *Infect Genet Evol* **11**, 646-53 (2011).
36. Barzon, L. *et al.* Genome sequence analysis of the first human West Nile virus isolated in Italy in 2009. *Euro Surveill* **14**(2009).
37. Barzon, L. *et al.* Human cases of West Nile Virus infection in north-eastern Italy, 15 June to 15 November 2010. *Euro Surveill* **16**(2011).

38. Barzon, L. *et al.* Novel West Nile virus lineage 1a full genome sequences from human cases of infection in north-eastern Italy, 2011. *Clin Microbiol Infect* **18**, E541-4 (2012).
39. Magurano, F. *et al.* Circulation of West Nile virus lineage 1 and 2 during an outbreak in Italy. *Clin Microbiol Infect* **18**, E545-7 (2012).
40. Rossini, G. *et al.* Heterogeneity of West Nile virus genotype 1a in Italy, 2011. *J Gen Virol* **94**, 314-7 (2013).
41. Brault, A.C. *et al.* A single positively selected West Nile viral mutation confers increased virogenesis in American crows. *Nat Genet* **39**, 1162-6 (2007).
42. Ebel, G.D. *et al.* Partial genetic characterization of West Nile virus strains, New York State, 2000. *Emerg Infect Dis* **7**, 650-3 (2001).
43. Lanciotti, R.S. *et al.* Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* **286**, 2333-7 (1999).
44. Bernard, K.A. *et al.* West Nile virus infection in birds and mosquitoes, New York State, 2000. *Emerg Infect Dis* **7**, 679-85 (2001).
45. Andreadis, T.G., Anderson, J.F. & Vossbrinck, C.R. Mosquito surveillance for West Nile virus in Connecticut, 2000: isolation from *Culex pipiens*, *Cx. restuans*, *Cx. salinarius*, and *Culiseta melanura*. *Emerg Infect Dis* **7**, 670-4 (2001).
46. Anderson, J.F. *et al.* A phylogenetic approach to following West Nile virus in Connecticut. *Proc Natl Acad Sci U S A* **98**, 12885-9 (2001).
47. U.S. Department of the Interior, U.S.G.S. West Nile virus: Historical Data. (2014).
48. (CDC), C.f.D.C.a.P. West Nile virus. (CDC, 2014).
49. Deardorff, E. *et al.* Introductions of West Nile virus strains to Mexico. *Emerg Infect Dis* **12**, 314-8 (2006).
50. Pepperell, C. *et al.* West Nile virus infection in 2002: morbidity and mortality among patients admitted to hospital in southcentral Ontario. *CMAJ* **168**, 1399-405 (2003).
51. Morales, M.A. *et al.* West Nile virus isolation from equines in Argentina, 2006. *Emerg Infect Dis* **12**, 1559-61 (2006).

52. Osorio, J.E. *et al.* Characterization of West Nile viruses isolated from captive American Flamingoes (*Phoenicopterus ruber*) in Medellin, Colombia. *Am J Trop Med Hyg* **87**, 565-72 (2012).
53. Estrada-Franco, J.G. *et al.* West Nile virus in Mexico: evidence of widespread circulation since July 2002. *Emerg Infect Dis* **9**, 1604-7 (2003).
54. Malkinson, M. *et al.* Introduction of West Nile virus in the Middle East by migrating white storks. *Emerg Infect Dis* **8**, 392-7 (2002).
55. Davis, C.T. *et al.* Genetic variation among temporally and geographically distinct West Nile virus isolates, United States, 2001, 2002. *Emerg Infect Dis* **9**, 1423-9 (2003).
56. Granwehr, B.P., Li, L., Davis, C.T., Beasley, D.W. & Barrett, A.D. Characterization of a West Nile virus isolate from a human on the Gulf Coast of Texas. *J Clin Microbiol* **42**, 5375-7 (2004).
57. Davis, C.T. *et al.* Emergence of attenuated West Nile virus variants in Texas, 2003. *Virology* **330**, 342-50 (2004).
58. Davis, C.T. *et al.* Phylogenetic analysis of North American West Nile virus isolates, 2001-2004: evidence for the emergence of a dominant genotype. *Virology* **342**, 252-65 (2005).
59. Ebel, G.D., Carricaburu, J., Young, D., Bernard, K.A. & Kramer, L.D. Genetic and phenotypic variation of West Nile virus in New York, 2000-2003. *Am J Trop Med Hyg* **71**, 493-500 (2004).
60. Beasley, D.W. *et al.* Limited evolution of West Nile virus has occurred during its southwesterly spread in the United States. *Virology* **309**, 190-5 (2003).
61. Davis, C.T. *et al.* Genetic stasis of dominant West Nile virus genotype, Houston, Texas. *Emerg Infect Dis* **13**, 601-4 (2007).
62. Snapinn, K.W. *et al.* Declining growth rate of West Nile virus in North America. *J Virol* **81**, 2531-4 (2007).
63. McMullen, A.R. *et al.* Evolution of new genotype of West Nile virus in North America. *Emerg Infect Dis* **17**, 785-93 (2011).

64. Anez, G. *et al.* Evolutionary dynamics of West Nile virus in the United States, 1999-2011: phylogeny, selection pressure and evolutionary time-scale analysis. *PLoS Negl Trop Dis* **7**, e2245 (2013).
65. Grinev, A. *et al.* Genetic analysis of West Nile virus isolates from an outbreak in Idaho, United States, 2006-2007. *Int J Environ Res Public Health* **10**, 4486-506 (2013).
66. Moudy, R.M., Meola, M.A., Morin, L.L., Ebel, G.D. & Kramer, L.D. A newly emergent genotype of West Nile virus is transmitted earlier and more efficiently by *Culex* mosquitoes. *Am J Trop Med Hyg* **77**, 365-70 (2007).
67. Anderson, J.F., Main, A.J., Cheng, G., Ferrandino, F.J. & Fikrig, E. Horizontal and vertical transmission of West Nile virus genotype NY99 by *Culex salinarius* and genotypes NY99 and WN02 by *Culex tarsalis*. *Am J Trop Med Hyg* **86**, 134-9 (2012).
68. Amore, G. *et al.* Multi-year evolutionary dynamics of West Nile virus in suburban Chicago, USA, 2005-2007. *Philos Trans R Soc Lond B Biol Sci* **365**, 1871-8 (2010).
69. Armstrong, P.M. *et al.* Molecular evolution of West Nile virus in a northern temperate region: Connecticut, USA 1999-2008. *Virology* **417**, 203-10 (2011).
70. Bowen, R.A. & Nemeth, N.M. Experimental infections with West Nile virus. *Curr Opin Infect Dis* **20**, 293-7 (2007).
71. Ciota, A.T. & Kramer, L.D. Vector-virus interactions and transmission dynamics of West Nile virus. *Viruses* **5**, 3021-47 (2013).
72. Hamer, G.L. *et al.* Fine-scale variation in vector host use and force of infection drive localized patterns of West Nile virus transmission. *PLoS One* **6**, e23767 (2011).
73. Styer, L.M., Meola, M.A. & Kramer, L.D. West Nile virus infection decreases fecundity of *Culex tarsalis* females. *J Med Entomol* **44**, 1074-85 (2007).
74. Ciota, A.T., Styer, L.M., Meola, M.A. & Kramer, L.D. The costs of infection and resistance as determinants of West Nile virus susceptibility in *Culex* mosquitoes. *BMC Ecol* **11**, 23 (2011).
75. Kilpatrick, A.M. *et al.* West Nile virus risk assessment and the bridge vector paradigm. *Emerg Infect Dis* **11**, 425-9 (2005).

76. Turell, M.J., Sardelis, M.R., Dohm, D.J. & O'Guinn, M.L. Potential North American vectors of West Nile virus. *Ann N Y Acad Sci* **951**, 317-24 (2001).
77. Tiawsirisup, S. *et al.* Susceptibility of fox squirrels (*Sciurus niger*) to West Nile virus by oral exposure. *Vector Borne Zoonotic Dis* **10**, 207-9 (2010).
78. Kiupel, M. *et al.* West Nile virus infection in Eastern fox squirrels (*Sciurus niger*). *Vet Pathol* **40**, 703-7 (2003).
79. Root, J.J. *et al.* Experimental infection of fox squirrels (*Sciurus niger*) with West Nile virus. *Am J Trop Med Hyg* **75**, 697-701 (2006).
80. Padgett, K.A. *et al.* West Nile virus infection in tree squirrels (Rodentia: Sciuridae) in California, 2004-2005. *Am J Trop Med Hyg* **76**, 810-3 (2007).
81. Platt, K.B. *et al.* West Nile virus viremia in eastern chipmunks (*Tamias striatus*) sufficient for infecting different mosquitoes. *Emerg Infect Dis* **13**, 831-7 (2007).
82. Tiawsirisup, S., Platt, K.B., Tucker, B.J. & Rowley, W.A. Eastern cottontail rabbits (*Sylvilagus floridanus*) develop West Nile virus viremias sufficient for infecting select mosquito species. *Vector Borne Zoonotic Dis* **5**, 342-50 (2005).
83. Molaei, G. *et al.* Host feeding pattern of *Culex quinquefasciatus* (Diptera: Culicidae) and its role in transmission of West Nile virus in Harris County, Texas. *Am J Trop Med Hyg* **77**, 73-81 (2007).
84. Campbell, R., Thiemann, T.C., Lemenager, D. & Reisen, W.K. Host-selection patterns of *Culex tarsalis* (Diptera: Culicidae) determine the spatial heterogeneity of West Nile virus enzootic activity in northern California. *J Med Entomol* **50**, 1303-9 (2013).
85. Kilpatrick, A.M., Kramer, L.D., Jones, M.J., Marra, P.P. & Daszak, P. West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. *PLoS Biol* **4**, e82 (2006).
86. Ebel, G.D., Rochlin, I., Longacker, J. & Kramer, L.D. *Culex restuans* (Diptera: Culicidae) relative abundance and vector competence for West Nile Virus. *J Med Entomol* **42**, 838-43 (2005).
87. Andreadis, T.G., Anderson, J.F., Vossbrinck, C.R. & Main, A.J. Epidemiology of West Nile virus in Connecticut: a five-year analysis of mosquito data 1999-2003. *Vector Borne Zoonotic Dis* **4**, 360-78 (2004).

