5-1-2012

Population dynamics of flaviviruses in the United States

Kendra Pesko

Kendra Pesko
Candidate

Biomedical Sciences
Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Gregory Ebel, *Co-chairperson*

Brian Hjelle, *Co-chairperson*

Kathryn Hanley

Bryce Chackerian

Antonito Panganiban

POPULATION DYNAMICS OF FLAVIVIRUSES IN THE UNITED STATES

BY

KENDRA PESKO

B.S. Biology, University of Michigan, 2003

M.S. Entomology, University of Florida, 2006

DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Biomedical Sciences

The University of New Mexico

Albuquerque, New Mexico

Spring, 2012

DEDICATION

For my wonderful husband, Dave, without whom this would not be possible.

ACKNOWLEDGEMENTS

I want to thank my committee members for their help and expertise in this endeavor: Antonito Panganiban and Bryce Chackerian for providing a fresh perspective and challenging my thinking, Kathryn Hanley for reading and reviewing all of my work and helping me develop better hypotheses and troubleshoot research problems, Brian Hjelle for stepping in as a co-mentor and most especially Gregory Ebel for all of his help – from welcoming me to the lab, to reading and revising everything with me, to supporting me as a student and researcher, and helping me develop a greater understanding of virus population biology. Also the remainder of our lab group: Doug Brackney, for many discussions and letting me use our paper on bottlenecking as part of this dissertation, Kelly Fitzpatrick, Eleanor Deardorff, Abhishek Prasad for technical help and emotional support, our undergrads, particularly Ivy Brown for making the lab a home my first two years at UNM, and Robert Nofchissey for taking care of everything all the time. Fernando Torres-Perez for teaching me phyogenetic analysis. The Broad institute for sequencing and analysis of WNV isolates. I want to thank the administrative people that make our jobs possible: Janet Kelly, Lauren Thal, and Lisa Dillon, without whom my research supplies and very stipend would not have been provided. Ignacio Ortiz, Natalie Bruner, and Mary Fenton at the BSGP office who have helped with whatever questions and aid I've needed throughout my time as a graduate student. I want to thank my parents, Bob and Ellen Pesko, for instilling a sense of curiosity about the world, teaching me how to think critically, and encouraging me in all my endeavors. And all of my friends who've made Albuquerque my favorite place to live thus far.

POPULATION DYNAMICS OF FLAVIVIRUSES IN THE UNITED STATES

BY

KENDRA PESKO

B.S. Biology, University of Michigan, 2003

M.S. Entomology, University of Florida, 2006

Ph.D. Biomedical Sciences, University of New Mexico, 2012

ABSTRACT

The genus Flavivirus comprises viral species with positive sense, single stranded RNA genomes, most of which are transmitted by mosquitoes and ticks. During replication, RNA viruses have higher mutation rates than DNA viruses, because their virally encoded RNA dependent RNA polymerase has no proofreading activity. These viruses have been shown to exist within and between hosts as swarms of closely related mutant genomes that can be examined through sequence analysis. In the work presented here we focus on the population dynamics of two zoonotic flaviviruses that cause human morbidity and mortality in the United States: tick-borne Powassan virus (POW) and mosquito-borne West Nile virus (WNV). We examine the molecular epidemiology of POW, which has increased in incidence over the past ten years, and exists as two distinct lineages in the

United States. We examine the sequence variants in WNV populations in different mosquito tissues to determine the influence of potential bottlenecks within mosquitoes on viral genetic variability. We also test influence of internally deleted WNV genomes isolated from naturally infected birds in NM on virus growth and pathogenesis, both *in vitro* and *in vivo*. The primary goal of this work is to develop a better understanding of the underlying evolutionary pressures that influence arbovirus populations and lead to disease emergence.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

The genus Flavivirus comprises viral species with positive sense, single stranded RNA genomes (Lindenbach and Rice, 2001). Flaviviruses are divided into three major groups, based on ecological and phylogenetic criteria: those vectored by mosquitoes, those vectored by ticks, and those with no known vector (Kuno et al. 1998). Flaviviruses in the mosquito vectored group include pathogens that cause major outbreaks of disease worldwide, including dengue (DENV), Japanese encephalitis (JEV), and West Nile viruses (WNV). Powassan virus (POW) is a member of the tick borne encephalitis serogroup, transmitted from the enzootic cycle to humans by ticks in the US and Russia. Several other tick borne flaviviruses have emerged as important pathogens over the past ten years, including Alkhurma virus in the Middle East and India, and Deer tick virus, a lineage of POW that is distributed in particular localities in the northern United States. Flaviviruses have been extremely successful in establishing themselves in new ecological niches and continue to produce a huge amount of infectious disease worldwide. Studies focused on understanding flavivirus population genetics can help predict and ameliorate outbreaks, allowing public health centers to better educate the public and prepare for dealing with epidemics.

WNV was introduced to the US in 1999, and rapidly spread across the nation, causing outbreaks of encephalitis (Kramer and Ebel, 2009, Kramer et al. 2008). Reported cases of POW have increased 4 fold in the past ten years. POW infection frequently leads to severe neurological disease and is fatal in approximately 10% of infected individuals,

making it one of the most clinically severe arboviral diseases transmitted in the US (Ebel, 2010, CDC, 2010). Although relying on very different hosts and vectors, these two viruses are both naturally transmitted in enzootic cycles, with periodic spillover into human populations. For both, virus has been isolated from arthropod vectors, infected vertebrates, and humans. Nucleotide sequence data from virus isolates can be used to track virus distribution, and examine evolutionary pressures that impact virus populations, such as bottlenecks within a vector, selection on genomic sites, and intrahost genetic variation. This dissertation focuses on enhancing basic understanding of flaviviral population dynamics in the US through observational and experimental approaches. First, we examine evolutionary pressures and virus demographic patterns in order to understand the current distribution and transmission cycles of POW and DTV in North America and to assess whether virus population structure has contributed to the apparent emergence of these agents as health threats. Second, we examine intrahost genetic variation to define the extent to which WNV populations undergo bottlenecks during physiological barriers traversed by virus prior to transmission. Last, we characterize naturally occurring WNV populations that contain genomes with large internal deletions and assess the role of these genomes in virus replication in cells, mosquitoes, and mice.

Powassan virus

POW (Flavivirus: Flaviviridae) is the sole North American representative of the tick-borne encephalitis (TBE) serological complex of flaviviruses. First isolated in Powassan, ON, from the brain of a fatally infected boy, it has since been isolated from ticks and/or seropositive animals in Canada, the US, and in the Primorsky region of Russia (McLean and Donohue, 1959, Artsob, 1988, Leonova et al., 2009). Deer tick virus (DTV) was first

described in the late 1990s, and is considered a distinct lineage of POW due to high nucleotide and amino acid sequence identity (84 and 93 percent, respectively), and strong serological cross-reactivity (Kuno et al., 2001, Beasley et al., 2001, Telford et al., 1997). POW and DTV are perpetuated in different transmission cycles, with the former maintained by *Ixodes cookei* and/or *Ix. marxi* and medium sized mammals, such as groundhogs, and the latter maintained by *Ix. scapularis* and white-footed mice (Artsob, 1988, Ebel et al., 2000). In the last ten years, POW encephalitis incidence has increased by 4 fold, from 0.7 cases per year from 1958 to 1998, to 2.8 cases per year from 2001 to 2010 (Hinten et al., 2008, CDC 2010). A recent fatality that occurred in New York state was unequivocally attributed to DTV (Tavakoli et al., 2009). Through its association with the aggressively human-biting deer tick, DTV has been considered a more significant public health threat compared with POW because POW is vectored mainly by the nidiculous *Ix. cookei*. Tick-transmitted flaviviruses may therefore be an emerging public health issue in regions such as Wisconsin, Minnesota, and New York where deer ticks are present.

In chapter one, we sought to determine whether the apparent emergence of tick-borne flavivirus disease in North America may be driven by increases in the virus population size, reasoning that the well-documented expansion in the population of deer ticks that has occurred in recent decades may have consequently increased the population size(s) of their pathogens. In particular, full genome sequences were determined and phylogenetic relationships of POW and DTV from North America and Russia reconstructed. We then estimated the divergence time between lineages, and tested for evidence of positive selection. Finally, population dynamics were inferred under a

Bayesian framework, shedding light on the mechanisms influencing the epidemiology of POW in North America. Our results demonstrated a stable population, with no recent increases in genetic variation, indicating the increase in human incidence was caused by other factors. Estimates for the most recent common ancestor for each lineage was around 200 years before present, with neither one nor the other being ancestral to the other, and both lineages sharing a common ancestor at around 500 years before present. Strong evidence for purifying selection was present in all alignments tested, lending further support for the stability of these populations.

Vector competence and viral population bottlenecks

Transmission of arboviruses is dependent on many environmental, host, and vector factors. When an infectious blood meal is ingested by a mosquito, the infecting viral population must traverse a number of barriers before transmission can take place: it infects the midgut and replicates in midgut epithelial cells, then disseminates to peripheral tissues such as the fat bodies where further replication takes place before infection of the salivary glands and, ultimately, transmission to another host. Studies with a fluorescently labeled Venezuelan equine encephalitis virus (VEEV) infectious clone in *Aedes taeniorynchus* have shown that a limited number of cells are susceptible to infection immediately after the blood meal has been taken, which may represent a bottleneck for viral populations (Smith et al. 2008). It is assumed that of the viral population initially imbibed, only a small number of individual clones may proceed through each physiological barrier during mosquito infection (especially midgut infection and escape barriers, salivary gland infection barrier), so the mosquito vector may reduce viral genetic diversity. Overall, arboviruses are thought to have diminished substitution

rates compared to other RNA viruses due to the diverse requirements of replicating in vertebrate and invertebrate cells, which provides a strong purifying pressure on viral populations (Jenkins et al. 2002). However, nonconsensus WNV variants have been isolated from different birds within the same transmission foci (Jerzak et al. 2005).The influence of vector physiological barriers on virus populations.

In the second chapter, we tracked a genetically diverse virus population from blood meal through saliva in *Culex pipiens quinquefasciastus* mosquitoes in order to determine the impact of physiological barriers on virus genetic diversity. We fed mosquitoes with a highly genetically diverse population of WNV generated by mixing 24 different isolates together at equal ratios, and sampled the viral sequences present in three infected tissues (midgut, legs, and salivary glands) from three mosquitoes at three timepoints (7, 14, and 21 days post infection), plus three mosquitoes immediately following the blood meal. By comparing sequence variants in each tissue, we tracked the movement of the specific virus genotypes and determined the impact of these barriers on overall viral genetic diversity over time. We utilized a technique that has been used previously to look at intrahost population dynamics in human immunodeficiency virus (HIV) infected patients by mapping tissues where individual clones were found onto a Bayesian analysis of virus population in each mosquito (Salemi et al. 2007). Surprisingly, although the viral genetic diversity was reduced from input to midguts, we found the salivary glands had the most genetically diverse population of viruses, including many individual clones that had been present in the input, but not represented in the midgut. Overall, our results show that the transmission "barriers" that are associated with

mosquito transmission do not seem to create population bottlenecks (i.e. genetic barriers). Overall, these studies suggest that founder effects in mosquitoes are relatively weak.

Defective genomes' impact on virus infection

One effect of the high error rate in RNA polymerases is the production of viral mutations that are deleterious. Viral variants with deleterious mutations occur readily in cell culture when virus is propagated at high multiplicities of infection, and are known as defective interfering particles (DIPs), because their presence can interfere with the production of full length homologous virus (Thompson et al. 2010). These DIPs have large internal deletions to their genomes that must be complemented by full length virus, and the interference observed may be due to resource competition (Kim and Yin, 2004). Deleterious flaviviral genomes have also been identified *in vivo* for dengue virus (DENV): a premature stop codon in the envelope coding region was been found in isolates taken from mosquitoes and humans in Myanmar over a two year period (Aaskov et al. 2006). Using phylogenetic analysis, this study determined the deleterious genomes may have been complemented *in vivo* through multiple transmission cycles. The impact of such deleterious genomes *in vivo* has not been characterized.

In chapter three we describe internally deleted genomes identified in samples of WNV taken from lorikeets that died in the NM zoo in 2005 and 2007. These genomes contained deletions that spanned from the premembrane to the first nonstructural protein coding region, occurred in frame, but began and ended at different sites. Furthermore, phylogenetic analysis showed the isolates did not form a monophyletic clade when compared with other isolates from the southwest region. We assayed virus production *in vitro* in Vero cells and observed internally deleted mutants interfering with full length

virus production. To measure the impact of these deleterious mutants on the WNV transmission cycle *in vivo*, we infected mosquitoes and mice with one of the isolates. Deletion mutant could be detected in bodies but not legs or saliva from infected mosquitoes, indicating a lack of transmissibility and inefficient complementation. Deletion mutant presence did diminish infection rates and body titers in mosquitoes infected with this isolate. We did not detect the presence of deletion mutant after infecting two different mouse strains, and saw no attenuation of morbidity or mortality. Deletion mutant was recovered from mosquito bodies but not from mice, indicating mosquito cells may be more permissible to coinfection than vertebrate cells. Overall, we hypothesize these deletion mutants arose independently in each lineage and probably do not substantially shape the endemic WNV transmission cycle.