88. Goldberg, T.L., Anderson, T.K. & Hamer, G.L. West Nile virus may have hitched a ride across the Western United States on *Culex tarsalis* mosquitoes. *Mol Ecol* **19**, 1518-9 (2010).
89. Venkatesan, M. & Rasgon, J.L. Population genetic data suggest a role for mosquito-mediated dispersal of West Nile virus across the western United States. *Mol Ecol* **19**, 1573-84 (2010).
90. Venkatesan, M., Westbrook, C.J., Hauer, M.C. & Rasgon, J.L. Evidence for a population expansion in the West Nile virus vector *Culex tarsalis*. *Mol Biol Evol* **24**, 1208-18 (2007).
91. Chamberlain, R.W. & Sudia, W.D. Mechanism of transmission of viruses by mosquitoes. *Annu Rev Entomol* **6**, 371-90 (1961).
92. Kramer, L.D., Hardy, J.L., Presser, S.B. & Houk, E.J. Dissemination barriers for western equine encephalomyelitis virus in *Culex tarsalis* infected after ingestion of low viral doses. *Am J Trop Med Hyg* **30**, 190-7 (1981).
93. Hardy, J.L., Houk, E.J., Kramer, L.D. & Reeves, W.C. Intrinsic factors affecting vector competence of mosquitoes for arboviruses. *Annu Rev Entomol* **28**, 229-62 (1983).
94. Grimstad, P.R., Paulson, S.L. & Craig, G.B., Jr. Vector competence of *Aedes hendersoni* (Diptera: Culicidae) for La Crosse virus and evidence of a salivary-gland escape barrier. *J Med Entomol* **22**, 447-53 (1985).
95. Dohm, D.J., O'Guinn, M.L. & Turell, M.J. Effect of environmental temperature on the ability of *Culex pipiens* (Diptera: Culicidae) to transmit West Nile virus. *J Med Entomol* **39**, 221-5 (2002).
96. Richards, S.L., Mores, C.N., Lord, C.C. & Tabachnick, W.J. Impact of extrinsic incubation temperature and virus exposure on vector competence of *Culex pipiens quinquefasciatus* Say (Diptera: Culicidae) for West Nile virus. *Vector Borne Zoonotic Dis* **7**, 629-36 (2007).
97. Kilpatrick, A.M., Fonseca, D.M., Ebel, G.D., Reddy, M.R. & Kramer, L.D. Spatial and temporal variation in vector competence of *Culex pipiens* and *Cx. restuans* mosquitoes for West Nile virus. *Am J Trop Med Hyg* **83**, 607-13 (2010).
98. Kilpatrick, A.M., Meola, M.A., Moudy, R.M. & Kramer, L.D. Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. *PLoS Pathog* **4**, e1000092 (2008).

99. Perez-Ramirez, E., Llorente, F. & Jimenez-Clavero, M.A. Experimental infections of wild birds with West Nile virus. *Viruses* **6**, 752-81 (2014).
100. Komar, N. *et al.* Experimental infection of North American birds with the New York 1999 strain of West Nile virus. *Emerg Infect Dis* **9**, 311-22 (2003).
101. Komar, N., Panella, N.A., Young, G.R., Brault, A.C. & Levy, C.E. Avian hosts of West Nile virus in Arizona. *Am J Trop Med Hyg* **89**, 474-81 (2013).
102. Work, T.H., Hurlbut, H.S. & Taylor, R.M. Indigenous wild birds of the Nile Delta as potential West Nile virus circulating reservoirs. *Am J Trop Med Hyg* **4**, 872-88 (1955).
103. Benson, T.J., Ward, M.P., Lampman, R.L., Raim, A. & Weatherhead, P.J. Implications of spatial patterns of roosting and movements of American robins for West Nile virus transmission. *Vector Borne Zoonotic Dis* **12**, 877-85 (2012).
104. Kilpatrick, A.M. Globalization, land use, and the invasion of West Nile virus. *Science* **334**, 323-7 (2011).
105. Molaei, G., Andreadis, T.G., Armstrong, P.M., Anderson, J.F. & Vossbrinck, C.R. Host feeding patterns of *Culex* mosquitoes and West Nile virus transmission, northeastern United States. *Emerg Infect Dis* **12**, 468-74 (2006).
106. Hamer, G.L. *et al.* Host selection by *Culex pipiens* mosquitoes and West Nile virus amplification. *Am J Trop Med Hyg* **80**, 268-78 (2009).
107. Dusek, R.J. *et al.* Prevalence of West Nile virus in migratory birds during spring and fall migration. *Am J Trop Med Hyg* **81**, 1151-8 (2009).
108. Langevin, S.A. *et al.* Host Competence and Helicase Activity Differences Exhibited by West Nile Viral Variants Expressing NS3-249 Amino Acid Polymorphisms. *PLoS One* **9**, e100802 (2014).
109. Brault, A.C. *et al.* Reduced avian virulence and viremia of West Nile virus isolates from Mexico and Texas. *Am J Trop Med Hyg* **85**, 758-67 (2011).
110. Murata, R. *et al.* Glycosylation of the West Nile Virus envelope protein increases in vivo and in vitro viral multiplication in birds. *Am J Trop Med Hyg* **82**, 696-704 (2010).

111. Totani, M., Yoshii, K., Kariwa, H. & Takashima, I. Glycosylation of the envelope protein of West Nile Virus affects its replication in chicks. *Avian Dis* **55**, 561-8 (2011).
112. Beasley, D.W. *et al.* Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. *J Virol* **79**, 8339-47 (2005).
113. Shirato, K. *et al.* Viral envelope protein glycosylation is a molecular determinant of the neuroinvasiveness of the New York strain of West Nile virus. *J Gen Virol* **85**, 3637-45 (2004).
114. Domingo, E. & Holland, J.J. RNA virus mutations and fitness for survival. *Annu Rev Microbiol* **51**, 151-78 (1997).
115. Ciota, A.T. & Kramer, L.D. Insights into arbovirus evolution and adaptation from experimental studies. *Viruses* **2**, 2594-617 (2010).
116. Deardorff, E.R. *et al.* West Nile virus experimental evolution in vivo and the trade-off hypothesis. *PLoS Pathog* **7**, e1002335 (2011).
117. Ciota, A.T., Ehrbar, D.J., Matarachiero, A.C., Van Slyke, G.A. & Kramer, L.D. The evolution of virulence of West Nile virus in a mosquito vector: implications for arbovirus adaptation and evolution. *BMC Evol Biol* **13**, 71 (2013).
118. Brackney, D.E. *et al.* West Nile virus genetic diversity is maintained during transmission by *Culex pipiens quinquefasciatus* mosquitoes. *PLoS One* **6**, e24466 (2011).
119. Fitzpatrick, K.A. *et al.* Population variation of West Nile virus confers a host-specific fitness benefit in mosquitoes. *Virology* **404**, 89-95 (2010).
120. Jerzak, G., Bernard, K.A., Kramer, L.D. & Ebel, G.D. Genetic variation in West Nile virus from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection. *J Gen Virol* **86**, 2175-83 (2005).
121. Jerzak, G.V., Bernard, K., Kramer, L.D., Shi, P.Y. & Ebel, G.D. The West Nile virus mutant spectrum is host-dependant and a determinant of mortality in mice. *Virology* **360**, 469-76 (2007).
122. Jerzak, G.V., Brown, I., Shi, P.Y., Kramer, L.D. & Ebel, G.D. Genetic diversity and purifying selection in West Nile virus populations are maintained during host switching. *Virology* **374**, 256-60 (2008).

123. Murgue, B., Zeller, H. & Deubel, V. The ecology and epidemiology of West Nile virus in Africa, Europe and Asia. *Curr Top Microbiol Immunol* **267**, 195-221 (2002).
124. Busch, M.P. *et al.* West Nile virus infections projected from blood donor screening data, United States, 2003. *Emerg Infect Dis* **12**, 395-402 (2006).
125. Carson, P.J. *et al.* Neuroinvasive disease and West Nile virus infection, North Dakota, USA, 1999-2008. *Emerg Infect Dis* **18**, 684-6 (2012).
126. Lindsey, N.P., Kuhn, S., Campbell, G.L. & Hayes, E.B. West Nile virus neuroinvasive disease incidence in the United States, 2002-2006. *Vector Borne Zoonotic Dis* **8**, 35-9 (2008).
127. Busch, M.P. *et al.* Screening the blood supply for West Nile virus RNA by nucleic acid amplification testing. *N Engl J Med* **353**, 460-7 (2005).
128. Pealer, L.N. *et al.* Transmission of West Nile virus through blood transfusion in the United States in 2002. *N Engl J Med* **349**, 1236-45 (2003).
129. Wengler, G. & Wengler, G. Terminal sequences of the genome and replicative-from RNA of the flavivirus West Nile virus: absence of poly(A) and possible role in RNA replication. *Virology* **113**, 544-55 (1981).
130. Shuman, S. Structure, mechanism, and evolution of the mRNA capping apparatus. *Prog Nucleic Acid Res Mol Biol* **66**, 1-40 (2001).
131. Brinton, M.A. The molecular biology of West Nile Virus: a new invader of the western hemisphere. *Annu Rev Microbiol* **56**, 371-402 (2002).
132. Perera, R. & Kuhn, R.J. Structural proteomics of dengue virus. *Curr Opin Microbiol* **11**, 369-77 (2008).
133. Brinton, M.A. Replication cycle and molecular biology of the West Nile virus. *Viruses* **6**, 13-53 (2014).
134. Mukhopadhyay, S., Kuhn, R.J. & Rossmann, M.G. A structural perspective of the flavivirus life cycle. *Nat Rev Microbiol* **3**, 13-22 (2005).
135. Ma, L., Jones, C.T., Groesch, T.D., Kuhn, R.J. & Post, C.B. Solution structure of dengue virus capsid protein reveals another fold. *Proc Natl Acad Sci U S A* **101**, 3414-9 (2004).

136. Wengler, G. & Wengler, G. Cell-associated West Nile flavivirus is covered with E+pre-M protein heterodimers which are destroyed and reorganized by proteolytic cleavage during virus release. *J Virol* **63**, 2521-6 (1989).
137. Murray, J.M., Aaskov, J.G. & Wright, P.J. Processing of the dengue virus type 2 proteins prM and C-prM. *J Gen Virol* **74** (Pt 2), 175-82 (1993).
138. Chambers, T.J., Halevy, M., Nestorowicz, A., Rice, C.M. & Lustig, S. West Nile virus envelope proteins: nucleotide sequence analysis of strains differing in mouse neuroinvasiveness. *J Gen Virol* **79** (Pt 10), 2375-80 (1998).
139. Beasley, D.W. & Barrett, A.D. Identification of neutralizing epitopes within structural domain III of the West Nile virus envelope protein. *J Virol* **76**, 13097-100 (2002).
140. Li, L., Barrett, A.D. & Beasley, D.W. Differential expression of domain III neutralizing epitopes on the envelope proteins of West Nile virus strains. *Virology* **335**, 99-105 (2005).
141. Volk, D.E. *et al.* Structure of yellow fever virus envelope protein domain III. *Virology* **394**, 12-8 (2009).
142. Zhang, S. *et al.* Role of BC loop residues in structure, function and antigenicity of the West Nile virus envelope protein receptor-binding domain III. *Virology* **403**, 85-91 (2010).
143. Langevin, S.A. *et al.* Envelope and pre-membrane protein structural amino acid mutations mediate diminished avian growth and virulence of a Mexican West Nile virus isolate. *J Gen Virol* **92**, 2810-20 (2011).
144. Ferlenghi, I. *et al.* Molecular organization of a recombinant subviral particle from tick-borne encephalitis virus. *Mol Cell* **7**, 593-602 (2001).
145. Lindenbach, B.D. & Rice, C.M. Molecular biology of flaviviruses. *Adv Virus Res* **59**, 23-61 (2003).
146. Welsch, S. *et al.* Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host Microbe* **5**, 365-75 (2009).
147. Flamand, M. *et al.* Dengue virus type 1 nonstructural glycoprotein NS1 is secreted from mammalian cells as a soluble hexamer in a glycosylation-dependent fashion. *J Virol* **73**, 6104-10 (1999).

148. Gutsche, I. *et al.* Secreted dengue virus nonstructural protein NS1 is an atypical barrel-shaped high-density lipoprotein. *Proc Natl Acad Sci U S A* **108**, 8003-8 (2011).
149. Hall, R.A. *et al.* Loss of dimerisation of the nonstructural protein NS1 of Kunjin virus delays viral replication and reduces virulence in mice, but still allows secretion of NS1. *Virology* **264**, 66-75 (1999).
150. Macdonald, J. *et al.* NS1 protein secretion during the acute phase of West Nile virus infection. *J Virol* **79**, 13924-33 (2005).
151. Winkler, G., Maxwell, S.E., Ruemmler, C. & Stollar, V. Newly synthesized dengue-2 virus nonstructural protein NS1 is a soluble protein but becomes partially hydrophobic and membrane-associated after dimerization. *Virology* **171**, 302-5 (1989).
152. Winkler, G., Randolph, V.B., Cleaves, G.R., Ryan, T.E. & Stollar, V. Evidence that the mature form of the flavivirus nonstructural protein NS1 is a dimer. *Virology* **162**, 187-96 (1988).
153. Lindenbach, B.D. & Rice, C.M. trans-Complementation of yellow fever virus NS1 reveals a role in early RNA replication. *J Virol* **71**, 9608-17 (1997).
154. Youn, S., Ambrose, R.L., Mackenzie, J.M. & Diamond, M.S. Non-structural protein-1 is required for West Nile virus replication complex formation and viral RNA synthesis. *Virol J* **10**, 339 (2013).
155. Whiteman, M.C. *et al.* Development and characterization of non-glycosylated E and NS1 mutant viruses as a potential candidate vaccine for West Nile virus. *Vaccine* **28**, 1075-83 (2010).
156. Whiteman, M.C. *et al.* Multiple amino acid changes at the first glycosylation motif in NS1 protein of West Nile virus are necessary for complete attenuation for mouse neuroinvasiveness. *Vaccine* **29**, 9702-10 (2011).
157. Avirutnan, P. *et al.* Vascular leakage in severe dengue virus infections: a potential role for the nonstructural viral protein NS1 and complement. *J Infect Dis* **193**, 1078-88 (2006).
158. Chung, K.M. *et al.* West Nile virus nonstructural protein NS1 inhibits complement activation by binding the regulatory protein factor H. *Proc Natl Acad Sci U S A* **103**, 19111-6 (2006).

159. Avirutnan, P. *et al.* Antagonism of the complement component C4 by flavivirus nonstructural protein NS1. *J Exp Med* **207**, 793-806 (2010).
160. Avirutnan, P. *et al.* Binding of flavivirus nonstructural protein NS1 to C4b binding protein modulates complement activation. *J Immunol* **187**, 424-33 (2011).
161. Somnuk, P., Hauhart, R.E., Atkinson, J.P., Diamond, M.S. & Avirutnan, P. N-linked glycosylation of dengue virus NS1 protein modulates secretion, cell-surface expression, hexamer stability, and interactions with human complement. *Virology* **413**, 253-64 (2011).
162. Chambers, T.J., Grakoui, A. & Rice, C.M. Processing of the yellow fever virus nonstructural polyprotein: a catalytically active NS3 proteinase domain and NS2B are required for cleavages at dibasic sites. *J Virol* **65**, 6042-50 (1991).
163. Wengler, G., Czaya, G., Farber, P.M. & Hegemann, J.H. In vitro synthesis of West Nile virus proteins indicates that the amino-terminal segment of the NS3 protein contains the active centre of the protease which cleaves the viral polyprotein after multiple basic amino acids. *J Gen Virol* **72** (Pt 4), 851-8 (1991).
164. Falgout, B., Miller, R.H. & Lai, C.J. Deletion analysis of dengue virus type 4 nonstructural protein NS2B: identification of a domain required for NS2B-NS3 protease activity. *J Virol* **67**, 2034-42 (1993).
165. Li, H., Clum, S., You, S., Ebner, K.E. & Padmanabhan, R. The serine protease and RNA-stimulated nucleoside triphosphatase and RNA helicase functional domains of dengue virus type 2 NS3 converge within a region of 20 amino acids. *J Virol* **73**, 3108-16 (1999).
166. Wengler, G. & Wengler, G. The NS 3 nonstructural protein of flaviviruses contains an RNA triphosphatase activity. *Virology* **197**, 265-73 (1993).
167. Patkar, C.G. & Kuhn, R.J. Yellow Fever virus NS3 plays an essential role in virus assembly independent of its known enzymatic functions. *J Virol* **82**, 3342-52 (2008).
168. Miller, S., Sparacio, S. & Bartenschlager, R. Subcellular localization and membrane topology of the Dengue virus type 2 Non-structural protein 4B. *J Biol Chem* **281**, 8854-63 (2006).
169. Zou, J. *et al.* Dimerization of flavivirus NS4B protein. *J Virol* **88**, 3379-91 (2014).

170. Kaufusi, P.H., Kelley, J.F., Yanagihara, R. & Nerurkar, V.R. Induction of endoplasmic reticulum-derived replication-competent membrane structures by West Nile virus non-structural protein 4B. *PLoS One* **9**, e84040 (2014).
171. Khromykh, A.A., Sedlak, P.L. & Westaway, E.G. cis- and trans-acting elements in flavivirus RNA replication. *J Virol* **74**, 3253-63 (2000).
172. Umareddy, I., Chao, A., Sampath, A., Gu, F. & Vasudevan, S.G. Dengue virus NS4B interacts with NS3 and dissociates it from single-stranded RNA. *J Gen Virol* **87**, 2605-14 (2006).
173. Youn, S. *et al.* Evidence for a genetic and physical interaction between nonstructural proteins NS1 and NS4B that modulates replication of West Nile virus. *J Virol* **86**, 7360-71 (2012).
174. Davis, C.T. *et al.* A combination of naturally occurring mutations in North American West Nile virus nonstructural protein genes and in the 3' untranslated region alters virus phenotype. *J Virol* **81**, 6111-6 (2007).
175. Evans, J.D. & Seeger, C. Differential effects of mutations in NS4B on West Nile virus replication and inhibition of interferon signaling. *J Virol* **81**, 11809-16 (2007).
176. Kakumani, P.K. *et al.* Role of RNA interference (RNAi) in dengue virus replication and identification of NS4B as an RNAi suppressor. *J Virol* **87**, 8870-83 (2013).
177. Munoz-Jordan, J.L. *et al.* Inhibition of alpha/beta interferon signaling by the NS4B protein of flaviviruses. *J Virol* **79**, 8004-13 (2005).
178. Welte, T. *et al.* Immune responses to an attenuated West Nile virus NS4B-P38G mutant strain. *Vaccine* **29**, 4853-61 (2011).
179. Wicker, J.A. *et al.* Mutational analysis of the West Nile virus NS4B protein. *Virology* **426**, 22-33 (2012).
180. Wicker, J.A. *et al.* A single amino acid substitution in the central portion of the West Nile virus NS4B protein confers a highly attenuated phenotype in mice. *Virology* **349**, 245-53 (2006).
181. Dong, H. *et al.* 2'-O methylation of internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathog* **8**, e1002642 (2012).

182. Egloff, M.P., Benarroch, D., Selisko, B., Romette, J.L. & Canard, B. An RNA cap (nucleoside-2'-O-)-methyltransferase in the flavivirus RNA polymerase NS5: crystal structure and functional characterization. *EMBO J* **21**, 2757-68 (2002).
183. Ray, D. *et al.* West Nile virus 5'-cap structure is formed by sequential guanine N-7 and ribose 2'-O methylations by nonstructural protein 5. *J Virol* **80**, 8362-70 (2006).
184. Daffis, S. *et al.* 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature* **468**, 452-6 (2010).
185. Malet, H. *et al.* Crystal structure of the RNA polymerase domain of the West Nile virus non-structural protein 5. *J Biol Chem* **282**, 10678-89 (2007).
186. Kapoor, M. *et al.* Association between NS3 and NS5 proteins of dengue virus type 2 in the putative RNA replicase is linked to differential phosphorylation of NS5. *J Biol Chem* **270**, 19100-6 (1995).
187. Yon, C. *et al.* Modulation of the nucleoside triphosphatase/RNA helicase and 5'-RNA triphosphatase activities of Dengue virus type 2 nonstructural protein 3 (NS3) by interaction with NS5, the RNA-dependent RNA polymerase. *J Biol Chem* **280**, 27412-9 (2005).
188. Yu, L., Takeda, K. & Markoff, L. Protein-protein interactions among West Nile non-structural proteins and transmembrane complex formation in mammalian cells. *Virology* **446**, 365-77 (2013).
189. Zou, G. *et al.* Functional analysis of two cavities in flavivirus NS5 polymerase. *J Biol Chem* **286**, 14362-72 (2011).
190. Ashour, J., Laurent-Rolle, M., Shi, P.Y. & Garcia-Sastre, A. NS5 of dengue virus mediates STAT2 binding and degradation. *J Virol* **83**, 5408-18 (2009).
191. Laurent-Rolle, M. *et al.* The NS5 protein of the virulent West Nile virus NY99 strain is a potent antagonist of type I interferon-mediated JAK-STAT signaling. *J Virol* **84**, 3503-15 (2010).
192. Roosendaal, J., Westaway, E.G., Khromykh, A. & Mackenzie, J.M. Regulated cleavages at the West Nile virus NS4A-2K-NS4B junctions play a major role in rearranging cytoplasmic membranes and Golgi trafficking of the NS4A protein. *J Virol* **80**, 4623-32 (2006).