References

Artsob, H. (1988). Powassan encephalitis. In The Arboviruses, Monath, T.P. (Ed.), CRC Press, Boca Raton, FL, pp. 29-50.

Beasley, D.W.C., Suderman, M.T., Holbrook, M.R., Barrett, A.D.T. (2001). Nucleotide Sequencing and Serological Evidence that the Recently Recognized Deer Tick Virus is a Genotype of Powassan Virus. Virus Res. 79, 81-89.

CDC. (2010). Powassan virus neuroinvasive disease cases reported by state, 2001-2010.

Ebel, G.D. (2010). Update on Powassan Virus: Emergence of a North American Tick-Borne Flavivirus. Annu.Rev.Entomol. 55, 95-110.

Ebel, G.D., Kramer, L.D. (2009). West Nile encephalitis virus infection. In: Diamond, M.S. (Ed.), West nile virus: Molecular epidemiology and diversity. Springer, pp. 25-43.

Ebel, G.D., Campbell, E.N., Goethert, H.K., Spielman, A., Telford III, S.R. (2000). Enzootic Transmission of Deer Tick Virus in New England and Wisconsin Sites. Am J Trop Med Hyg 63, 36-42.

Hinten, S.R., Beckett, G.A., Gensheimer, K.F., Pritchard, E., Courtney, T.M., Sears, S.D., Woytowicz, J.M., Preston, D.G., Smith, R.P.,Jr, Rand, P.W., Lacombe, E.H., Holman, M.S., Lubelczyk, C.B., Kelso, P.T., Beelen, A.P., Stobierski, M.G., Sotir, M.J., Wong, S., Ebel, G., Kosoy, O., Piesman, J., Campbell, G.L., Marfin, A.A.(2008). Increased Recognition of Powassan Encephalitis in the United States, 1999-2005. Vector Borne Zoonotic Dis. 8, 733-740.

Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C. (2002). Rates of Molecular Evolution in RNA Viruses: A Quantitative Phylogenetic Analysis. J.Mol.Evol. 54, 156-165.

Jerzak, G., Bernard, K.A., Kramer, L.D., Ebel, G.D. (2005). Genetic Variation in West Nile Virus from Naturally Infected Mosquitoes and Birds Suggests Quasispecies Structure and Strong Purifying Selection. J Gen Virol 86, 2175-2183.

Kim, H., Yin, J. (2004). Quantitative Analysis of a Parasitic Antiviral Strategy. Antimicrob.Agents Chemother. 48, 1017.

Kramer, L.D., Styer, L.M., Ebel, G.D. (2008). A Global Perspective on the Epidemiology of West Nile Virus. Annu. Rev. Entomol. 63, 81.

Kuno, G., Chang, G.J.J., Tsuchiya , K.R., Karabatsos, N., Cropp, C.B. (1998). Phylogeny of the Genus Flavivirus. J Virol 72, 73-83.

Kuno, G., Artsob, H., Karabatsos, N., Tsuchiya, K.R., Chang, G.J. (2001). Genomic Sequencing of Deer Tick Virus and Phylogeny of Powassan-Related Viruses of North America. Am.J.Trop.Med.Hyg. 65, 671-676.

Leonova, G.N., Kondratov, I.G., Ternovoi, V.A., Romanova, E.V., Protopopova, E.V., Chausov, E.V., Pavlenko, E.V., Ryabchikova, E.I., Belikov, S.I., Loktev, V.B. (2009). Characterization of Powassan Viruses from Far Eastern Russia. Arch.Virol. 154, 811-820.

Lindenbach, B., Rice, C. (2001). Flaviviridae: The Viruses and their Replication. In Fields Virology, 1, 991-1041.

McLean, D.M., Donohue, W.L. (1959). Powassan Virus: Isolation of Virus from a Fatal Case of Encephalitis. Canadian Medical Association Journal 80, 708.

Salemi, M., Burkhardt, B.R., Gray, R.R., Ghaffari, G., Sleasman, J.W., Goodenow, M.M. (2007). Phylodynamics of HIV-1 in Lymphoid and Non-Lymphoid Tissues Reveals a Central Role for the Thymus in Emergence of CXCR4-using Quasispecies. PLoS One 2, 950.

Tavakoli, N.P., Wang, H., Dupuis, M., Hull, R., Ebel, G.D., Gilmore, E.J., Faust, P.L. (2009). Fatal Case of Deer Tick Virus Encephalitis. N.Engl.J.Med. 360, 2099-2107.

Thompson, K.A.S., Rempala, G.A., Yin, J. (2009). Multiple-Hit Inhibition of Infection by Defective Interfering Particles. J.Gen.Virol. 90, 888-899.

CHAPTER 2: MOLECULAR EPIDEMIOLOGY OF POWASSAN VIRUS IN NORTH AMERICA

Pesko, K.N., Torres-Perez, F, Hjelle, B.J., and G.D. Ebel. 2010. Molecular epidemiology of Powassan virus in North America. Journal of General Virology, 91 (11): 2698-2705.

MOLECULAR EPIDEMIOLOGY OF POWASSAN VIRUS IN NORTH AMERICA

KENDRA N. PESKO[1], FERNANDO TORRES-PEREZ [1,2], BRIAN L. HJELLE [1] AND GREGORY D. EBEL[1]

[1]Department of Pathology, University of New Mexico School of Medicine, Albuquerque, NM

[2]Museum of Southwestern Biology, University of New Mexico, Albuquerque, NM

Running head: Powassan virus in North America

Address for Correspondence:

Gregory D. Ebel

gebel@salud.unm.edu

The Genbank accession numbers for the sequences reported in this manuscript are: HM440558- HM440563.

Abstract

Powassan virus (POW) is a tick borne flavivirus distributed in Canada, the northern United States, and the Primorsky region of Russia. POW is the only tick-borne flavivirus endemic to the western hemisphere, where it is transmitted mainly between *Ixodes cookei*, and groundhogs (*Marmota monax*). Deer tick virus (DTV), a genotype of POW that has been frequently isolated from deer ticks (*Ix. scapularis*), appears to be maintained in an enzootic cycle between these ticks and white-footed mice (*Peromyscus leucopus*). DTV has been isolated from ticks in several regions of North America, including the upper Midwest and the eastern seaboard. The incidence of human disease due to POW is apparently increasing. Previous analyses of tick-borne flaviviruses endemic to North America have been limited to relatively short genome fragments. We therefore assessed the evolutionary dynamics of POW using newly generated complete and partial sequences. Maximum likelihood and Bayesian phylogenetic inferences showed two well supported, reciprocally monophyletic lineages corresponding to POW and DTV. Bayesian skyline plots based on year-of-sampling data indicate no significant population size change for both virus lineages. Statistical model-based selection analyses showed evidence of purifying selection in both lineages. Positive selection was detected in NS-5 sequences for both lineages, and envelope sequences for POW. Our findings confirm that POW and DTV sequences are relatively stable through time, suggesting strong evolutionary constraint, and support field observations suggesting that tick-borne flavivirus populations are extremely stable in enzootic foci.

Introduction

Powassan virus (POW, Flaviviridae: Flavivirus) is the sole North American representative of the tick-borne encephalitis (TBE) serological complex of the flaviviruses. POW is comprised of two genetic lineages including the prototype lineage (POW, Lineage I) and a second lineage that was first described in the late 1990s (Deer Tick virus (DTV), Lineage II) (Telford et al., 1997). POW and DTV appear to perpetuate in nature in relatively distinct transmission cycles, with the former maintained by *Ixodes cookei* and/or *Ix. marxi* and medium sized mammals, and the latter maintained by *Ix. scapularis* and white-footed mice (Artsob, 1988, Ebel et al., 2000). Whereas POW has been considered to be a minor public health concern due to relative host-specificity of its arthropod vector, DTV could be a greater threat since it is mainly associated with the tick vectors that have driven the ongoing epidemics of Lyme disease, human babesiosis, and human granulocytic anaplasmosis (Rodgers and Mather, 2007, Daniels et al., 1998, Falco et al., 1995). DTV has been isolated from ticks in several regions of North America (Brackney et al., 2008, Ebel et al., 1999, Telford et al., 1997), and is considered a distinct lineage of POW due to high nucleotide and amino acid sequence identity (84 and 93 percent, respectively) (Kuno et al., 2001), and strong serological cross-reactivity (Beasley et al., 2001). Powassan virus has been isolated from ticks and/or seropositive animals in Canada, the US, and in the Primorsky region of Russia (Artsob, 1988, Leonova et al., 2009). Although data from field-based studies of POW and DTV are scant, two recent reports demonstrated that infection rates in ticks in a DTV focus in Northern Wisconsin were remarkably stable over a ten year period (Brackney et al., 2008, Ebel et al., 1999) and that infections within ticks were surprisingly genetically homogeneous, exhibiting

little or no quasispecies structure (Brackney et al., 2010). Therefore, POW and DTV appear to be similar to many other tick-borne agents in that they are nidal infections, maintained in highly stable enzootic foci for long periods of time. Paradoxically, human disease due to infection by POW and/or DTV appears to be increasing. Human cases have increased markedly between the period 1958-1998 and 1999-2007 (Hinten et al., 2008) and a recent fatality was unequivocally attributed to DTV (Tavakoli et al., 2009). These observations raise the possibility that these agents may be an emerging burden to public health in regions of North America where they are enzootic. Accordingly, we sought to determine whether the apparent emergence of tick-borne flavivirus disease in North America may be driven by increases in the virus population size, reasoning that the well-documented expansion in the population of deer ticks that has occurred in recent decades may have consequently increased the population size(s) of their pathogens. In particular, full genome sequences were determined and phylogenetic relationships of POW and DTV from North America and Russia reconstructed. We then estimated the divergence time between lineages, and tested for evidence of positive selection. Finally, population dynamics were inferred under a Bayesian framework, shedding light on the mechanisms influencing the epidemiology of POW in North America.

Results

Phylogenetic analyses. To assess phylogenetic relationships of POW virus, we used 11 full genome, and 23 envelope and 22 NS-5 viral sequences. Kyasanur forest disease virus (KFDV) and Alkhurma virus (AHV) were included as outgroups. Maximum likelihood topologies were congruent in full genome, NS-5, and envelope partitions in recovering

two supported monophyletic groups which correspond to lineage I POW and DTV (lineage II POW) (Fig. 1 A, B). Full genome topology showed that within lineage I POW, samples were subdivided into Russia - Ontario (Maximum likelihood bootstrap (MLB) =100), and the sample from New York was basal. DTV showed two subclades clustering all samples from Wisconsin (MLB=99), and Massachusetts with Connecticut (MLB=100). Envelope and NS-5 topologies supported these divisions, with DTV from CO being basal to the Wisconsin strains, and the isolate of DTV from WV basal to New England strains.

Substitution rates, divergence times, and population size changes. Bayesian Skyline plots (BSP) were generated for envelope (Fig 2A) and NS-5 coding sequences (Fig 2B), using relaxed (uncorrelated log normal) molecular clock models. Overall, results indicate that the effective number of POW and DTV infections through time was stable over the distant past (years 1950 – 1100) with recent declines, apparently beginning around 1950. BSPs for envelope and NS-5 partitions were congruent in placing the time to the most recent common ancestor (MRCA) for DTV and POW around 1000 years before present (ybp) (Fig 2). Estimated nucleotide substitution rates using envelope sequences were 2.2 x 10^{-4} substitutions/site/year, and 3.9 x 10^{-5} substitutions/site/year using NS-5 sequences.

Using concatenated envelope and NS-5 sequences, we estimated divergence times of POW and DTV lineages under a Bayesian Markov Chain Monte Carlo (BMCMC) framework (Fig. 3). Time to the MRCA of DTV isolates was estimated at 195 ybp (highest probability density [HPD]: 73-382), and for lineage I POW, 196 ybp (HPD: 67-

382); the analysis placed the time for the MRCA between the two lineages around 484 ybp (HPD: 114-1048).