193. Miller, S., Kastner, S., Krijnse-Locker, J., Buhler, S. & Bartenschlager, R. The non-structural protein 4A of dengue virus is an integral membrane protein inducing membrane alterations in a 2K-regulated manner. *J Biol Chem* **282**, 8873-82 (2007).
194. Leung, J.Y. *et al.* Role of nonstructural protein NS2A in flavivirus assembly. *J Virol* **82**, 4731-41 (2008).
195. Liu, W.J., Chen, H.B. & Khromykh, A.A. Molecular and functional analyses of Kunjin virus infectious cDNA clones demonstrate the essential roles for NS2A in virus assembly and for a nonconservative residue in NS3 in RNA replication. *J Virol* **77**, 7804-13 (2003).
196. Xie, X., Gayen, S., Kang, C., Yuan, Z. & Shi, P.Y. Membrane topology and function of dengue virus NS2A protein. *J Virol* **87**, 4609-22 (2013).
197. Liu, W.J., Chen, H.B., Wang, X.J., Huang, H. & Khromykh, A.A. Analysis of adaptive mutations in Kunjin virus replicon RNA reveals a novel role for the flavivirus nonstructural protein NS2A in inhibition of beta interferon promoter-driven transcription. *J Virol* **78**, 12225-35 (2004).
198. Liu, W.J. *et al.* A single amino acid substitution in the West Nile virus nonstructural protein NS2A disables its ability to inhibit alpha/beta interferon induction and attenuates virus virulence in mice. *J Virol* **80**, 2396-404 (2006).
199. Liu, W.J. *et al.* Inhibition of interferon signaling by the New York 99 strain and Kunjin subtype of West Nile virus involves blockage of STAT1 and STAT2 activation by nonstructural proteins. *J Virol* **79**, 1934-42 (2005).
200. Markoff, L. 5'- and 3'-noncoding regions in flavivirus RNA. *Adv Virus Res* **59**, 177-228 (2003).
201. Beasley, D.W., Li, L., Suderman, M.T. & Barrett, A.D. West Nile virus strains differ in mouse neurovirulence and binding to mouse or human brain membrane receptor preparations. *Ann N Y Acad Sci* **951**, 332-5 (2001).
202. Alvarez, D.E., De Lella Ezcurra, A.L., Fucito, S. & Gamarnik, A.V. Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication. *Virology* **339**, 200-12 (2005).
203. Brinton, M.A. & Dispoto, J.H. Sequence and secondary structure analysis of the 5'-terminal region of flavivirus genome RNA. *Virology* **162**, 290-9 (1988).

204. Khromykh, A.A., Kondratieva, N., Sgro, J.Y., Palmenberg, A. & Westaway, E.G. Significance in replication of the terminal nucleotides of the flavivirus genome. *J Virol* **77**, 10623-9 (2003).
205. Khromykh, A.A., Meka, H., Guyatt, K.J. & Westaway, E.G. Essential role of cyclization sequences in flavivirus RNA replication. *J Virol* **75**, 6719-28 (2001).
206. Lo, M.K., Tilgner, M., Bernard, K.A. & Shi, P.Y. Functional analysis of mosquito-borne flavivirus conserved sequence elements within 3' untranslated region of West Nile virus by use of a reporting replicon that differentiates between viral translation and RNA replication. *J Virol* **77**, 10004-14 (2003).
207. Chu, J.J., Leong, P.W. & Ng, M.L. Analysis of the endocytic pathway mediating the infectious entry of mosquito-borne flavivirus West Nile into *Aedes albopictus* mosquito (C6/36) cells. *Virology* **349**, 463-75 (2006).
208. Chu, J.J. & Ng, M.L. Infectious entry of West Nile virus occurs through a clathrin-mediated endocytic pathway. *J Virol* **78**, 10543-55 (2004).
209. Lee, E., Hall, R.A. & Lobigs, M. Common E protein determinants for attenuation of glycosaminoglycan-binding variants of Japanese encephalitis and West Nile viruses. *J Virol* **78**, 8271-80 (2004).
210. Davis, C.W. *et al.* West Nile virus discriminates between DC-SIGN and DC-SIGNR for cellular attachment and infection. *J Virol* **80**, 1290-301 (2006).
211. Lee, J.W., Chu, J.J. & Ng, M.L. Quantifying the specific binding between West Nile virus envelope domain III protein and the cellular receptor alphaVbeta3 integrin. *J Biol Chem* **281**, 1352-60 (2006).
212. Schmidt, K. *et al.* Integrins modulate the infection efficiency of West Nile virus into cells. *J Gen Virol* **94**, 1723-33 (2013).
213. van der Schaar, H.M. *et al.* Dissecting the cell entry pathway of dengue virus by single-particle tracking in living cells. *PLoS Pathog* **4**, e1000244 (2008).
214. Allison, S.L. *et al.* Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *J Virol* **69**, 695-700 (1995).
215. Pierson, T.C. & Kielian, M. Flaviviruses: braking the entering. *Curr Opin Virol* **3**, 3-12 (2013).

216. Corver, J. *et al.* Membrane fusion activity of tick-borne encephalitis virus and recombinant subviral particles in a liposomal model system. *Virology* **269**, 37-46 (2000).
217. Modis, Y., Ogata, S., Clements, D. & Harrison, S.C. Structure of the dengue virus envelope protein after membrane fusion. *Nature* **427**, 313-9 (2004).
218. Stiasny, K., Allison, S.L., Marchler-Bauer, A., Kunz, C. & Heinz, F.X. Structural requirements for low-pH-induced rearrangements in the envelope glycoprotein of tick-borne encephalitis virus. *J Virol* **70**, 8142-7 (1996).
219. Stiasny, K., Allison, S.L., Schlich, J. & Heinz, F.X. Membrane interactions of the tick-borne encephalitis virus fusion protein E at low pH. *J Virol* **76**, 3784-90 (2002).
220. Elshuber, S., Allison, S.L., Heinz, F.X. & Mandl, C.W. Cleavage of protein prM is necessary for infection of BHK-21 cells by tick-borne encephalitis virus. *J Gen Virol* **84**, 183-91 (2003).
221. Clyde, K. & Harris, E. RNA secondary structure in the coding region of dengue virus type 2 directs translation start codon selection and is required for viral replication. *J Virol* **80**, 2170-82 (2006).
222. Clyde, K., Barrera, J. & Harris, E. The capsid-coding region hairpin element (cHP) is a critical determinant of dengue virus and West Nile virus RNA synthesis. *Virology* **379**, 314-23 (2008).
223. Dong, H., Zhang, B. & Shi, P.Y. Terminal structures of West Nile virus genomic RNA and their interactions with viral NS5 protein. *Virology* **381**, 123-35 (2008).
224. Davis, W.G., Blackwell, J.L., Shi, P.Y. & Brinton, M.A. Interaction between the cellular protein eEF1A and the 3'-terminal stem-loop of West Nile virus genomic RNA facilitates viral minus-strand RNA synthesis. *J Virol* **81**, 10172-87 (2007).
225. Blackwell, J.L. & Brinton, M.A. Translation elongation factor-1 alpha interacts with the 3' stem-loop region of West Nile virus genomic RNA. *J Virol* **71**, 6433-44 (1997).
226. Mackenzie, J.M., Jones, M.K. & Young, P.R. Immunolocalization of the dengue virus nonstructural glycoprotein NS1 suggests a role in viral RNA replication. *Virology* **220**, 232-40 (1996).

227. Westaway, E.G., Mackenzie, J.M., Kenney, M.T., Jones, M.K. & Khromykh, A.A. Ultrastructure of Kunjin virus-infected cells: colocalization of NS1 and NS3 with double-stranded RNA, and of NS2B with NS3, in virus-induced membrane structures. *J Virol* **71**, 6650-61 (1997).
228. Zhou, Y. *et al.* Structure and function of flavivirus NS5 methyltransferase. *J Virol* **81**, 3891-903 (2007).
229. Cleaves, G.R., Ryan, T.E. & Schlesinger, R.W. Identification and characterization of type 2 dengue virus replicative intermediate and replicative form RNAs. *Virology* **111**, 73-83 (1981).
230. Johnston, L.J., Halliday, G.M. & King, N.J. Langerhans cells migrate to local lymph nodes following cutaneous infection with an arbovirus. *J Invest Dermatol* **114**, 560-8 (2000).
231. Lim, P.Y., Behr, M.J., Chadwick, C.M., Shi, P.Y. & Bernard, K.A. Keratinocytes are cell targets of West Nile virus in vivo. *J Virol* **85**, 5197-201 (2011).
232. Colton, L., Biggerstaff, B.J., Johnson, A. & Nasci, R.S. Quantification of West Nile virus in vector mosquito saliva. *J Am Mosq Control Assoc* **21**, 49-53 (2005).
233. Styer, L.M. *et al.* Mosquitoes inoculate high doses of West Nile virus as they probe and feed on live hosts. *PLoS Pathog* **3**, 1262-70 (2007).
234. Vanlandingham, D.L. *et al.* Real-time reverse transcriptase-polymerase chain reaction quantification of West Nile virus transmitted by *Culex pipiens quinquefasciatus*. *Am J Trop Med Hyg* **71**, 120-3 (2004).
235. Ben-Nathan, D., Huitinga, I., Lustig, S., van Rooijen, N. & Kobiler, D. West Nile virus neuroinvasion and encephalitis induced by macrophage depletion in mice. *Arch Virol* **141**, 459-69 (1996).
236. Bai, F. *et al.* A paradoxical role for neutrophils in the pathogenesis of West Nile virus. *J Infect Dis* **202**, 1804-12 (2010).
237. Suthar, M.S., Diamond, M.S. & Gale, M., Jr. West Nile virus infection and immunity. *Nat Rev Microbiol* **11**, 115-28 (2013).
238. Verma, S., Kumar, M., Gurjav, U., Lum, S. & Nerurkar, V.R. Reversal of West Nile virus-induced blood-brain barrier disruption and tight junction proteins degradation by matrix metalloproteinases inhibitor. *Virology* **397**, 130-8 (2010).

239. Verma, S. *et al.* West Nile virus infection modulates human brain microvascular endothelial cells tight junction proteins and cell adhesion molecules: Transmigration across the in vitro blood-brain barrier. *Virology* **385**, 425-33 (2009).
240. Wang, P. *et al.* Matrix metalloproteinase 9 facilitates West Nile virus entry into the brain. *J Virol* **82**, 8978-85 (2008).
241. Wang, S. *et al.* Drak2 contributes to West Nile virus entry into the brain and lethal encephalitis. *J Immunol* **181**, 2084-91 (2008).
242. Wang, T. *et al.* Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. *Nat Med* **10**, 1366-73 (2004).
243. Garcia-Tapia, D., Hassett, D.E., Mitchell, W.J., Jr., Johnson, G.C. & Kleiboeker, S.B. West Nile virus encephalitis: sequential histopathological and immunological events in a murine model of infection. *J Neurovirol* **13**, 130-8 (2007).
244. Samuel, M.A., Wang, H., Siddharthan, V., Morrey, J.D. & Diamond, M.S. Axonal transport mediates West Nile virus entry into the central nervous system and induces acute flaccid paralysis. *Proc Natl Acad Sci U S A* **104**, 17140-5 (2007).
245. Rawal, A., Gavin, P.J. & Sturgis, C.D. Cerebrospinal fluid cytology in seasonal epidemic West Nile virus meningo-encephalitis. *Diagn Cytopathol* **34**, 127-9 (2006).
246. Tyler, K.L., Pape, J., Goody, R.J., Corkill, M. & Kleinschmidt-DeMasters, B.K. CSF findings in 250 patients with serologically confirmed West Nile virus meningitis and encephalitis. *Neurology* **66**, 361-5 (2006).
247. Armah, H.B. *et al.* Systemic distribution of West Nile virus infection: postmortem immunohistochemical study of six cases. *Brain Pathol* **17**, 354-62 (2007).
248. Cheeran, M.C. *et al.* Differential responses of human brain cells to West Nile virus infection. *J Neurovirol* **11**, 512-24 (2005).
249. Shrestha, B., Gottlieb, D. & Diamond, M.S. Infection and injury of neurons by West Nile encephalitis virus. *J Virol* **77**, 13203-13 (2003).
250. Solomon, T. & Vaughn, D.W. Pathogenesis and clinical features of Japanese encephalitis and West Nile virus infections. *Curr Top Microbiol Immunol* **267**, 171-94 (2002).

251. Spigland, I., Jasinska-Klingberg, W., Hofshi, E. & Goldblum, N. [Clinical and laboratory observations in an outbreak of West Nile fever in Israel in 1957]. *Harefuah* **54**, 275-80; English & French abstracts 280-1 (1958).
252. Tsai, T.F., Popovici, F., Cernescu, C., Campbell, G.L. & Nedelcu, N.I. West Nile encephalitis epidemic in southeastern Romania. *Lancet* **352**, 767-71 (1998).
253. Davis, L.E. *et al.* West Nile virus neuroinvasive disease. *Ann Neurol* **60**, 286-300 (2006).
254. Asnis, D.S., Conetta, R., Teixeira, A.A., Waldman, G. & Sampson, B.A. The West Nile Virus outbreak of 1999 in New York: the Flushing Hospital experience. *Clin Infect Dis* **30**, 413-8 (2000).
255. Sejvar, J.J. *et al.* West Nile Virus-associated flaccid paralysis outcome. *Emerg Infect Dis* **12**, 514-6 (2006).
256. Sejvar, J.J. *et al.* Neurologic manifestations and outcome of West Nile virus infection. *JAMA* **290**, 511-5 (2003).
257. Bode, A.V., Sejvar, J.J., Pape, W.J., Campbell, G.L. & Marfin, A.A. West Nile virus disease: a descriptive study of 228 patients hospitalized in a 4-county region of Colorado in 2003. *Clin Infect Dis* **42**, 1234-40 (2006).
258. Murray, K. *et al.* Risk factors for encephalitis and death from West Nile virus infection. *Epidemiol Infect* **134**, 1325-32 (2006).
259. Murray, K. *et al.* Persistent infection with West Nile virus years after initial infection. *J Infect Dis* **201**, 2-4 (2010).
260. Tesh, R.B. *et al.* Persistent West Nile virus infection in the golden hamster: studies on its mechanism and possible implications for other flavivirus infections. *J Infect Dis* **192**, 287-95 (2005).
261. Tonry, J.H. *et al.* Persistent shedding of West Nile virus in urine of experimentally infected hamsters. *Am J Trop Med Hyg* **72**, 320-4 (2005).
262. Appler, K.K. *et al.* Persistence of West Nile virus in the central nervous system and periphery of mice. *PLoS One* **5**, e10649 (2010).
263. Isaacs, A. & Lindenmann, J. Virus interference. I. The interferon. *Proc R Soc Lond B Biol Sci* **147**, 258-67 (1957).

264. Isaacs, A., Lindenmann, J. & Valentine, R.C. Virus interference. II. Some properties of interferon. *Proc R Soc Lond B Biol Sci* **147**, 268-73 (1957).
265. Isaacs, A. & Westwood, M.A. Duration of protective action of interferon against infection with West Nile virus. *Nature* **184(Suppl 16)**, 1232-3 (1959).
266. Porterfield, J. A SIMPLE PLAQUE INHIBITION TEST FOR ANTIVIRAL AGENTS: APPLICATION TO ASSAY OF INTERFERON. *The Lancet* **274**, 326-327 (1959).
267. Keller, B.C. *et al.* Resistance to alpha/beta interferon is a determinant of West Nile virus replication fitness and virulence. *J Virol* **80**, 9424-34 (2006).
268. Samuel, M.A. & Diamond, M.S. Alpha/beta interferon protects against lethal West Nile virus infection by restricting cellular tropism and enhancing neuronal survival. *J Virol* **79**, 13350-61 (2005).
269. Lin, R.J., Liao, C.L., Lin, E. & Lin, Y.L. Blocking of the alpha interferon-induced Jak-Stat signaling pathway by Japanese encephalitis virus infection. *J Virol* **78**, 9285-94 (2004).
270. Munoz-Jordan, J.L., Sanchez-Burgos, G.G., Laurent-Rolle, M. & Garcia-Sastre, A. Inhibition of interferon signaling by dengue virus. *Proc Natl Acad Sci U S A* **100**, 14333-8 (2003).
271. Jiang, D. *et al.* Identification of five interferon-induced cellular proteins that inhibit west nile virus and dengue virus infections. *J Virol* **84**, 8332-41 (2010).
272. Szretter, K.J. *et al.* The interferon-inducible gene viperin restricts West Nile virus pathogenesis. *J Virol* **85**, 11557-66 (2011).
273. Quicke, K.M. & Suthar, M.S. The innate immune playbook for restricting West Nile virus infection. *Viruses* **5**, 2643-58 (2013).
274. Loo, Y.M. *et al.* Distinct RIG-I and MDA5 signaling by RNA viruses in innate immunity. *J Virol* **82**, 335-45 (2008).
275. Hornung, V. *et al.* 5'-Triphosphate RNA is the ligand for RIG-I. *Science* **314**, 994-7 (2006).
276. Lazear, H.M. *et al.* IRF-3, IRF-5, and IRF-7 coordinately regulate the type I IFN response in myeloid dendritic cells downstream of MAVS signaling. *PLoS Pathog* **9**, e1003118 (2013).

277. Schlee, M. *et al.* Recognition of 5' triphosphate by RIG-I helicase requires short blunt double-stranded RNA as contained in panhandle of negative-strand virus. *Immunity* **31**, 25-34 (2009).
278. Errett, J.S., Suthar, M.S., McMillan, A., Diamond, M.S. & Gale, M., Jr. The essential, nonredundant roles of RIG-I and MDA5 in detecting and controlling West Nile virus infection. *J Virol* **87**, 11416-25 (2013).
279. Daffis, S., Suthar, M.S., Szretter, K.J., Gale, M., Jr. & Diamond, M.S. Induction of IFN-beta and the innate antiviral response in myeloid cells occurs through an IPS-1-dependent signal that does not require IRF-3 and IRF-7. *PLoS Pathog* **5**, e1000607 (2009).
280. Fredericksen, B.L., Keller, B.C., Fornek, J., Katze, M.G. & Gale, M., Jr. Establishment and maintenance of the innate antiviral response to West Nile Virus involves both RIG-I and MDA5 signaling through IPS-1. *J Virol* **82**, 609-16 (2008).
281. Suthar, M.S. *et al.* IPS-1 is essential for the control of West Nile virus infection and immunity. *PLoS Pathog* **6**, e1000757 (2010).
282. Shipley, J.G., Vandergaast, R., Deng, L., Mariuzza, R.A. & Fredericksen, B.L. Identification of multiple RIG-I-specific pathogen associated molecular patterns within the West Nile virus genome and antigenome. *Virology* **432**, 232-8 (2012).
283. Diamond, M.S. & Gale, M., Jr. Cell-intrinsic innate immune control of West Nile virus infection. *Trends Immunol* **33**, 522-30 (2012).
284. Daffis, S., Samuel, M.A., Suthar, M.S., Gale, M., Jr. & Diamond, M.S. Toll-like receptor 3 has a protective role against West Nile virus infection. *J Virol* **82**, 10349-58 (2008).
285. Wilson, J.R., de Sessions, P.F., Leon, M.A. & Scholle, F. West Nile virus nonstructural protein 1 inhibits TLR3 signal transduction. *J Virol* **82**, 8262-71 (2008).
286. Baronti, C., Sire, J., de Lamballerie, X. & Querat, G. Nonstructural NS1 proteins of several mosquito-borne Flavivirus do not inhibit TLR3 signaling. *Virology* **404**, 319-30 (2010).
287. Town, T. *et al.* Toll-like receptor 7 mitigates lethal West Nile encephalitis via interleukin 23-dependent immune cell infiltration and homing. *Immunity* **30**, 242-53 (2009).

288. Welte, T. *et al.* Toll-like receptor 7-induced immune response to cutaneous West Nile virus infection. *J Gen Virol* **90**, 2660-8 (2009).
289. van Marle, G. *et al.* West Nile virus-induced neuroinflammation: glial infection and capsid protein-mediated neurovirulence. *J Virol* **81**, 10933-49 (2007).
290. Shirato, K., Miyoshi, H., Kariwa, H. & Takashima, I. The kinetics of proinflammatory cytokines in murine peritoneal macrophages infected with envelope protein-glycosylated or non-glycosylated West Nile virus. *Virus Res* **121**, 11-6 (2006).
291. Ramos, H.J. *et al.* IL-1beta signaling promotes CNS-intrinsic immune control of West Nile virus infection. *PLoS Pathog* **8**, e1003039 (2012).
292. Byrne, S.N., Halliday, G.M., Johnston, L.J. & King, N.J. Interleukin-1beta but not tumor necrosis factor is involved in West Nile virus-induced Langerhans cell migration from the skin in C57BL/6 mice. *J Invest Dermatol* **117**, 702-9 (2001).
293. Ambrose, R.L. & Mackenzie, J.M. West Nile virus differentially modulates the unfolded protein response to facilitate replication and immune evasion. *J Virol* **85**, 2723-32 (2011).
294. Xie, X. *et al.* Inhibition of dengue virus by targeting viral NS4B protein. *J Virol* **85**, 11183-95 (2011).
295. Xie, G. *et al.* A West Nile virus NS4B-P38G mutant strain induces adaptive immunity via TLR7-MyD88-dependent and independent signaling pathways. *Vaccine* **31**, 4143-51 (2013).
296. Elizondo-Quiroga, D. *et al.* West Nile Virus isolation in human and mosquitoes, Mexico. *Emerg Infect Dis* **11**, 1449-52 (2005).
297. Lorono-Pino, M.A. *et al.* Antibodies to influenza and West Nile viruses in horses in Mexico. *Vet Rec* **166**, 22-3 (2010).
298. Rios-Ibarra, C. *et al.* Fatal human case of West Nile disease, Mexico, 2009. *Emerg Infect Dis* **16**, 741-3 (2010).
299. Rodriguez Mde, L. *et al.* Serologic surveillance for West Nile virus and other flaviviruses in febrile patients, encephalitic patients, and asymptomatic blood donors in northern Mexico. *Vector Borne Zoonotic Dis* **10**, 151-7 (2010).