Selection analyses in POW strains. Sliding window analysis of dN/dS and π conducted with complete genomic coding sequences of both POW and DTV indicated purifying selection throughout the genome, with values distributed below 0.3 for dN/dS throughout (Fig 4). Likelihood-based approaches mostly agreed in showing that most substitutions were silent and likely reflect strong purifying selection acting over POW genome. However, the SLAC, REL, and PARRIS models detected purifying selection throughout the sequences, while the FEL model found evidence for positive selection at two different sites in two different NS-5 partitions (one in POW lineage I only, the other evident in comparisons of the two lineages) (Table 2). Using the Selecton server we found evidence for positive selection in the envelope and NS-5 sequences using the MEC model and in the DTV NS-5 sequences using the M8 model (Table 2). McDonald-Kreitman tests showed significant values for partitions of full genome sequences (p=0.035) and NS-5 sequences (p=0.001), but did not conclude significant difference from neutral evolution for envelope sequences (p=0.077) (Table 2).

Discussion

Our phylogenetic analyses using both partial sequences and whole genome sequences found support for two distinct lineages of POW, similar to previous findings (Ebel et al., 2001, Kuno et al., 2001). The clustering of Russian samples with the

prototypic LB strain from Ontario, 1958 supports assertions that a strain similar to this was recently imported into Russia, most likely in the past 100 years, potentially from importation of mink for fur farms in the area (Leonova et al. 2009). Envelope and NS-5 topologies showed DTV from CO basal to the Wisconsin strains, and the isolate of DTV from WV basal to New England strains. Isolates taken from overlapping geographical areas in both Ontario and New England were placed into the two separate lineages (POW and DTV) suggesting that these two partially-sympatric lineages exploit different ecological niches, which supports the association of each lineage with a distinct transmission cycle (Ebel et al., 1999).

Bayesian skyline plots (BSPs), which measure population dynamics based on genetic diversity, show a long, stable history for both POW and DTV, with some evidence of recent declines in populations. These results support field data that suggests a stable, persistent virus population is present in the upper Midwest (Brackney et al., 2008, Ebel et al., 2000). Although the BSPs show a recent decline in populations of POW virus, an increase in human cases of POW encephalitis has been reported over the past ten years (Hinten et al., 2008). This may be the result of more human contact with areas where the virus is present (Lloyd-Smith et al., 2009). In addition or alternatively, there could be an enhanced awareness of arthropod borne flaviviruses due to the introduction of WNV to the US in 1999, which lead to increased testing and better diagnosis (Hinten et al., 2008). Genetic diversity is very low within tick hosts (Brackney et al, 2010) especially compared to mosquito-borne flaviviruses, where intrahost populations have greater genetic diversity (Jerzak et al., 2005). The lack of diversity within POW populations may limit the available number of unique genomes and consequently the

adaptability of POW virus to new ecological niches. Ultimately this could produce the patterns of strong genetic conservation and population stability that we document in this study. Other factors that may influence population size as estimated by genetic diversity could be the low levels of replication achieved in ticks, resulting in smaller populations generated at this stage of the life cycle, transstadial or transovarial transmission that may occur as part of the viral life cycle, providing a bottleneck for virus transmission, and the potential for cofeeding transmission which would bypass the need for replication in vertebrate hosts from the population dynamics of the virus overall (Ebel and Kramer, 2004, Nuttall et al., 1994). Finally, our sampling effort might not be sufficient to detect very recent (i.e. in the past 30 yrs) changes in virus population size that would be associated with the well-documented expansion in the populations of their main vector, *Ix. scapularis*. In sum, our results are concordant with previously published data suggesting that populations of POW (and DTV), like tick-borne encephalitis and other tick-borne pathogens, are very stable in nature (Blaskovic and Nosek, 1972, Goethert and Telford III, 2009, Gresikova and Calisher, 1989).

Estimated nucleotide substitution rates ranged from 2×10^{-4} to 4×10^{-5} subs/site/yr, with NS-5-derived estimates approximately one log10 lower than those derived from E sequences. Nucleotide substitution estimates for the segmented negative sense RNA virus transmitted by ticks, Crimean Congo hemorrhagic fever (CCHF; Bunyaviridae: *Nairovirus*), found a rate of 5.8×10^{-5} for the L segment, which encodes the RNA dependent RNA polymerase, and 1.09×10^{-4} for the S segment, which encodes the structural nucleocapsid protein, although authors concluded this could be a result of more taxon sampling for S segment, which is not true of our data set (Carroll et al.,

2010). Nucleotide substitution rates for other vector-borne flaviviruses, such as KFDV and TBEV, tick borne flaviviruses related to POW, or the mosquito borne flaviviruses yellow fever virus (YFV), SLEV, and WNV fall within this range (Baillie et al., 2008, Bryant et al., 2007, Mehla et al., 2009, Snapinn et al., 2007). These estimates of nucleotide substitution rates for POW and DTV are both lower than the average of ~10^{-3} subs/site/yr estimate for RNA viruses (Jenkins et al. 2002). However, vector borne RNA viruses have been found overall to display lower mutation rates than other RNA viruses, most likely because of the requirement for replication in two distinct cell types (vertebrate and invertebrate) which may lead to stronger evolutionary constraint and lower overall accumulation of mutant types (Jenkins et al., 2002). Although the mutation rates are lower than expected for RNA viruses, they are 20-fold higher than those found in a recent study of POW using only the Russian and LB sequence data (Leonova et al., 2009). This difference may be due to a restricted focus of Russian transmission, or that the strain of POW in Russia is closely related to the LB strain because it was only recently introduced there from North America, leading to lower genetic diversity in those samples (Leonova et al., 2009). The apparently non-native Russian focus may prove a useful laboratory, over time, for studying the actual rate of divergence in a new viral focus in nature, and provide an interesting contrast to literature concerning the similar importation of the mosquito-borne WNV into North America (Ebel and Kramer, 2009, Beasley et al., 2003, Bertolotti et al., 2007). In any case, our data show that nucleotide substitution rates for POW are similar to other flaviviruses and variable along the length of the genome.

Divergence times were then estimated to assess the likelihood that DTV and POW diverged recently as a result of expanding deer tick populations. Using the BMCMC approach, and concatenated complete NS-5 and envelope sequences (around 4000 nt total), we estimated the MRCA for POW from DTV at around 484 years ago (95% HPD=114-1048). This estimate globally agrees with previous studies based on an approximately 2000 nt sequence comprising portions of the envelope, NS-3, and NS-5 coding sequences (Charrel et al., 2005). Previously, greater genetic diversity was found within the DTV lineage, leading researchers to speculate that this may be the older lineage (Ebel et al., 2001). Our divergence estimates show a nearly simultaneous radiation of DTV and POW at around 200 years ago, with no indication that one clade is ancestral to the other. Divergence of DTV and POW may have occurred well before the current expansion of deer tick populations, and must have been driven by other factors.

We observed strong purifying selection, and some evidence of positive selection in the sequences we studied. A signal of positive selection has previously been detected in partial POW envelope alignments, but not for NS-5 alignments, using a McDonald-Kreitman (M-K) test (Ebel et al., 2001). We found significant signals for positive selection occurring in full genome sequences and NS-5 coding sequences using M-K tests, and approaching significance for envelope coding sequences. The likelihood models we employed were also able to identify positive selection in certain sites. However, we were unable to consistently detect specific codons that appeared to be subject to positive selection using multiple programs. All methods which were capable of doing so detected purifying selection at multiple sites. Our results indicate that most substitutions in POW and DTV genomes were synonymous, indicating that strong

purifying selection limits the accumulation of genetic changes in POW virus populations. Similar results have been found in other arthropod borne flaviviruses (Baillie et al., 2008). Despite some incongruence in the methods identifying positive selection, we suggest that this process is acting on POW and DTV genomes.

Methods

Virus strains. RNA samples were generated from three isolates of ticks collected in Wisconsin in 2008 (Brackney et al., 2008), one isolate from the Nantucket field station in 2000 (Ebel et al., 2000), cf9901 from Wisconsin, 1999 (Ebel et al., 1999), and POWV 64-7062 (Whitney and Jamnback, 1965) (Table 1).

RNA extraction, RT-PCR, and sequencing. RNA was extracted using the RNeasy Protect Mini-Kit according to the manufacturer's protocol (Qiagen, Valencia, CA) from 100 µl of viral stock and eluted in 100 µl of ddH_2O. RT-PCR was conducted using Superscript III One-Step Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) with Platinum Taq (Life Technologies, Carlsbad, CA) using the following parameters: 55 °C for 30 min (reverse transcription), 95°C for 15 min (initial denaturation), and 40 cycles of 94 °C for 15 s, 55 °C for 30 s, and 72°C for 30 s, followed by final extension at 72°C for 10 min. Primers amplifying approximately 1 kb segments were used, sequences available on request. PCR products were purified using Strataprep PCR purification kits (Stratagene, Agilent Technologies, Garden Grove, CA) and sequenced using the primers used for RT-PCR for a 2x plus coverage by the University of New Mexico DNA Research Services. Sequences were assembled using SeqMan (DNAstar, Madison, WI), and aligned with ClustalW in Bioedit (Hall, 1999). The full genome sequences from the POW and DTV

strains sequenced for this study have been deposited in GenBank and have been assigned accession numbers HM440558 to HM440563. Other previously published sequences used in the study were obtained from GenBank. Sequences were tested for evidence of recombination using GARD (Genetic Algorithm for Recombination Assessment) available in the datamonkey.org website.

Phylogenetic analysis. Maximum likelihood trees were generated in Paup* (Swofford, 2002) considering all characters as unordered with four possible states (A,C,T,G). We employed heuristic searches with 10 random stepwise additions of sequences and tree bisection–reconnection (TBR) branch swapping. Node support was evaluated with 1000 non-parametric bootstrap pseudoreplicates (Felsenstein, 1985), using the same searching conditions as described above. Analysis was performed with 11 full genome sequences.

Divergence times, substitution rates, and population size changes. Bayesian phylogenetic analyses were performed to investigate the time-scale of diversification of POWV and DTV lineages using BEAST v. 1.5.1 (Drummond and Rambaut, 2007). For these analyses, 24 sequences of partial sequences including ENV and NS-5 (4000-nt) were included. To assess the extent of temporal structure in the sequence data, we performed a regression analysis of tree root-to-tip genetic distance against sampling date using the program Path-O-Gen (http://tree.bio.ed.ac.uk/software/pathogen/) based on a Neighbor Joining tree. Regression analysis showed sufficient temporal structure in the viral sequences of this study (range of 1958 to 2008) to estimate rates and dates. Analyses were performed comparing uniform rates across branches (strict clock) and uncorrelated

relaxed clock assumptions. We used the uncorrelated log normal assumption after comparing the marginal posterior distribution with the prior distribution of the standard deviation of the uncorrelated log normal distribution (Suchard et al., 2001).

The HKY model of nucleotide substitution was used with gamma rate heterogeneity among sites as selected by comparison of Akaike Information Criterion scores. Two independent MCMC chains were run for 50,000,000 steps with sampling every 5000 steps, following a discarded burn-in of 5,000,000 steps of the posterior samples. The samples from the two runs were combined, and the convergence of the chain, sampling and mixing was confirmed by inspection of the MCMC samples using the program Tracer v. 1.5 (Rambaut and Drummond, 2007).

Selection and recombination detection. Alignments of POW sequences were tested for possible breakpoints using GARD (Genetic Algorithm Recombination Detection) (Pond et al., 2006), available on the Datamonkey web-based gateway (Pond and Frost, 2005a). Models of nucleotide substitution were calculated using Modeltest (Posada and Crandall, 1998). Non-synonymous to synonymous rate ratios were estimated in HyPhy, using the maximum likelihood phylogeny and the GTR model of nucleotide substitution. Tests for positive selection were conducted on the datamonkey server using the following methods: single-likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), random effects likelihood (REL), and PARRIS (a PARtitioning approach for Robust Inference of Selection). SLAC counts the number of nonsynonymous changes per nonsynonymous sites (dN) and tests whether it is significantly different from the number of synonymous changes per synonymous site (dS). FEL and REL estimate ratios of

nonsynonymous to synonymous for each site in an alignment, but FEL fixes estimates of branch lengths and substitution rate parameters for each site independently, while REL allows rate variation in both synonymous and nonsynonymous rates and the nucleotide substitution model (Pond and Frost, 2005b). PARRIS is able to take both recombination and differences in synonymous evolution rate into account (Scheffler et al., 2006). Additionally, we tested for selection using three models from the Selecton server: MEC, M7+β, and M8. DNAsp (Librado and Rozas, 2009) was used to calculate McDonald-Kreitman tests, the ratio of the mean number of nonsynonymous (dN) to synonymous (dS) nucleotide substitutions (dN/dS), and the genetic diversity (π). Analyses were performed using full coding sequences.

Acknowledgements:

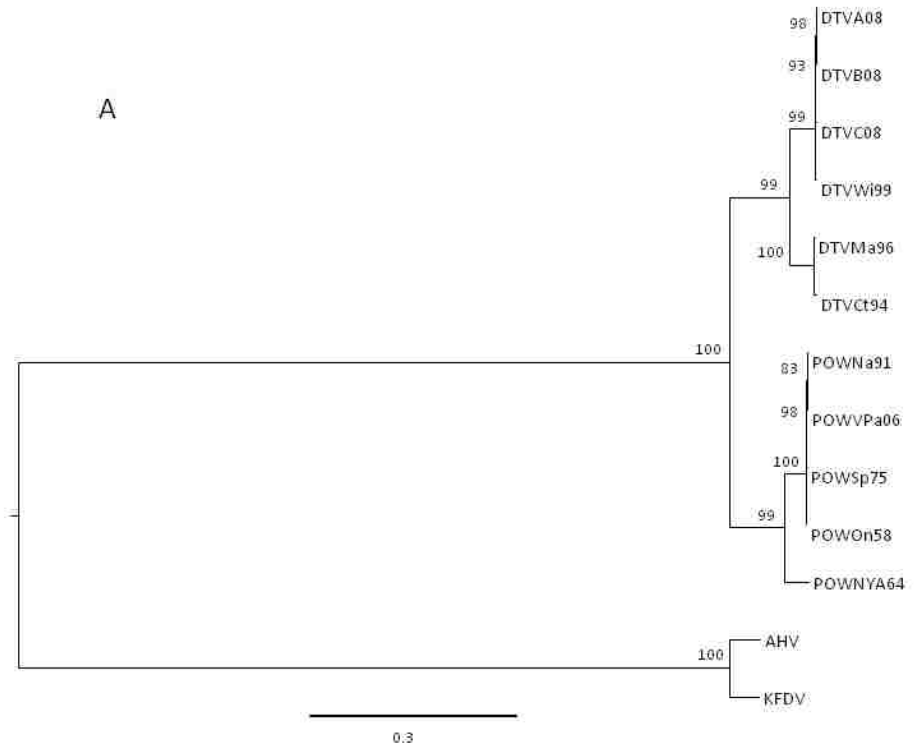
This work was supported by the University of New Mexico, school of medicine, department of Pathology. F.T.P was supported by the Fogarty Actions for Building Capacity Award 5D43 TW01133.

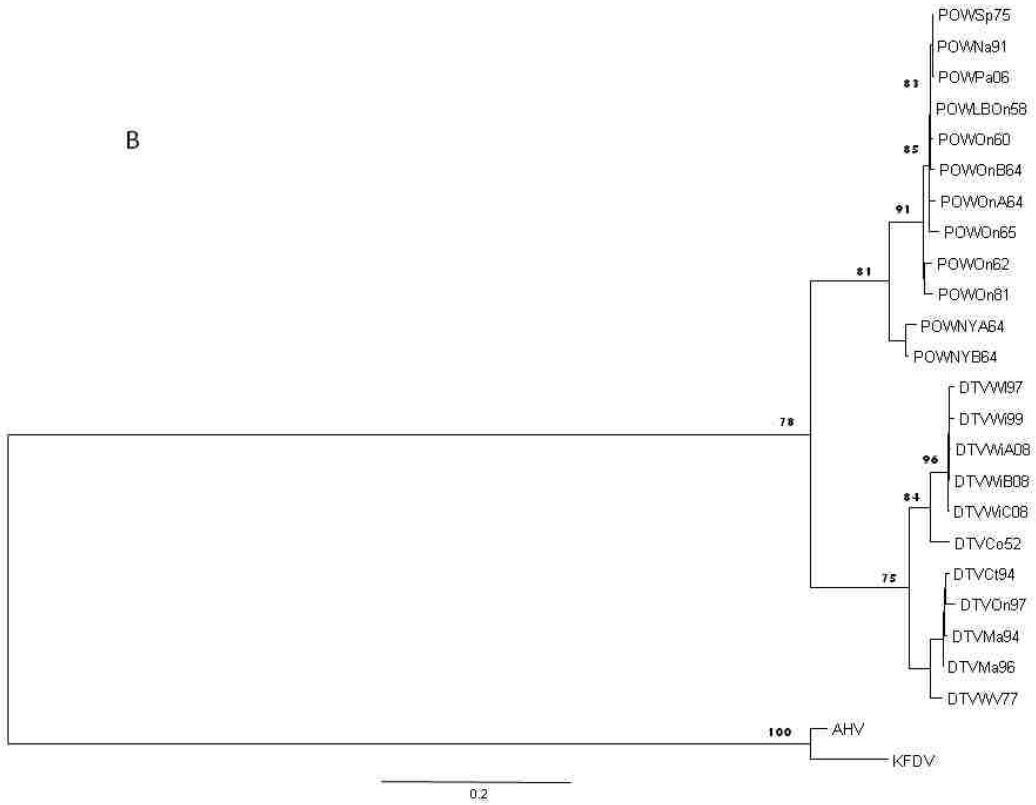
Code	Strain	Collection location	Year	Source	Accession numbers
DTVWiA08		Spooner, WI	2008	*Ixodes scapularis*	HM440560
DTVWiB08		Spooner, WI	2008	*I. scapularis*	HM440561
DTVWiC08		Spooner, WI	2008	*I. scapularis*	HM440562
POWPa06	Part/2006	Partizansk, Russia	2006	Human	EU543649
DTVWi99	wicf9901	Chippewa Falls, WI	1999	*I. scapularis*	HM440558
DTVWi97	DTV-SPO	Spooner, WI	1997	*I. scapularis*	AF310921, AF310938
DTVMa96	NFS001	Nantucket, MA	1996	*I. scapularis*	AF310918, AF310947, HM440559
DTVCt94	DTV-CT	Connecticut	1994	*I. scapularis*	AF310919, AF311056
DTVMa94	DTV-IPS	Ipswich, MA	1994	*I. scapularis*	AF310947, AF310918
POWNad91	Nad-1991	Nadezdinsk, Russia	1991	Human	EU670438
POWOn81	t18-23-81	Ontario, Canada	1981	*I. cookei* on *Marmota* sp.	AF310909, AF310943
DTVWV77	12542	West Virginia	1977	*Vulpes* spp.	AF310920, AF310949
POWSp75	Sp-9	Spassk, Russia	1975	*Dermacentor silvarum*	EU770575
POWOn65	m11665	Ontario, Canada	1965	Tick	AF310910, AF310937
POWNYA64	64-7062	New York	1964	Ticks on *Marmota* sp.	AF310915, AF310944, HM440563
POWNYB64	64-7483	New York	1964	Unknown	AF310916, AF310945
POWOn64A	M8998	Ontario, Canada	1964	Unknown	AF310911, no NS-5
POWOn64B	1982-64	Ontario, Canada	1964	*M. monax*	AF310913, AF310939
POWOn62	1247-62	Ontario, Canada	1962	*T. hudsonicus*	AF310914, AF310942
POWOn60s	m1409	Ontario, Canada	1960s	Tick	AF310912, AF310940
POWOn58	LB	Powassan, ON	1958	Human brain	NC003687
DTVCo52	791A-52	Colorado	1952	*Dermacentor andersonii*	AF310922, AF310950

10 Table 1: Isolates of POW used in this study

Model			**Datamonkey**		**Selecton**			**M-K test**
Partition	SLAC	FEL	REL	PARRIS	MEC (vsM8a)	M8 (vsM8a)	M7	Neutrality index (p value)
NS5 POWV	Purifying	**Positive (735)**	Purifying	No +	**Significant**	NS	No +	
NS5 DTV	Purifying	Purifying	Purifying	No +	NS	**Significant**	No +	
NS5 All	Purifying	**Positive (383)**	Purifying	No +	**Significant**	NS	No+	2.089 (0.00107)
ENV POWV	Purifying	Purifying	Purifying	No +	**Significant**	NS	No +	
ENV DTV	Purifying	Purifying	Purifying	No +	NS	NS	No +	
ENV All	Purifying	Purifying	Purifying	No +	**Significant**	NS	No+	1.82 (0.0753)
POWV CDS	Purifying	Purifying	Purifying	No +	NS	NS	No +	
DTV CDS	Purifying	Purifying	Purifying	No +	NS	NS	No +	
CDS All	Purifying	Purifying	Purifying	No +	NS	No +	No+	0.753 (0.03539)

Table 2: Results from selection analyses





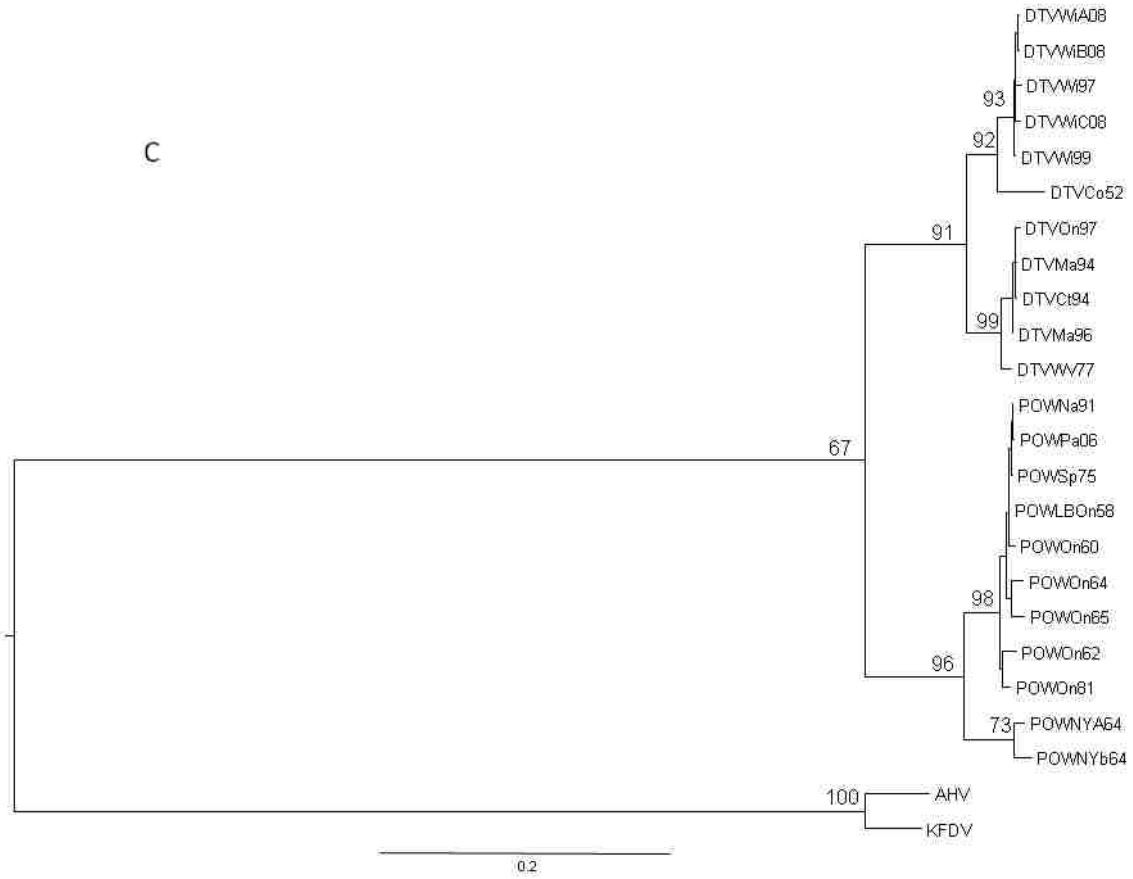


Figure 1**:** Maximum likelihood tree obtained from (A) full coding sequences (10245-nt), (B) envelope (1491-nt), and (C) NS-5 (2709-nt) partial sequences of POW and DTV strains. Numbers above branches indicate pseudoreplicate bootstrap values.