300. Hanley, K.A., Lee, J.J., Blaney, J.E., Jr., Murphy, B.R. & Whitehead, S.S. Paired charge-to-alanine mutagenesis of dengue virus type 4 NS5 generates mutants with temperature-sensitive, host range, and mouse attenuation phenotypes. *J Virol* **76**, 525-31 (2002).
301. Blaney, J.E., Jr. *et al.* Chemical mutagenesis of dengue virus type 4 yields mutant viruses which are temperature sensitive in vero cells or human liver cells and attenuated in mice. *J Virol* **75**, 9731-40 (2001).
302. Ciota, A.T., Ehrbar, D.J., Van Slyke, G.A., Willsey, G.G. & Kramer, L.D. Cooperative interactions in the West Nile virus mutant swarm. *BMC Evol Biol* **12**, 58 (2012).
303. Coffey, L.L., Beeharry, Y., Borderia, A.V., Blanc, H. & Vignuzzi, M. Arbovirus high fidelity variant loses fitness in mosquitoes and mice. *Proc Natl Acad Sci U S A* **108**, 16038-43 (2011).
304. Zeng, J. *et al.* Ribavirin-resistant variants of foot-and-mouth disease virus: the effect of restricted quasispecies diversity on viral virulence. *J Virol* **88**, 4008-20 (2014).
305. Giard, D.J. *et al.* In vitro cultivation of human tumors: establishment of cell lines derived from a series of solid tumors. *J Natl Cancer Inst* **51**, 1417-23 (1973).
306. Lieber, M., Smith, B., Szakal, A., Nelson-Rees, W. & Todaro, G. A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. *Int J Cancer* **17**, 62-70 (1976).
307. Hall, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl.Acids Symp. Ser.* **41**, 95-98 (1999).
308. Edgar, R.C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).
309. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792-7 (2004).
310. Sievers, F. & Higgins, D.G. Clustal Omega, accurate alignment of very large numbers of sequences. *Methods Mol Biol* **1079**, 105-16 (2014).
311. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).

312. McWilliam, H. *et al.* Analysis Tool Web Services from the EMBL-EBI. *Nucleic Acids Res* **41**, W597-600 (2013).
313. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221-4 (2010).
314. Stamatakis, A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* **22**, 2688-90 (2006).
315. Stamatakis, A., Hoover, P. & Rougemont, J. A rapid bootstrap algorithm for the RAxML Web servers. *Syst Biol* **57**, 758-71 (2008).
316. Miller, M.A., Pfeiffer, W. & Schwartz, T. The CIPRES science gateway: a community resource for phylogenetic analyses. in *Proceedings of the 2011 TeraGrid Conference: Extreme Digital Discovery* 1-8 (ACM, Salt Lake City, Utah, 2011).
317. Drummond, A.J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**, 214 (2007).
318. Drummond, A.J., Suchard, M.A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* **29**, 1969-73 (2012).
319. Delpont, W., Poon, A.F., Frost, S.D. & Kosakovsky Pond, S.L. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* **26**, 2455-7 (2010).
320. Pond, S.L. & Frost, S.D. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics* **21**, 2531-3 (2005).
321. Kosakovsky Pond, S.L. & Frost, S.D. Not so different after all: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol* **22**, 1208-22 (2005).
322. Pond, S.L. *et al.* Adaptation to different human populations by HIV-1 revealed by codon-based analyses. *PLoS Comput Biol* **2**, e62 (2006).
323. Bolger, A.M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* (2014).

324. Simpson, J.T. *et al.* ABySS: a parallel assembler for short read sequence data. *Genome Res* **19**, 1117-23 (2009).
325. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-9 (2012).
326. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078-9 (2009).
327. Wei, Z., Wang, W., Hu, P., Lyon, G.J. & Hakonarson, H. SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. *Nucleic Acids Res* **39**, e132 (2011).
328. Koboldt, D.C. *et al.* VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* **22**, 568-76 (2012).
329. Gerstung, M. *et al.* Reliable detection of subclonal single-nucleotide variants in tumour cell populations. *Nat Commun* **3**, 811 (2012).
330. Nishijima, N. *et al.* Dynamics of hepatitis B virus quasispecies in association with nucleos(t)ide analogue treatment determined by ultra-deep sequencing. *PLoS One* **7**, e35052 (2012).
331. Wright, C.F. *et al.* Beyond the consensus: dissecting within-host viral population diversity of foot-and-mouth disease virus by using next-generation genome sequencing. *J Virol* **85**, 2266-75 (2011).
332. Giulietti, A. *et al.* An overview of real-time quantitative PCR: applications to quantify cytokine gene expression. *Methods* **25**, 386-401 (2001).
333. 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**, 402-8 (2001).
334. Alonso-Padilla, J. *et al.* The continuous spread of West Nile virus (WNV): seroprevalence in asymptomatic horses. *Epidemiol Infect* **137**, 1163-8 (2009).
335. Blitvich, B.J. *et al.* Phylogenetic analysis of West Nile virus, Nuevo Leon State, Mexico. *Emerg Infect Dis* **10**, 1314-7 (2004).
336. Blitvich, B.J. *et al.* Serologic evidence of West Nile virus infection in horses, Coahuila State, Mexico. *Emerg Infect Dis* **9**, 853-6 (2003).

337. Ibarra-Juarez, L. *et al.* Detection of West Nile virus-specific antibodies and nucleic acid in horses and mosquitoes, respectively, in Nuevo Leon State, northern Mexico, 2006-2007. *Med Vet Entomol* **26**, 351-4 (2012).
338. Lorono-Pino, M.A. *et al.* Serologic evidence of West Nile virus infection in horses, Yucatan State, Mexico. *Emerg Infect Dis* **9**, 857-9 (2003).
339. Farfan-Ale, J.A. *et al.* Antibodies to West Nile virus in asymptomatic mammals, birds, and reptiles in the Yucatan Peninsula of Mexico. *Am J Trop Med Hyg* **74**, 908-14 (2006).
340. Aguirre, A.A., McLean, R.G., Cook, R.S. & Quan, T.J. Serologic survey for selected arboviruses and other potential pathogens in wildlife from Mexico. *J Wildl Dis* **28**, 435-42 (1992).
341. Tesh, R.B., Travassos da Rosa, A.P., Guzman, H., Araujo, T.P. & Xiao, S.Y. Immunization with heterologous flaviviruses protective against fatal West Nile encephalitis. *Emerg Infect Dis* **8**, 245-51 (2002).
342. Kopp, A. *et al.* Provenance and geographic spread of St. Louis encephalitis virus. *MBio* **4**, e00322-13 (2013).
343. Guerrero-Sanchez, S. *et al.* West Nile virus infection of birds, Mexico. *Emerg Infect Dis* **17**, 2245-52 (2011).
344. Jia, X.Y. *et al.* Genetic analysis of West Nile New York 1999 encephalitis virus. *Lancet* **354**, 1971-2 (1999).
345. Lillibridge, K.M. *et al.* The 2002 introduction of West Nile virus into Harris County, Texas, an area historically endemic for St. Louis encephalitis. *Am J Trop Med Hyg* **70**, 676-81 (2004).
346. Shackelford, C.E., Rozenburg, E.R., Hunter, W.C. & Lockwood, M.W. *Migration and the Migratory Birds of Texas: Who They Are and Where They Are Going*, 34 (Texas Parks and Wildlife, 2005).
347. Bertolotti, L., Kitron, U. & Goldberg, T.L. Diversity and evolution of West Nile virus in Illinois and the United States, 2002-2005. *Virology* **360**, 143-9 (2007).
348. Bertolotti, L. *et al.* Fine-scale genetic variation and evolution of West Nile Virus in a transmission "hot spot" in suburban Chicago, USA. *Virology* **374**, 381-9 (2008).

349. Services, T.D.o.S.H. Arbovirus activity in Texas: 2012 surveillance report. (ed. Infectious Disease Control Unit, Z.C.B.) (Texas Department of State Health Services, 2014).
350. Mann, B.R., McMullen, A.R., Guzman, H., Tesh, R.B. & Barrett, A.D. Dynamic transmission of West Nile virus across the United States-Mexican border. *Virology* **436**, 75-80 (2013).
351. Mann, B.R. *et al.* Continued evolution of West Nile virus, Houston, Texas, USA, 2002-2012. *Emerg Infect Dis* **19**, 1418-27 (2013).
352. Nasu, A. *et al.* Genetic heterogeneity of hepatitis C virus in association with antiviral therapy determined by ultra-deep sequencing. *PLoS One* **6**, e24907 (2011).
353. Henn, M.R. *et al.* Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. *PLoS Pathog* **8**, e1002529 (2012).
354. Forbi, J.C. *et al.* Intra-host diversity and evolution of hepatitis C virus endemic to Cote d'Ivoire. *J Med Virol* **86**, 765-71 (2014).
355. Acevedo, A., Brodsky, L. & Andino, R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* **505**, 686-90 (2014).
356. Beck, A. *et al.* Comparison of the live attenuated yellow fever vaccine 17D-204 strain to its virulent parental strain Asibi by deep sequencing. *J Infect Dis* **209**, 334-44 (2014).
357. Coffey, L.L. *et al.* Arbovirus evolution in vivo is constrained by host alternation. *Proc Natl Acad Sci U S A* **105**, 6970-5 (2008).
358. Ciota, A.T. *et al.* Experimental passage of St. Louis encephalitis virus in vivo in mosquitoes and chickens reveals evolutionarily significant virus characteristics. *PLoS One* **4**, e7876 (2009).
359. Vasilakis, N. *et al.* Mosquitoes put the brake on arbovirus evolution: experimental evolution reveals slower mutation accumulation in mosquito than vertebrate cells. *PLoS Pathog* **5**, e1000467 (2009).
360. Forrester, N.L., Guérbois, M., Seymour, R.L., Spratt, H. & Weaver, S.C. Vector-borne transmission imposes a severe bottleneck on an RNA virus population. *PLoS Pathog* **8**, e1002897 (2012).