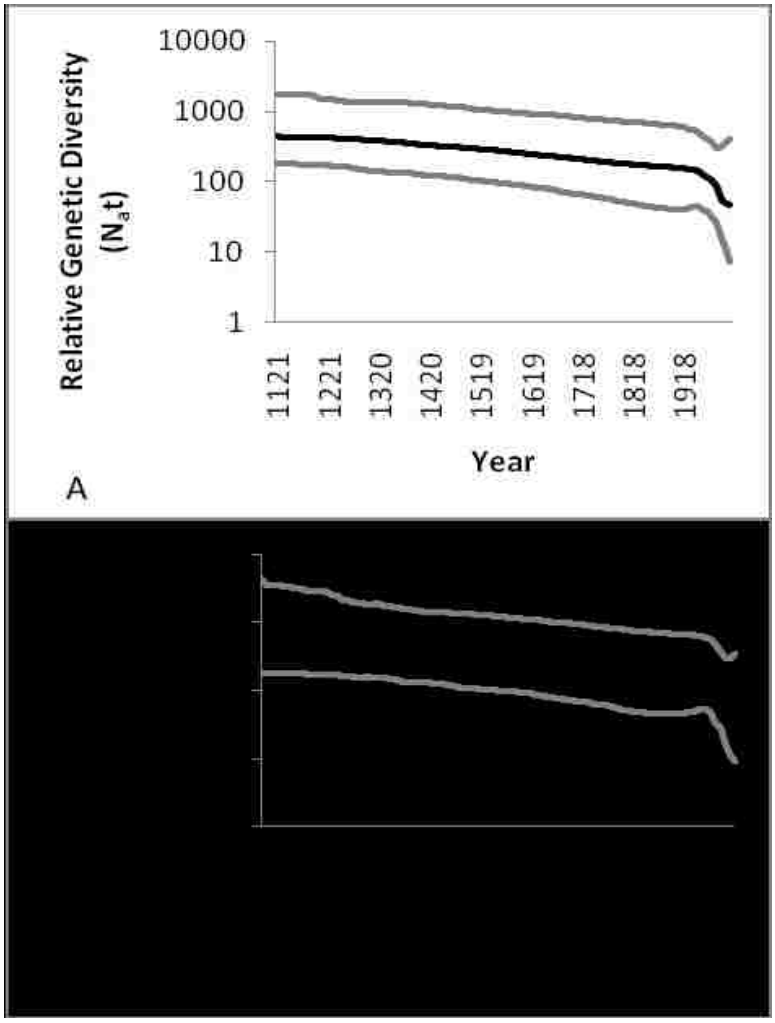


Figure 2: Population dynamics of POW and DTV depicted using Bayesian skyline plots (BSP) derived from (A) envelope coding sequences and (B) NS-5 coding sequences. The plots show changes in relative genetic diversity, depicted as the effective number of infections (Neτ, where Ne is the effective population size and τ the generation time from infected host to infected host) through time. The thick solid line is the median estimate of Neτ, while the 95% highest posterior density (HPD) intervals are shown in grey.

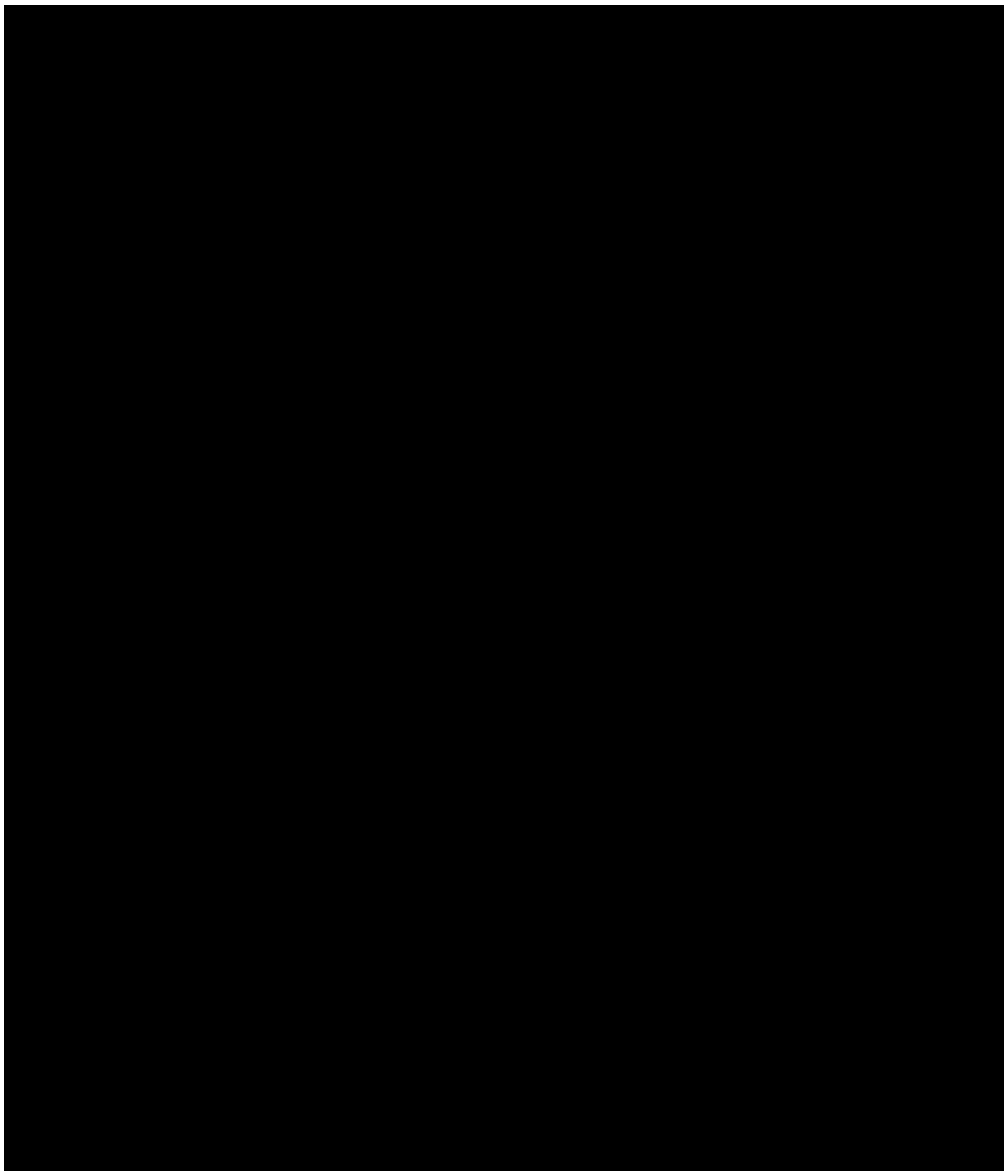


Figure 3: Maximum Clade Credibility (MCC) tree generated from concatenated envelope and NS-5 sequences (4200 nt) of POW and DTV strains. The median node ages and their 95% HPD values are shown on the major nodes, with 95% HPD values shown in parenthesis.

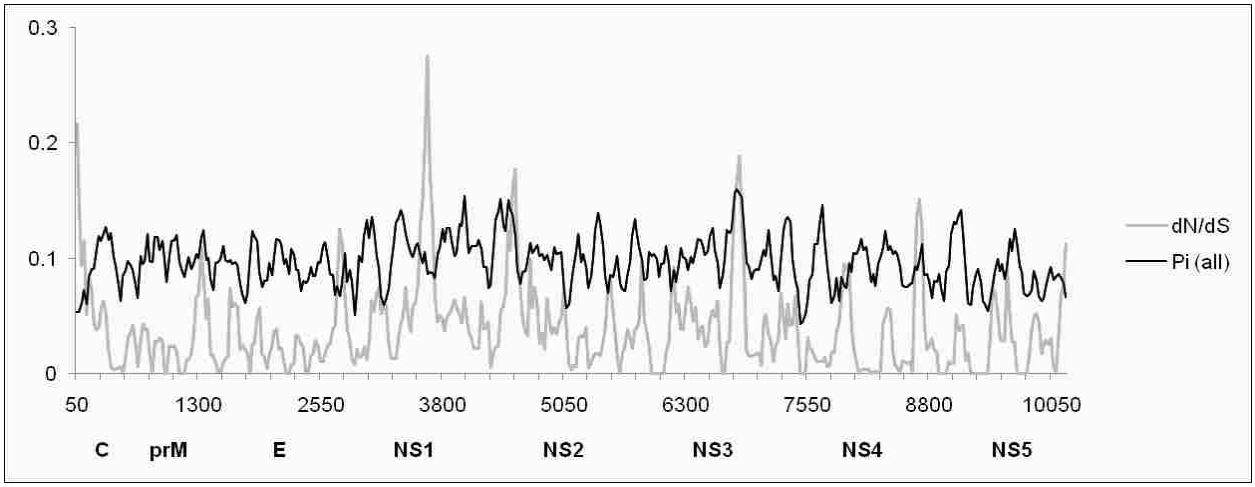


Figure 4: Estimates of dN/dS and genetic diversity (π) across POW and DTV full genome coding sequences, with estimates averaged over 50 nt windows.

References

Artsob H. (1988). Powassan encephalitis. In The Arboviruses pp. 29-50. Edited by T. P. Monath. Boca Raton, FL: CRC Press.

Baillie G. J., Kolokotronis S., Waltari E., Maffei J. G., Kramer L. D., Perkins S. L. (2008). Phylogenetic and evolutionary analyses of St. Louis encephalitis virus genomes. Mol Phylogenet Evol 47, 717-728.

Beasley D. W. C., Davis C. T., Guzman H., Vanlandingham D. L., Travassos da Rosa A., Parsons R. E., Higgs S., Tesh R. B., Barrett A. D. T. (2003). Limited evolution of West Nile virus has occurred during its southwesterly spread in the United States. Virology 309, 190-195.

Beasley D. W. C., Suderman M. T., Holbrook M. R., Barrett A. D. T. (2001). Nucleotide sequencing and serological evidence that the recently recognized deer tick virus is a genotype of Powassan virus. Virus Res 79, 81-89.

Bertolotti L., Kitron U., Goldberg T. L. (2007). Diversity and evolution of West Nile virus in Illinois and the United States, 2002-2005. Virology 360, 143-149.

Blaskovic D. and Nosek J. (1972). The ecological approach to the study of tick-borne encephalitis. Prog Med Virol 14, 275-320.

Brackney D. E., Brown I. K., Nofchissey R. A., Fitzpatrick K. A., Ebel G. D. (2010). Homogeneity of Powassan virus populations in naturally infected Ixodes scapularis. Virology 402, 366-371.

Brackney D. E., Nofchissey R. A., Fitzpatrick K. A., Brown I. K., Ebel G. D. (2008). Stable prevalence of Powassan virus in Ixodes scapularis in a northern Wisconsin focus. Am J Trop Med Hyg 79, 971-973.

Bryant J. E., Holmes E. C., Barrett A. D. T. (2007). Out of Africa: a molecular perspective on the introduction of yellow fever virus into the Americas. PLoS Pathog 3, e75.

Carroll S. A., Bird B. H., Rollin P. E., Nichol S. T. (2010). Ancient common ancestry of Crimean-Congo hemorrhagic fever virus. Mol Phylogenet Evol 55, 1103-1110.

Charrel R. N., Zaki A. M., Fakeeh M., Yousef A. I., de Chesse R., Attoui H., de Lamballerie X. (2005). Low diversity of Alkhurma hemorrhagic fever virus, Saudi Arabia, 1994-1999. Emerg Infect Dis 11, 683-688.

Daniels T. J., Boccia T. M., Varde S., Marcus J., Le J., Bucher D. J., Falco R. C., Schwartz I. (1998). Geographic risk for Lyme disease and human granulocytic ehrlichiosis in southern New York state. Appl Environ Microbiol 64, 4663.

Drummond A. J. and Rambaut A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol 7, 214.

Ebel G. D. and Kramer L. D. (2009). West Nile Virus: Molecular Epidemiology and Diversity. In West Nile Encephalitis Virus Infection pp. 25-43. Edited by M. S. Diamond: Springer.

Ebel G. D. and Kramer L. D. (2004). Short report: duration of tick attachment required for transmission of Powassan virus by deer ticks. Am J Trop Med Hyg 71, 268-271.

Ebel G. D., Spielman A., Telford S. R. (2001). Phylogeny of North American Powassan virus. J Gen Virol 82, 1657-1665.

Ebel G. D., Foppa I., Spielman A., Telford III S. R. (1999). A focus of deer tick virus transmission in the northcentral United States. Emerg Infect Dis 87, 570-574.

Ebel G. D., Campbell E. N., Goethert H. K., Spielman A., Telford III S. R. (2000). Enzootic transmission of deer tick virus in New England and Wisconsin sites. Am J Trop Med Hyg 63, 36-42.

Falco R. C., Daniels T. J., Fish D. (1995). Increase in abundance of immature Ixodes scapularis (Acari: Ixodidae) in an emergent Lyme disease endemic area. J Med Entomol 32, 522-526.

Goethert H. K. and Telford III S. R. (2009). Nonrandom distribution of vector ticks (Dermacentor variabilis) infected by Francisella tularensis. PLoS Pathogens 5, e1000319.

Gresikova M. and Calisher C. (1989). Tick-borne encephalitis. In The Arboviruses pp. 177-202. Edited by T. P. Monath. Boca Raton, FL: CRC Press.

Hall T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. 41, 95–98.

Hinten S. R., Beckett G. A., Gensheimer K. F., Pritchard E., Courtney T. M., Sears S. D., Woytowicz J. M., Preston D. G., Smith R. P.,Jrand other authors. (2008). Increased recognition of Powassan encephalitis in the United States, 1999-2005. Vector Borne Zoonotic Dis 8, 733-740.

Jenkins G. M., Rambaut A., Pybus O. G., Holmes E. C. (2002). Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J Mol Evol 54, 156-165.

Jerzak G., Bernard K. A., Kramer L. D., Ebel G. D. (2005). Genetic variation in West Nile virus from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection. J Gen Virol 86, 2175-2183.

Kuno G., Artsob H., Karabatsos N., Tsuchiya K. R., Chang G. J. (2001). Genomic sequencing of deer tick virus and phylogeny of Powassan-related viruses of North America. Am J Trop Med Hyg 65, 671-676.

Leonova G. N., Kondratov I. G., Ternovoi V. A., Romanova E. V., Protopopova E. V., Chausov E. V., Pavlenko E. V., Ryabchikova E. I., Belikov S. I., Loktev V. B. (2009). Characterization of Powassan viruses from Far Eastern Russia. Arch Virol 154, 811-820.

Librado P. and Rozas J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451-1452.

Lloyd-Smith J. O., George D., Pepin K. M., Pitzer V. E., Pulliam J. R. C., Dobson A. P., Hudson P. J., Grenfell B. T. (2009). Epidemic dynamics at the human-animal interface. Science 326, 1362-1367.

Mehla R., Kumar S. R. P., Yadav P., Barde P. V., Yergolkar P. N., Erickson B. R., Carroll S. A., Mishra A. C., Nichol S. T., Mourya D. T. (2009). Recent ancestry of Kyasanur Forest disease virus. Emerg Infect Dis 15, 1431-1437.

Nuttall P. A., Jones L. D., Labuda M., Kaufman W. R. (1994). Adaptations of arboviruses to ticks. J Med Entomol 31, 1-9.

Pond S. L. K. and Frost S. D. (2005a). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21, 2531-2533.

Pond S. L. K. and Frost S. D. W. (2005b). Not so different after all: a comparison of methods for detecting amino acid sites under selection. Mol Biol Evol 22, 1208-1222.

Pond S. L. K., Posada D., Gravenor M. B., Woelk C. H., Frost S. D. W. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol 23, 1891-1901.

Posada D. and Crandall K. (1998). Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817-818.

Rambaut A. and Drummond A.J. (2007). Tracer v 1.5. Available at http://beast.bio.ed.ac.uk/Tracer.

Rodgers S. E. and Mather T. N. (2007). Human Babesia microti incidence and Ixodes scapularis distribution, Rhode Island, 1998–2004. Emerg Infect Dis [serial on the Internet]. Available at http://www.cdc.gov/EID/content/13/4/633.htm.

Scheffler K., Martin D. P., Seoighe C. (2006). Robust inference of positive selection from recombining coding sequences. Bioinformatics 22, 2493-2499.

Snapinn K. W., Holmes E. C., Young D. S., Bernard K. A., Kramer L. D., Ebel G. D. (2007). Declining growth rate of West Nile virus in North America. J Virol 81, 2531.

Suchard M. A., Weiss R. E., Sinsheimer J. S. (2001). Bayesian selection of continuous-time Markov chain evolutionary models. Mol Biol Evol 18, 1001-1013.

Tavakoli N. P., Wang H., Dupuis M., Hull R., Ebel G. D., Gilmore E. J., Faust P. L. (2009). Fatal case of deer tick virus encephalitis. N Engl J Med 360, 2099-2107.

Telford S. R., Armstrong P. M., Katavolos P., Foppa I., Garcia A. S., Wilson M. L., Spielman A. (1997). A new tick-borne encephalitis-like virus infecting New England deer ticks, Ixodes dammini. Emerg Infect Dis 3, 165-70.

Whitney E. and Jamnback H. (1965). The first isolations of Powassan virus in New York state. Proc Soc Exp Biol Med 119, 432-435.

CHAPTER 3: WEST NILE VIRUS GENETIC DIVERSITY IS MAINTAINED DURING TRANSMISSION BY *CULEX PIPIENS QUINQUEFASCIATUS* MOSQUITOES

Brackney, D.E., Pesko, K.N., Brown, I.K., Deardorff, E.R., Kawatachi, J. and G.D. Ebel. 2011. West Nile virus genetic diversity is maintained during transmission by *Culex pipiens quinquefasciatus* mosquitoes. PlosONE, 6(9): e24466

WEST NILE VIRUS GENETIC DIVERSITY IS MAINTAINED DURING TRANSMISSION BY *CULEX PIPIENS QUINQUEFASCIATUS* MOSQUITOES

DOUG E. BRACKNEY, KENDRA N. PESKO, IVY K. BROWN, ELEANOR R. DEARDORFF, JON KAWATACHI AND GREGORY D. EBEL

Department of Pathology, University of New Mexico School of Medicine

Albuquerque, New Mexico, U.S.A.

Address for Correspondence:

Gregory D. Ebel

University of New Mexico School of Medicine

Department of Pathology

1 University of New Mexico

MSC 08-4640

Albuquerque, NM 87131

TEL: 505-272-3163

FAX: 505-272-9912

gebel@salud.unm.edu

Abstract

Due to error-prone replication, RNA viruses exist within hosts as a heterogeneous population of non-identical, but related viral variants. These populations may undergo bottlenecks during transmission that stochastically reduce variability leading to fitness declines. Such bottlenecks have been documented for several single-host RNA viruses, but their role in the population biology of obligate two-host viruses such as arthropod-borne viruses (arboviruses) *in vivo* is unclear, but of central importance in understanding arbovirus persistence and emergence. Therefore, we tracked the composition of West Nile virus (WNV; *Flaviviridae*, *Flavivirus*) populations during infection of the vector mosquito, *Culex pipiens quinquefasciatus* to determine whether WNV populations undergo bottlenecks during transmission by this host. Quantitative, qualitative and phylogenetic analyses of WNV sequences in mosquito midguts, hemolymph and saliva failed to document reductions in genetic diversity during mosquito infection. Further, migration analysis of individual viral variants revealed that while there was some evidence of compartmentalization, anatomical barriers do not impose genetic bottlenecks on WNV populations. Together, these data suggest that the complexity of WNV populations are not significantly diminished during the extrinsic incubation period of mosquitoes.

Introduction

West Nile virus (WNV; *Flaviviridae, Flavivirus*) was introduced into North America in 1999 and has since spread across the continental United States and into Canada, Mexico, the Carribean, and South America (Randolph and Rogers, 2010). Molecular epidemiologic studies of WNV in the US revealed that minor changes at the genomic level were associated with a dramatic shift in the genotypic composition of WNV circulating in North America (Anderson et al. 2001, Beasley et al. 2003, Davis et al. 2003, Ebel et al. 2001, 2004). Specifically, the introduced genotype, termed NY99, was displaced by a new variant, WN02. The WN02 genotype differs from NY99 by only a few nucleotide and/or amino acid changes, but is more efficiently transmitted by native *Culex* mosquitoes (Ebel et al. 2004, Moudy et al. 2006, Moudy et al. 2007). It was determined that the WN02 genotype requires a shorter extrinsic incubation period in mosquitoes (EIP, time from vector infection to transmission) thereby resulting in an increased vectorial capacity of local mosquitoes. Similarly, the emergence of Chikungunya virus (CHIKV; *Togaviridae, Alphavirus*) seems to have been facilitated by analogous mutations that result in increased transmission efficiency by the vector *Aedes albopictus* (Tsetsarkin et al. 2007, 2009). Thus, relatively minor consensus genetic changes can significantly influence arbovirus transmission patterns and disease emergence. Determining the mechanistic underpinnings of genetic change in arboviruses is therefore critical to understanding their persistence and emergence.

RNA viruses exist within hosts as a dynamic distribution of non-identical, but related viral variants (Domingo et al. 1996, Duarte et al. 1994, Eigen et al. 1988, Jerzak et al. 2005). High genetic diversity profoundly influences the population biology of RNA

viruses, including WNV, polio, mumps and hepatitis C viruses (Jerzak et al. 2007, Sauder et al. 2006, Sullivan et al. 2007, Vignuzzi et al., 2006). In the case of WNV, high genetic diversity is associated with increased fitness in mosquitoes (Fitzpatrick et al. 2010). Population bottlenecks may reduce fitness by stochastically reducing the genetic diversity of the virus population. *In vitro* studies of vesicular stomatitis virus, an RNA virus, have demonstrated that repeated bottlenecks can lead to fitness loss through the action of Muller's ratchet (Duarte et al. 1992). The extent to which mosquitoes impose such population bottlenecks on arthropod-borne viruses (arboviruses) is unclear. Analysis of WNV populations from naturally infected birds revealed that non-consensus, minority genotypes were shared among samples collected from multiple birds, suggesting that WNV populations may not be subject to bottlenecks during the natural transmission cycle (Jerzak et al. 2005). Similarly, it was suggested that dengue virus type 1 (DENV1; *Flaviviridae*, *Flavivirus*) is not subject to widespread population bottlenecks during the natural transmission cycle because putatively defective genomes persist through complementation, requiring frequent coinfection of cells in both mosquitoes and humans (Aaskov et al. 2006). Similarly, a high frequency of coinfection of midgut cells has been reported for Venezuelan equine encephalitis virus (VEEV; *Togaviridae*, *Alphavirus*) in *Aedes taeniorhynchus* (Smith et al. 2008). Conversely, studies examining early infection of mosquitoes by WNV and VEEV demonstrated that only a few (~15) midgut cells are susceptible to arbovirus infection (Smith et al. 2008, Scholle et al. 2004). These findings suggest that anatomical barriers, specifically cells of the midgut, may act as genetic bottlenecks by restricting the population of infecting virions thereby diminishing the genetic diversity of the population. Importantly, these observations are not mutually

exclusive as several viral genomes may coinfect a single midgut cell. Importantly, population bottlenecks associated with mosquito transmission have not been assessed from a virus genetics perspective.

Therefore, we determined whether WNV experiences genetic bottlenecks during the EIP in the vector mosquito *Culex pipiens quinquefasciatus*. We hypothesized that WNV experiences genetic bottlenecks during the EIP in mosquitoes, and reasoned that sequential reductions in viral genetic diversity would occur as infection progressed throughout the mosquito. To assess this, WNV genetic diversity was quantified in mosquito midguts, hemolymph, and salivary secretions, compartments that represent three well-characterized infection stages (midgut colonization, dissemination, and transmission). Three mosquitoes at three time points (7, 14, and 21 days post infection (dpi)) were sampled. Our data suggest that stochastic reduction of genetic diversity in mosquitoes is at most a minor component of WNV population biology during horizontal transmission.

Results

Mosquito infection rates. Mosquito tissues were screened for the presence of WNV RNA by one-step RT-PCR. All freshly fed mosquitoes were positive for WNV RNA representing 'input', blood-meal associated virus. The infection rates for midguts at 7, 14, and 21 days post infection (dpi), reflecting viral populations able to overcome the midgut infection barrier, were 88% (21/24), 86% (19/22), and 70% (14/20), respectively. The percentage of mosquitoes positive for WNV RNA in the legs, indicating virus dissemination from the midgut and into surrounding hemolymph, was 58% (14/24), 36%

(8/22), and 55% (11/20) at 7, 14, and 21 dpi, respectively. In order for mosquitoes to transmit WNV, the virus must be able to overcome the salivary gland infection and escape barriers. The percentage of mosquitoes with WNV in salivary secretions was 25% (6/24), 14% (3/22), and 35% (7/20) at 7, 14, and 21 dpi, respectively. Three mosquitoes per time point with WNV RNA in midgut, hemolymph and saliva were selected for further analysis and WNV genome equivalents quantified (Text S1). Genome equivalents were highest in midguts and progressively decreased in the hemolymph and saliva. Further, genome equivalents increased with time post infection (21 dpi > 14 dpi > 7dpi) (Figure S1). In addition, three mosquitoes, representing the 'input' group, were collected immediately post-engorgment. WNV genome equivalents determined for each of these individuals were 3.2, 4.1 and 5.9 x 10^5 genome equivalents/ ml (Figure S1). The bloodmeal contained 6 x 10^6 pfu/ ml, assuming 10-100 genomes per infectious particle and an engorgment volume of ~3 µl, engorged mosquitoes would be expected to contain ~1.4 x $10^{5 \text{ or } 6}$ genome equivalents. The concentrations for the three individuals in the 'input' group are in agreement with these calculations and thus faithfully represent the population of the bloodmeal as a whole.

Some arboviruses may enter mosquito hemolymph directly, bypassing midgut infection via a 'leaky midgut' (Houk and Hardy 1979, Weaver et al. 1991). In order to determine whether this occurred in the WNV-*Cx. quinquefasciatus* system, hemolymph was removed from mosquitoes at 1, 3, 24, and 48 hpi as well as 8 and 16 dpi and tested for WNV by plaque assay (Text S1). Hemolymph collected at 8 and 16 dpi commonly held high titers of WNV. In contrast, hemolymph collected at early timepoints after feeding almost never contained infectious WNV (Figure S2).