361. Ciota, A.T. *et al.* Cell-specific adaptation of two flaviviruses following serial passage in mosquito cell culture. *Virology* **357**, 165-74 (2007).
362. Ciota, A.T. *et al.* Characterization of mosquito-adapted West Nile virus. *J Gen Virol* **89**, 1633-42 (2008).
363. Gray, R.R., Veras, N.M., Santos, L.A. & Salemi, M. Evolutionary characterization of the West Nile Virus complete genome. *Mol Phylogenet Evol* **56**, 195-200 (2010).
364. Hahn, C.S. *et al.* Conserved elements in the 3' untranslated region of flavivirus RNAs and potential cyclization sequences. *J Mol Biol* **198**, 33-41 (1987).
365. Zhang, B., Dong, H., Stein, D.A., Iversen, P.L. & Shi, P.Y. West Nile virus genome cyclization and RNA replication require two pairs of long-distance RNA interactions. *Virology* **373**, 1-13 (2008).
366. Eigen, M. Error catastrophe and antiviral strategy. *Proc Natl Acad Sci U S A* **99**, 13374-6 (2002).
367. Vignuzzi, M., Stone, J.K. & Andino, R. Ribavirin and lethal mutagenesis of poliovirus: molecular mechanisms, resistance and biological implications. *Virus Res* **107**, 173-81 (2005).
368. Pfeiffer, J.K. & Kirkegaard, K. Increased fidelity reduces poliovirus fitness and virulence under selective pressure in mice. *PLoS Pathog* **1**, e11 (2005).
369. Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E. & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344-8 (2006).
370. Vignuzzi, M., Wendt, E. & Andino, R. Engineering attenuated virus vaccines by controlling replication fidelity. *Nat Med* **14**, 154-61 (2008).
371. Deo, S. *et al.* Activation of 2' 5'-oligoadenylate synthetase by stem loops at the 5'-end of the West Nile virus genome. *PLoS One* **9**, e92545 (2014).
372. Hsu, Y.L., Shi, S.F., Wu, W.L., Ho, L.J. & Lai, J.H. Protective roles of interferon-induced protein with tetratricopeptide repeats 3 (IFIT3) in dengue virus infection of human lung epithelial cells. *PLoS One* **8**, e79518 (2013).

373. Fredericksen, B.L. & Gale, M., Jr. West Nile virus evades activation of interferon regulatory factor 3 through RIG-I-dependent and -independent pathways without antagonizing host defense signaling. *J Virol* **80**, 2913-23 (2006).
374. Kuno, G. & Chang, G.J. Biological transmission of arboviruses: reexamination of and new insights into components, mechanisms, and unique traits as well as their evolutionary trends. *Clin Microbiol Rev* **18**, 608-37 (2005).
375. Puig-Basagoiti, F. *et al.* A mouse cell-adapted NS4B mutation attenuates West Nile virus RNA synthesis. *Virology* **361**, 229-41 (2007).
376. Rossi, S.L., Fayzulin, R., Dewsbury, N., Bourne, N. & Mason, P.W. Mutations in West Nile virus nonstructural proteins that facilitate replicon persistence in vitro attenuate virus replication in vitro and in vivo. *Virology* **364**, 184-95 (2007).
377. Daffis, S. *et al.* Interferon regulatory factor IRF-7 induces the antiviral alpha interferon response and protects against lethal West Nile virus infection. *J Virol* **82**, 8465-75 (2008).
378. Gack, M.U. *et al.* TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature* **446**, 916-920 (2007).
379. Wang, P. *et al.* Caspase-12 controls West Nile virus infection via the viral RNA receptor RIG-I. *Nat Immunol* **11**, 912-9 (2010).
380. Melian, E.B. *et al.* West Nile virus NS2A protein facilitates virus-induced apoptosis independently of interferon response. *J Gen Virol* **94**, 308-13 (2013).
381. Tapia, K. *et al.* Defective viral genomes arising in vivo provide critical danger signals for the triggering of lung antiviral immunity. *PLoS Pathog* **9**, e1003703 (2013).
382. Kumar, M., Verma, S. & Nerurkar, V.R. Pro-inflammatory cytokines derived from West Nile virus (WNV)-infected SK-N-SH cells mediate neuroinflammatory markers and neuronal death. *J Neuroinflammation* **7**, 73 (2010).
383. Qian, F. *et al.* Identification of genes critical for resistance to infection by West Nile virus using RNA-Seq analysis. *Viruses* **5**, 1664-81 (2013).
384. Kawai, T. *et al.* IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. *Nat Immunol* **6**, 981-8 (2005).

385. Suthar, M.S. *et al.* A systems biology approach reveals that tissue tropism to West Nile virus is regulated by antiviral genes and innate immune cellular processes. *PLoS Pathog* **9**, e1003168 (2013).
386. Shrestha, B., Zhang, B., Purtha, W.E., Klein, R.S. & Diamond, M.S. Tumor necrosis factor alpha protects against lethal West Nile virus infection by promoting trafficking of mononuclear leukocytes into the central nervous system. *J Virol* **82**, 8956-64 (2008).
387. Lin, R., Heylbroeck, C., Pitha, P.M. & Hiscott, J. Virus-dependent phosphorylation of the IRF-3 transcription factor regulates nuclear translocation, transactivation potential, and proteasome-mediated degradation. *Mol Cell Biol* **18**, 2986-96 (1998).
388. Grandvaux, N. *et al.* Transcriptional profiling of interferon regulatory factor 3 target genes: direct involvement in the regulation of interferon-stimulated genes. *J Virol* **76**, 5532-9 (2002).
389. Fredericksen, B.L., Smith, M., Katze, M.G., Shi, P.Y. & Gale, M., Jr. The host response to West Nile Virus infection limits viral spread through the activation of the interferon regulatory factor 3 pathway. *J Virol* **78**, 7737-47 (2004).
390. Genin, P., Algarte, M., Roof, P., Lin, R. & Hiscott, J. Regulation of RANTES chemokine gene expression requires cooperativity between NF-kappa B and IFN-regulatory factor transcription factors. *J Immunol* **164**, 5352-61 (2000).
391. May, F.J., Li, L., Davis, C.T., Galbraith, S.E. & Barrett, A.D. Multiple pathways to the attenuation of West Nile virus in south-east Texas in 2003. *Virology* **405**, 8-14 (2010).
392. Brault, A.C. *et al.* Differential virulence of West Nile strains for American crows. *Emerg Infect Dis* **10**, 2161-8 (2004).
393. Langevin, S.A., Brault, A.C., Panella, N.A., Bowen, R.A. & Komar, N. Variation in virulence of West Nile virus strains for house sparrows (*Passer domesticus*). *Am J Trop Med Hyg* **72**, 99-102 (2005).
394. Molaei, G. *et al.* Vector-host interactions governing epidemiology of West Nile virus in Southern California. *Am J Trop Med Hyg* **83**, 1269-82 (2010).
395. Duggal, N.K. *et al.* Sequence analyses of 2012 West Nile virus isolates from Texas fail to associate viral genetic factors with outbreak magnitude. *Am J Trop Med Hyg* **89**, 205-10 (2013).

396. Perales, C., Martin, V., Ruiz-Jarabo, C.M. & Domingo, E. Monitoring sequence space as a test for the target of selection in viruses. *J Mol Biol* **345**, 451-9 (2005).
397. Perales, C., Mateo, R., Mateu, M.G. & Domingo, E. Insights into RNA virus mutant spectrum and lethal mutagenesis events: replicative interference and complementation by multiple point mutants. *J Mol Biol* **369**, 985-1000 (2007).
398. Gregori, J. *et al.* Inference with viral quasispecies diversity indices: clonal and NGS approaches. *Bioinformatics* (2014).
399. Schlick, P. *et al.* Characterization of West Nile virus live vaccine candidates attenuated by capsid deletion mutations. *Vaccine* **28**, 5903-9 (2010).
400. Monath, T.P. *et al.* A live, attenuated recombinant West Nile virus vaccine. *Proc Natl Acad Sci U S A* **103**, 6694-9 (2006).
401. Arroyo, J. *et al.* ChimeriVax-West Nile virus live-attenuated vaccine: preclinical evaluation of safety, immunogenicity, and efficacy. *J Virol* **78**, 12497-507 (2004).
402. Widman, D.G., Ishikawa, T., Fayzulin, R., Bourne, N. & Mason, P.W. Construction and characterization of a second-generation pseudoinfectious West Nile virus vaccine propagated using a new cultivation system. *Vaccine* **26**, 2762-71 (2008).
403. Guirakhoo, F. *et al.* Immunogenicity, genetic stability, and protective efficacy of a recombinant, chimeric yellow fever-Japanese encephalitis virus (ChimeriVax-JE) as a live, attenuated vaccine candidate against Japanese encephalitis. *Virology* **257**, 363-72 (1999).
404. McGee, C.E. *et al.* Substitution of wild-type yellow fever Asibi sequences for 17D vaccine sequences in ChimeriVax-dengue 4 does not enhance infection of *Aedes aegypti* mosquitoes. *J Infect Dis* **197**, 686-92 (2008).
405. Ni, H. *et al.* Comparison of nucleotide and deduced amino acid sequence of the 5' non-coding region and structural protein genes of the wild-type Japanese encephalitis virus strain SA14 and its attenuated vaccine derivatives. *J Gen Virol* **75** (Pt 6), 1505-10 (1994).
406. Ni, H., Chang, G.J., Xie, H., Trent, D.W. & Barrett, A.D. Molecular basis of attenuation of neurovirulence of wild-type Japanese encephalitis virus strain SA14. *J Gen Virol* **76** (Pt 2), 409-13 (1995).

407. Hoenen, A., Liu, W., Kochs, G., Khromykh, A.A. & Mackenzie, J.M. West Nile virus-induced cytoplasmic membrane structures provide partial protection against the interferon-induced antiviral MxA protein. *J Gen Virol* **88**, 3013-7 (2007).
408. Lin, R., Heylbroeck, C., Genin, P., Pitha, P.M. & Hiscott, J. Essential role of interferon regulatory factor 3 in direct activation of RANTES chemokine transcription. *Mol Cell Biol* **19**, 959-66 (1999).
409. Kulkarni, A.B., Mullbacher, A. & Blanden, R.V. Functional analysis of macrophages, B cells and splenic dendritic cells as antigen-presenting cells in West Nile virus-specific murine T lymphocyte proliferation. *Immunol Cell Biol* **69** (Pt 2), 71-80 (1991).
410. Kong, K.F., Wang, X., Anderson, J.F., Fikrig, E. & Montgomery, R.R. West Nile virus attenuates activation of primary human macrophages. *Viral Immunol* **21**, 78-82 (2008).
411. Winkelmann, E.R. *et al.* Subcapsular sinus macrophages limit dissemination of West Nile virus particles after inoculation but are not essential for the development of West Nile virus-specific T cell responses. *Virology* **450-451**, 278-89 (2014).
412. Kesson, A.M., Blanden, R.V. & Mullbacher, A. The primary in vivo murine cytotoxic T cell response to the flavivirus, West Nile. *J Gen Virol* **68** (Pt 7), 2001-6 (1987).
413. Kulkarni, A.B., Mullbacher, A. & Blanden, R.V. In vitro T-cell proliferative response to the flavivirus, west Nile. *Viral Immunol* **4**, 73-82 (1991).
414. Shrestha, B. & Diamond, M.S. Role of CD8+ T cells in control of West Nile virus infection. *J Virol* **78**, 8312-21 (2004).
415. Sitati, E.M. & Diamond, M.S. CD4+ T-cell responses are required for clearance of West Nile virus from the central nervous system. *J Virol* **80**, 12060-9 (2006).
416. Wang, Y., Lobigs, M., Lee, E. & Mullbacher, A. CD8+ T cells mediate recovery and immunopathology in West Nile virus encephalitis. *J Virol* **77**, 13323-34 (2003).
417. Brown, A.N., Kent, K.A., Bennett, C.J. & Bernard, K.A. Tissue tropism and neuroinvasion of West Nile virus do not differ for two mouse strains with different survival rates. *Virology* **368**, 422-30 (2007).