WNV genetic diversity. The percent nucleotide diversity and proportion of unique viral variants were used as indicators of viral genetic diversity in each of the samples. The percent nucleotide diversity was determined by calculating the total number of nucleotide changes for all clones within a given sample divided by the total number of nucleotides sequenced per sample. The data was grouped either by days post infection (Figure 1 A, B, and C) or by tissue type (Figure 1 D, E, and F). Analysis of the data set by days post infection revealed that there was no significant difference in the percent nucleotide diversity among the viral populations sequenced at 7 and 14 dpi between 'input', midgut, legs or saliva ($p=0.2739$ and $p=0.2662$, respectively) (Figures 1 A and B). Interestingly, genetic diversity seemed to decrease with time post infection as there was a significant reduction in diversity from the 'input' to the three tissue types at 21 dpi (ANOVA $p=0.0015$; Tukey's HSD post test, 'input' vs midgut $q=7.262$ $p<0.05$, 'input' vs legs $q=8.493$ $p<0.05$, and 'input' vs saliva $q=5.293$ $p<0.05$) , but no difference between tissue types (Figure 1C). Analyzing the data by tissue type revealed that there was a significant reduction in diversity from the 'input' to midguts at 14 and 21 dpi (ANOVA $p=0.0125$, Tukey's HSD post test, 'input' vs 14 dpi $q=5.404$ $p<0.05$ and 'input' vs 21 dpi $q=5.694$ $p<0.05$), but no significant difference between midguts at 7 dpi and 'input' (Figure 1D). There was no statistical differences between any of the leg or saliva samples at 7, 14 or 21 dpi (legs $p=0.0996$, saliva $p=0.3563$) (Figure 1E and 1F).

The second indicator of genetic diversity used in these studies was the proportion of unique viral variants. This was determined by calculating the number of unique clones per sample and dividing by the total number of clones sequenced per sample. Again the

data was grouped either by days post infection (Figure 2 A, B, and C) or tissue type (Figure 2 D, E, and F). At 7 dpi, the midguts and saliva were significantly lower than the 'input' (ANOVA $p=0.0052$, Tukey's HSD post test, 'input' vs midguts $q=7.038$ $p<0.05$, 'input' vs saliva $q=5.841$ $p<0.05$), but the tissues were not significantly different from one another (Figure 2A). Interestingly, by 14 dpi the proportion of unique viral variants between the tissues and 'input' was not significant (ANOVA $p=0.0517$), but at 21 dpi each of three tissue types were significantly lower than the 'input' (ANOVA $p=0.0021$, Tukey's HSD post test, 'input' vs midguts $q=6.898$ $p<0.05$, 'input' vs legs $q=8.028$ $p<0.05$, and 'input' vs saliva $q=5.306$ $p<0.05$) (Figure 2B and 2C). Analysis by tissue type revealed that midguts from all three time points were significantly lower than the 'input', but not different between time points (ANOVA $p=0.0014$, Tukey's HSD post test, 'input' vs 7dpi $q=7.298$ $p<0.05$, 'input' vs 14 dpi $q=7.838$ $p<0.05$, and 'input' vs 21 dpi $q=7.576$ $p<0.05$) (Figure 2D). Like the midguts, the legs at 14 and 21 dpi contained significantly less diversity than the 'input' (ANOVA $p=0.0018$, Tukey's HSD post test, 'input' vs 14 dpi $q=6.597$ $p<0.05$ and 'input' vs 21 dpi $q=8.458$ $p<0.05$), but was not different from the 7 dpi time point. Further, there was no difference between the time points (Figure 2E). Finally, comparison of the saliva samples at each of the time points revealed no significant difference between the three time points and the 'input' or between time points (ANOVA $p=0.1431$) (Figure 2F).

Because our frequency and location analysis of viral variants revealed that numerous variants were found in both the 'input' and saliva, but not the midgut or legs, we performed a correlation analysis between the genetic diversity metrics and log transformed viral genome equivalents. The Pearson correlation analysis revealed that

percent nucleotide diversity is significantly inversely correlated to viral genome equivalents ($p=0.003$; Pearson $r^2=-0.2747$) (Figure 3A). Similarly, viral genome equivalents are inversely correlated to the proportion of unique viral variants ($p=0.0205$; Pearson $r^2=0.1772$) (Figure 3B).

Frequency and migration analysis. Analysis of the frequency and location of viral variants revealed that 78 of 883 sequences sampled were unique. These variants were found in all three tissue types and at all three time points. There were 16 variants unique to the 'input' mosquitoes, 6 were found in all four categories (input, midgut, legs, saliva), 19 were unique to saliva, 15 unique to legs, and 9 unique to midguts, and the remaining variants were found in multiple tissues. The 14 most common variants were then plotted to display their relative proportion in each mosquito sample (Figure 4). By analyzing the data by this approach we were able to track individual variants from 'input' through infection (midguts), dissemination (legs), and transmission (saliva). The 'input' set is a combination of all three 0 hpi mosquitoes and as expected represents a complex population of multiple variants. Generally, the midgut populations, at all three time points, are composed of only a few variants with no one variant dominating in all samples. Likewise, the WNV populations recovered from the legs had, in general, low intrahost variability, but with no overrepresentation of any one variant between mosquitoes. Interestingly, there was an expansion in the total number of variants identified in the saliva compared to the midguts and legs. These findings are supported by the proportion of unique viral variants analysis (Figure 2). Further, many of the variants that were present in the input samples and subsequently undetected in the midguts and

legs were recovered from the saliva (Figure 4). Included in this data was a variant that contained a single nucleotide deletion at nucleotide 2194 in the E-glycoprotein. This deletion mutant was found in the legs or saliva of three different mosquitoes at 7 and 14 dpi (black colored sections of Figures 4A and 4B).

Migration analyses were performed in order to more closely look for evidence of genetic bottlenecks and test for tissue compartmentalization (Figure 5). A hypothetical tree was generated to represent what would be expected if strong genetic bottlenecks were influencing WNV populations during the EIP (Figure 5A). Under this scenario, a single 'input' variant initiates infection in the midgut. Subsequently, a single midgut variant establishes an infection in the hemolymph from which a single variant invades the salivary glands and is transmitted. However, this is not what we observed. Three representative mosquitoes, one from each time point, are shown and the remaining six trees are provided in a supplement (Figure S3). The migration analysis indicates some evidence for compartmentalization based on the p-values for ordered character states as estimated from 10,000 randomly sampled trees. A cluster of closely related variants isolated from legs was identified in our representative mosquito at 7 dpi (Figure 5B $p<0.0001$). Similarly, clusters of saliva variants were identified at 14 dpi and 21 dpi ($p<0.0001$ and $p<0.0001$, respectively) (Figures 5C and 5D).

In vivo competition assay. To test for the presence of genetic bottlenecks using a less genetically complex virus population, mosquitoes were fed on chicks infected with a known mixture of marked-reference and wild-type WNV (Text S1). The mean proportion of wild-type to reference virus in chick viremia was 0.73 (n=4, SEM 0.011) (Figure S4).

Subsequent analysis of the mosquito samples revealed that there was no change in the proportion of wild-type to reference WNV in any mosquito tissue compared to chick viremia (bodies 0.8 SEM 0.031, legs 0.73 SEM 0.07, and saliva 0.75 SEM 0.087).

Discussion

Population bottlenecks during transmission may profoundly influence the evolution of arboviruses by stochastically reducing population variation, thereby selecting random genomes that may be less fit than the overall population. Currently, it is unclear whether arboviruses experience genetic bottlenecks during infection of the mosquito vector. Therefore we tracked the composition of WNV populations during mosquito infection to quantify genetic bottlenecks associated with infection of these hosts.

Cx. p. quinquefasciatus mosquitoes were exposed to a bloodmeal containing a highly fit, genetically diverse WNV population (M24) that has been described in detail elsewhere (Fitzpatrick et al. 2010). Most mosquitoes exposed to M24 had WNV RNA in midgut tissues after either 7, 14 or 21 days EI. Although fewer mosquitoes had WNV in hemolymph and salivary secretions, at all timepoints at least three individual mosquitoes had WNV in midgut, hemolymph and salivary secretions. WNV from these mosquitoes was used to assess population bottlenecks associated with mosquito transmission.

Extreme care was taken to minimize the possibility that WNV RNA from one tissue would contaminate other tissues. Hemolymph was first sampled from newly anesthetized mosquitoes by gently removing their legs. Second, the mosquito mouthparts were inserted into a pulled capillary tube charged with buffer and the mosquito was

allowed to salivate for approximately 30 minutes. Finally, the midgut was removed from the mosquito and washed three times in PBS to remove hemolymph-associated WNV. This approach was validated by intrathoracically inoculating adult female mosquitoes with $2x10^4$ pfu/ml of WNV. Subsequently, five whole mosquitoes, five unwashed midguts and three times washed midguts were collected 45 minutes post inoculation. The presence of WNV RNA was determined by one-step RT-PCR. Expectedly, all five whole body mosquitoes were positive for WNV RNA along with two of five unwashed midguts. All five of the washed midguts were negative for WNV RNA (data not shown). Since the greatest concentration of WNV RNA tended to be in the mosquito midguts, handling this tissue last minimized the possibility of contaminating samples from the same mosquito. Several mosquitoes were detected that had midgut-limited infections, or WNV in hemolymph but not salivary secretions (data not shown). These results indicate that our efforts to minimize contamination were effective and that the samples selected for this study were not compromized by contaminating WNV RNA.

Quantitative analysis of viral genetic diversity in mosquito midguts compared with the highly genetically diverse 'input' WNV M24 clearly demonstrated that virus population diversity is restricted in this tissue. Both the percent nucleotide diversity and proportion of unique viral variants in midguts are significantly lower than in the 'input' population (Figures 1D and 2D). Surprisingly, however, the genetic diversity of peripheral WNV was not significantly different from virus in the bloodmeal, with saliva-associated WNV (i.e. the WNV that would be transmitted by a feeding mosquito) tending to be the most genetically diverse of the three tissue types sampled (Figures 1 and 2). The mechanism(s) that lead to increased genetic diversity outside of the midgut are not clear.

The presence of a 'leaky midgut' may explain this discrepency between the percent nucleotide diversity in legs and saliva compared to that of the midguts (Houk and Hardy 1979, Weaver et al. 1991). However, we found little evidence of WNV bypassing the midgut and directly infecting secondary tissues (Figure S2). In addition, we observed a significant inverse correlation between viral genome equivalents and the genetic diversity metrics (Figure 3). Taken together, these findings suggest that although WNV populations appear to be restricted in the midguts, and to a lesser extent in hemolymph, the genetic diversity of transmitted WNV was similar to that of the ingested virus population, and that variables other than tissue of origin determine viral genetic diversity in mosquitoes.

It may be that WNV accumulates mutations during the course of mosquito infection: relaxation of purifying selection on WNV sequences has been associated with mosquito infection (Jerzak et al. 2005). To assess this possibility, we compared the frequency and location of viral variants present in our mosquito tissue samples to those of the 'input' population. Not surprisingly, the majority of the viral variants identified in the midguts were also present in the 'input' population (Figure 4). Interestingly, however, variants found in legs and saliva were also represented in our 'input' dataset, without being present in the mosquito midguts. These findings support our quantitative analysis of genetic diversity, in these tissues, and indicate that the increased variation observed in peripheral WNV populations was more attributable to genetic diversity in the 'input' WNV population than to the generation of novel mutants during mosquito infection.

Additional evidence supporting the infection of a single cell by multiple WNV variants was obtained through examination of a defective WNV sequence in our dataset.

Specifically, we identified a single nucleotide deletion mutant that was found in multiple mosquito samples, including peripheral compartments, but not in the 'input' (Figure 4). Although it is possible that these mutants arose independently, it seems more likely that an ancestral mutant was present but undetected in the M24 population and was maintained in mosquitoes by complementation. Numerous studies have observed complementation of defective *Flavivirus* genomes in cell passage experiments (Tsai et al. 2007, Yoon et al. 2006). Typically, these studies have found large, ~2 kb, in-frame deletions at the 5'-end of the genome in the structural genes. Interestingly, one study found long-term transmission of a defective DENV-1 virus with a premature stop codon in the E gene (Aaskov et al. 2006). This data suggests that defective WNV particles can infect mosquitoes, propagate through complementation and ultimately be transmitted (Figure 4; mosquito 3 saliva 7 dpi). This implies that multiple WNV virions may frequently infect a single midgut cell, providing a mechanism by which WNV genetic diversity may be maintained in mosquitoes despite limitations in the number of susceptible midgut cells (Smith et al. 2008, Scholle et al. 2004).