Vitae

Brian Richard Mann was born in Toledo, Ohio on May 13, 1988 to Cindy Lynn Rowe and William Lee Mann. He graduated from Pike-Delta-York High School, Delta, Ohio in 2006 and from Wittenberg University, Springfield, Ohio in 2010 with a Bachelor of Science in Biochemistry and Molecular Biology. He commenced a Doctor of Philosophy in Experimental Pathology in 2010 at the University of Texas Medical Branch, Galveston, TX. He completed his dissertation work in the laboratory of Dr. Alan D.T. Barrett, investigating genotypic and molecular determinants involved in the continued evolution of West Nile virus in the United States using a combination of *in silico*, *in vitro*, and *in vivo* studies.

Permanent address: 6533 County Road 4, Swanton, OH 43558

This dissertation was typed by Brian Richard Mann.

Curriculum Vitae

NAME: Brian Richard Mann

DATE: 12/12/2014

CURRENT POSITION:

08/2010 - Present Doctoral Candidate
Graduate School of Biomedical Sciences
c/o University of Texas Medical Branch
301 University Blvd.
Galveston, TX 77555-0609
Laboratory Phone: (409) 772-2547
brmann@utmb.edu

BIOGRAPHICAL INFORMATION:

Date of Birth: May 13, 1988
Citizenship: United States

Home Address: 215 Post Office Street #1101
Galveston, TX 77550

Mobile Phone: (419) 654-0536
Email: brmann@utmb.edu

EDUCATION:

08/2010 - Present Ph.D. in Experimental Pathology
Department of Pathology
Graduate School of Biomedical Sciences
University of Texas Medical Branch
Galveston, TX

09/2006 - 05/2010 B.S. in Biochemistry/Molecular Biology
(*Summa cum Laude*)
Department of Biochemistry/Molecular Biology
Wittenberg University
Springfield, OH

GRANT SUPPORT:

ASM/CDC Postdoctoral Research Fellowship from the American Society for
Microbiology and the Centers for Disease Control and Prevention
“Phenotypic characterization of Babanki virus host range and vector competence”
(PI: Ann Powers from 12/31/2014 to 09/30/2016)

T32AI007526 from the National Institute of Allergy and Infectious Diseases
“Molecular determinants of attenuation in the WNV NS4B protein
C-terminal domain”
(PI: Alan D.T. Barrett from 09/01/2012 to 08/31/2014)

PREDOCTORAL RESEARCH ACTIVITIES:

08/2011 - Present

Doctoral dissertation research, UTMB, Galveston, TX
Advisor: Dr. Alan D.T. Barrett
“Genotypic and molecular determinants in the continued evolution of
West Nile virus in the United States”

05/2011 - 07/2011

Doctoral-level research rotation, UTMB, Galveston, TX
Advisors: Dr. Stephen Higgs and Dr. Saravanan Thangamani
“Effect of novel larval toxins on 2nd-instar *Aedes aegypti* mosquitoes”

03/2011 - 05/2011

Doctoral-level research rotation, UTMB, Galveston, TX
Advisor: Dr. David Beasley
“Antigenicity of dengue virus serotype IV (DENV4) ED3 mutant constructs”

01/2011 - 03/2011

Doctoral-level research rotation, UTMB, Galveston, TX
Advisor: Dr. Alan D.T. Barrett
“Cross-neutralization of a Japanese encephalitis (JE) genotype III ED3-specific
mAb to recombinant ED3 proteins of all four JE genotypes (I-IV)”

COMMITTEE RESPONSIBILITIES:

10/2012 - 10/2014	Co-President, Experimental Pathology Graduate Student Organization
05/2013 - 05/2014	President, Graduate Student Organization
05/2011 - 05/2013	Vice-President, Graduate Student Organization
06/2012 - 05/2013	Student Representative, Basic Biomedical Science Curriculum Committee
09/2010 - 05/2011	Student Representative, Graduate Student Organization: Basic Biomedical Science Curriculum

TEACHING RESPONSIBILITIES:

08/2012 - 05/2013	Krystal Fuentes (Mentee), Scientific Research and Design: Bench Tutorials Program, Ball High School
-------------------	--

AD-HOC JOURNAL REVIEWER:

09/2013 - Present	Viruses (Open Access)
-------------------	-----------------------

MEMBERSHIP IN SCIENTIFIC SOCIETIES/ PROFESSIONAL ORGANIZATIONS:

01/2012 - Present	American Society for Virology, Associate Student Member
01/2012 - Present	Sigma Xi, Associate Member

HONORS AND AWARDS:

04/2010 - Present	Presidential Scholar, University of Texas Medical Branch
11/2014	Robert Shope Endowed Scholarship (\$1000)
11/2014	Center for Tropical Diseases Graduate Student Award (\$500)
05/2014	Edward S. Reynolds Award (20th Annual Trainee Research Day, Department of Pathology, UTMB, \$500)
04/2014	James W. McLaughlin Travel Award (IHII/McLaughlin Colloquium, UTMB; \$1000)
04/2014	Sigma Xi: Excellence in Graduate Research Award (IHII/McLaughlin Colloquium, UTMB; \$200)
11/2013	Dean's Service Award
11/2013	Graduate Student Organization (GSO) Student Award
11/2013	Zhou Sisters Great Expectations Scholarship (\$1500)
11/2013	David and Janet Niesel Scholarship (\$1000)
08/2012 - 05/2013	UTMB Bromberg Scholar, High School mentorship program
05/2013	Team Science Award (Scientific Research and Design: Bench Tutorials Program, Ball High School)
04/2013	James W. McLaughlin Travel Award (IHII/McLaughlin Colloquium, UTMB; \$1200)
04/2012	American Society for Virology: Student Travel Grant (\$500)
01/2012	Who's Who Among Students in American Universities and Colleges

PUBLISHED:

A. ARTICLES IN PEER-REVIEWED JOURNALS

1. Nisha K. Duggal, Angela Bosco-Lauth, Richard A. Bowen, Sarah S. Wheeler, William K. Reisen, Todd A. Felix, **Brian R. Mann**, Daniele M. Swetnam, Alan D.T. Barrett, Aaron C. Brault. (2014) Evidence for co-evolution of West Nile virus and house sparrows in North America. PLoS Neglected Tropical Diseases 8(10): e3262.
2. Jessica A. Plante, Kristen L. Burkhalter, **Brian R. Mann**, Marvin S. Godsey, John-Paul Mutebi, David W.C. Beasley. (2014) Co-circulation of West Nile virus sequence variants, Arizona, USA, 2010. Emerging Infectious Diseases 20(2): 272-275.
3. **Brian R. Mann**, Allison R. McMullen, Daniele M. Swetnam, Vence Salvato, Martin Reyna, Hilda Guzman, Rudy Buen°C o, Jr., James A. Dennett, Robert B. Tesh, Alan D.T. Barrett. (2013) Evidence for the continued evolution of West Nile virus in the

- United States based on studies in Houston, Texas from 2002-2012. *Emerging Infectious Diseases* 19(9):1418-1427.
4. **Brian R. Mann**, Allison R. McMullen, Hilda Guzman, Robert B. Tesh, Alan D.T. Barrett. (2013) Dynamic transmission of West Nile virus across the United States-Mexican border. *Virology* 436(1):75-80.

B. REVIEWS

1. **Brian R. Mann**, Allison R. McMullen, Daniele M. Swetnam, Alan D.T. Barrett. (2013) Molecular epidemiology and evolution of West Nile virus in North America. *International Journal of Environmental Research and Public Health* 10(10):5111-29.

C. ABSTRACTS

1. **Brian R. Mann**, Steve G. Widen, Andrew Beck, Hilda Guzman, James A. Dennett, Rudy Bueno, Jr., Thomas G. Wood, Robert B. Tesh, Alan D.T. Barrett. Genetic determinants of attenuation and evolution in natural West Nile virus isolates. 23 June 2014. **[Oral Presentation]** American Society for Virology (ASV) Conference, Colorado State University, Fort Collins, Co
2. **Brian R. Mann**, Allison R. McMullen, Hilda Guzman, Robert B. Tesh, and Alan D.T. Barrett. Dynamic transmission of West Nile virus across the United States-Mexican border. 22 July 2012. **[Oral presentation]** American Society for Virology (ASV) Conference, University of Wisconsin, Madison, WI
3. **Brian R. Mann**, Steve G. Widen, Andrew Beck, Hilda Guzman, James A. Dennett, Rudy Bueno, Jr., Thomas G. Wood, Robert B. Tesh, Alan D.T. Barrett. Genetic determinants of attenuation and evolution in natural West Nile virus isolates. 11 Apr 2014. **[Poster Presentation]** IHII/McLaughlin Colloquium, The University of Texas Medical Branch, Galveston, TX
4. **Brian R. Mann**, Allison R. McMullen, Daniele M. Swetnam, Vence Salvato, Martin Reyna, Hilda Guzman, Rudy Bueno, Jr., James A. Dennett, Robert B. Tesh, Alan D.T. Barrett. West Nile virus in Houston: Back to the Future. 27 Apr 2013. **[Poster Presentation]** Keystone Symposium: Positive Strand RNA Viruses, Boston, MA
5. **Brian R. Mann**, Allison R. McMullen, Hilda Guzman, Robert B. Tesh, and Alan D.T. Barrett. Dynamic transmission of West Nile virus across the United States-Mexican border. 13 Apr 2012. **[Poster Presentation]** IHII/McLaughlin Colloquium, The University of Texas Medical Branch, Galveston, TX

PUBLICATIONS – ACCEPTED:

A. ARTICLES IN PEER-REVIEWED JOURNALS

1. Jinping Zhang, Jia Wang, Lan Pang, Guorui Xie, Thomas Welte, Vandana Saxena, Jason Wicker, **Brian R. Mann**, Lynn Soong, Alan Barrett, Willi Born, Rebecca O'Brien, Tian Wang. (2014) The co-stimulatory effects of MyD88-dependent toll-like receptor signaling on activation of murine $\gamma\delta$ T cells. *PLOS One*.