Finally, we performed a migration analysis to formally test for the presence of bottlenecks and compartmentalization. We detected compartmentalization in legs and saliva, but found no evidence of genetic bottlenecks (Figure 5). In this analysis, if genetic bottlenecks exist, viral variants from the tissue samples would originate from a single 'input' variant as demonstrated in our hypothetical tree. Rather, variants identified in the saliva were found to originate from multiple 'input' variants indicating the ability of numerous 'input' variants to overcome multiple mosquito barriers to infection (i.e. midgut infection, midgut escape, and salivary gland escape barriers). The artificial nature

of this experimental system may explain the discrepancies between our tests. Mosquitoes were offered a bloodmeal containing WNV M24 which contains an approximately 10 fold increase in the genetic diversity compared to natural WNV populations (Jerzak et al. 2005, Fitzpatrick et al. 2010). This approach was implemented as a means to more easily track variation in our populations. It may be that the perceived bottlenecks were artificial due to saturating the system. As a more realistic approach to testing for bottlenecks we performed an *in vivo* competition assay in which infectious clone-derived wild-type WNV was competed against a marked reference virus (Fitzpatrick et al. 2010). It was observed that the proportion of marked refernce virus to wild-type WNV remained unchanged from 'input' to bodies, legs or saliva (Figure S4). Together, these data suggest that genetic bottlenecks do not significantly influence WNV populations during the EIP in *Cx. p. quinquefasciatus*.

Our genetic approach to transmission bottlenecks provides an intersting contrast to previous studies of bottlenecks in arbovirus transmission cycles (Smith et al. 2008, Scholle et al. 2004). Using virus-like particles to track binding and internalization, one study demonstrated that WNV infects only a few midgut epithelial cells during infection of *Cx. quinquefasciatus* (Scholle et al. 2004). Similar results were found during VEEV infection of *Aedes taeniorhynchus* (Smith et al. 2008). By virtue of the small number of infected cells it was concluded that arbovirus populations may be stochastically reduced at the point of infection. Our genetics studies of WNV do not support this obsevation. Notably, these conclusions are not necessarily mutually exclusive: It may be that the small proportion of susceptible midgut cells are infected with more than one virus particle or that an undetectable level of infection occurred in a higher proportion of cells.

In fact, a high frequency of dual infections were observed in the VEEV-*Aedes* system (Smith et al. 2008). Essentially, only a few susceptible midgut cells may be needed to propagate a diverse arbovirus population. Our observation of a deletion mutant persisting, apparently through complementation, during mosquito infection supports this possibility.

The literature regarding the role of bottlenecks in natural transmission cycles of RNA viruses is currently ambiguous. Bottlenecks are seemingly unimportant for Cauliflower mosaic virus in plants, but may exist for other RNA viruses (Ali and Roossinck, 2010, Kuss et al. 2008, Monsion et al. 2008, Wang et al. 2010). Numerous factors may contribute to this discrepency such as virus species, single vs two-host systems, mode of transmission and/ or site of inoculation. Further, environmental or host genetic factors may influence differences between individual hosts within a given population and likely explain the high variablity observed between individual mosquitoes in this experiment (Cadwell et al. 2010). Nevertheless, our data establish that transmitted WNV populations are at least as diverse as those of the imbibed population and therefore suggests that genetic bottlenecks are unlikely to significantly influence WNV population biology during horizontal transmission.

Methods

Virus and mosquito infection. The highly genetically diverse WNV population, WNV M24, used for these studies has been previously described (Fitzpatrick et al. 2010). Briefly, 24 WNV isolates from naturally infected mosquitoes and birds were passaged once on C6/36 *Aedes albopictus* cells (Jerzak et al. 2005). Titers were determined by plaque assay on Vero cells and mixed at a 1:1:1… ratio. This mixture was amplified once

on C6/36 cells at an MOI of 0.1 and the resultant population titered and genetically characterized.

To infect mosquitoes, WNV M24 was mixed 1:1 with defibrinated goose blood and offered to adult female *Culex pipiens quinquefasciatus* 7-8 days post emergence. The virus titer in the bloodmeal was 6 x 10^6 pfu/ml. Fully engorged females were separated from the remaining unfed mosquitoes and housed in an environmental chamber (27˚C, 16:8 L:D photoperiod) for the remainder of the experiment.

Sample collection and virus detection. To quantify viral genetic diversity of the 'input' virus population, three fully engorged mosquitoes were placed in RNA extraction buffer immediately after feeding and homogenized using a mixer mill. Viral genetic diversity was quantified as described below. At 7, 14, and 21 days post-infection paired tissues samples (midguts, legs, and saliva) were collected from 20-25 mosquitoes. To ensure that contaminating WNV from the hemoceol was not introduced into our midgut samples, dissected midguts were washed three times in PBS before placing the samples in RNA extraction buffer. Dissecting forceps were flame sterilized between dissections to avoid cross contamination between samples. Total RNA was extracted from mosquito hemolymph and tissues using the RNeasy Mini Protect kit (Qiagen, Valencia, CA) and screened for the presence of WNV RNA by one-step RT-PCR using the Superscript III kit with platinum *Taq* (Invitrogen, Carlsbad, CA). WNV specific primers used in this study spanned a 934 nt. region corresponding to the E-NS1 junction (1971 nt-2928 nt). Three mosquitoes with detectable WNV RNA in all three tissue types were selected from each time point for further analyses.

Quantification of viral genetic diversity. Viral genetic diversity was determined according to methods previously described (Jerzak et al. 2005). Briefly, cDNA was generated from 5 µl of total RNA using the High Fidelity Reverse Transcription kit (Stratagene, Cedar Creek, TX) according to the manufacturers specifications and WNV specific primers, WNV 1971 F and WNV 2928 R. Subsequently, the cDNA served as template for high fidelity *Pfu* Ultra polymerase amplification (Stratagene). Amplicons were PCR purified and cloned into the pCR Script Amp $^{(+)}$ vector (Stratagene). Between 21-30 individual clones from each of the samples were sequenced using the M13F, M13R, WNV 2369 F, and WNV 2768 R primers. DNAStar's SeqMan module (DNAStar Inc., Madison, WI) was used for sequence alignment and analysis of genetic diversity. Only clones with two-fold sequencing coverage were considered complete. As a means to estimate genetic diversity, consensus sequences for each sample were determined and individual clones within that sample were then compared to the specimen-specific consensus sequence. The percent nucleotide diversity (total number of mutations from all clones within a sample divided by the total number of nulcoetides sequenced per sample) and the proportion of unique viral variants (the number of unique clones differing from the consensus divided by the total number of clones sequenced per sample) were calculated and used as indicators of genetic diversity.

Quantification of viral genome equivalents. WNV genome equivalents were determined by quantitative-RT-PCR (Q-RT-PCR). As a standard control for this assay a ~2 kb fragment from the WNV E gene was amplified using the WNV 1031 F and WNV3430 R

primers. The resultant amplicon was cloned into the pCR2.1-TOPO vector (Invitrogen) downstream of the T7 promoter. The recombinant vector was linearized with *Kpn* I, purified and used as template for *in vitro* transcription using the T7 Megascript kit according the manufacturer's instructions (Ambion, Austin, TX). The resultant RNA was quantified and aliquoted in serial ten-fold dilutions. Using a probe specific for the E gene, the WNV 1160 F and WNV 1229 R primers, and the TaqMan ® One-Step RT-PCR Master Mix Reagent (Applied Biosystems, Foster City, CA) viral RNA copy numbers were determined (Lanciotti et al. 2000). Samples were run on the ABI Prism 7000 Sequence Detection System (Applied Biosystems).

Frequency and migration analysis. The presence of genetic bottlenecks and/ or compartmentalization was further assessed by migration analyses. To determine the frequency and location of viral variants, sequences from each sample were aligned in DNAStar's SeqMan module, exported as FASTA files and duplicates removed using BioEdit (Hall, 1999). Alignments were generated for each mosquito and tested for recombination using the Genetic Algorithm for Recombination detection program implemented on the datamonkey.org website (Pond et al. 2006). Evidence of recombination was not detected, so the alignments were used to perform a migration analysis. To test the null hypothesis of panmixis versus the alternative that there are distinct WNV sub-populations within different mosquito tissues, we used the Slatkin-Maddison test for gene flow in MacClade version 4 (Sinauer Associates, Sunderland, MA) (Slatkin and Maddison, 1989). Tissue of origin was assigned to each taxon in a one-character data matrix. 'Input' sequences from freshly-fed mosquitoes were included

as an estimate of the population of variants present in the infectious bloodmeal. In total there were four character states (input, midgut, legs, and saliva). The Slatkin-Maddison test was performed independently for each mosquito. This analysis was performed on Bayesian phylogenies, generated with MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). These were run with a general time reversible (GTR) model with invariable rates with substitution rates following a gamma plus invariants distribution. Two Markov Chains Monte Carlo (MCMC) tree searches of 5 million generations each were run in parallel with sampling one in every 250 trees. 50% majority-rule consensus trees are shown based on the last 19,000 trees. Briefly, the phylogenetic tree resulting from the nucleotide data was loaded into MacClade and the most parsimonious reconstruction of this ancestral character inferred with the Fitch algorithm in order to estimate the minimum number of steps required to explain the distribution of tissue states on the tree of interest (Fitch 1971). We then generated 10,000 random trees by random joining and splitting of the input tree and compared the number of steps on our input tree to those calculated in the random trees, as described previously for HIV-1, using ordered tissue states (Salemi et al. 2007).

Statistical analysis. Statistical analyses were completed in Microsoft Excel and GraphPad Prism. A one-way analysis of variance (ANOVA) with the Tukey's multiple comparison post-test with a significance level of $\alpha=0.05$ was used for analysis of the percent nucleotide diversity and proportion of unique viral variants. A Pearson correlation analyses was perfomed on log transformed viral genome equivalents versus percent

nucleotide diversity and proportion of unique viral variants. Figures were generated in GraphPad.

Supplemental information

Quantification of viral genome equivalents. WNV genome equivalents were determined by quantitative-RT-PCR (Q-RT-PCR). As a standard control for this assay a ~2 kb fragment from the WNV E gene was amplified using the WNV 1031 F and WNV3430 R primers. The resultant amplicon was cloned into the pCR2.1-TOPO vector (Invitrogen) downstream of the T7 promoter. The recombinant vector was linearized with *Kpn* I, purified and used as template for *in vitro* transcription using the T7 Megascript kit according the manufacturer's instructions (Ambion, Austin, TX). The resultant RNA was quantified and aliquoted in serial ten-fold dilutions. Using a probe specific for the E gene, the WNV 1160 F and WNV 1229 R primers, and the TaqMan ® One-Step RT-PCR Master Mix Reagent (Applied Biosystems, Foster City, CA) viral RNA copy numbers were determine (Lanciotti et al. 2002). Samples were run on the ABI Prism 7000 Sequence Detection System (Applied Biosystems).

Hemolymph Collection and Titration

Colonized *Cx. p. quinquefasciatus* mosquitoes were reared according to standard laboratory procedures. Females that had been starved for ~36h were allowed to feed on infectious clone-derived WNV mixed 1:1 with defibrinated goose blood (Rockland Immunochemicals Inc., Gilbertsville, PA). Engorged mosquitoes were separated and held for 1, 3, 24 or 48 hours; or 8 or 16 days extrinsic incubation. Hemolymph was sampled from mosquitoes anesthetized with triethylamine (Sigma) by inserting a glass needle into the posterior portion of the thorax, adjacent to the abdomen. Care was taken to avoid rupturing the midgut, and coloration of the hemolymph that might indicate such was noted. Approximately 0.3-0.5ul of hemolymph was sampled per mosquito. Hemolymph

was expelled from the needle into 100ul of mosquito diluent and stored at -80 until use. Infectious virus in hemolymph samples was quantified by plaque assay on Vero cells according to standard methods.

In vivo competition assays. Equal titers of an infectious clone-derived marked reference virus and infectious clone derived wild-type virus were mixed and ~2000 PFU subcutaneously inoculated into chicks at a concentration (Fitzpatrick et al. 2010, Shi et al. 2002). Two days post infection female *Culex pipiens quinquefasciatus* were fed on the infected chicks. At 7 dpi mosquito bodies, legs and saliva were collected, RNA extracted, and the proportion of WT-WNV determined by RT-PCR followed by SNPS analysis (Hall and Little, 2007).

Acknowledgements

The authors thank Kelly Fitzpatrick for technical assistance, and Marcelo Jacobs-Lorena for helpful discussions regarding leaky midgut experiments. In addition we would like to thank the reviewers and editors for their thoughful and insightful suggestions.

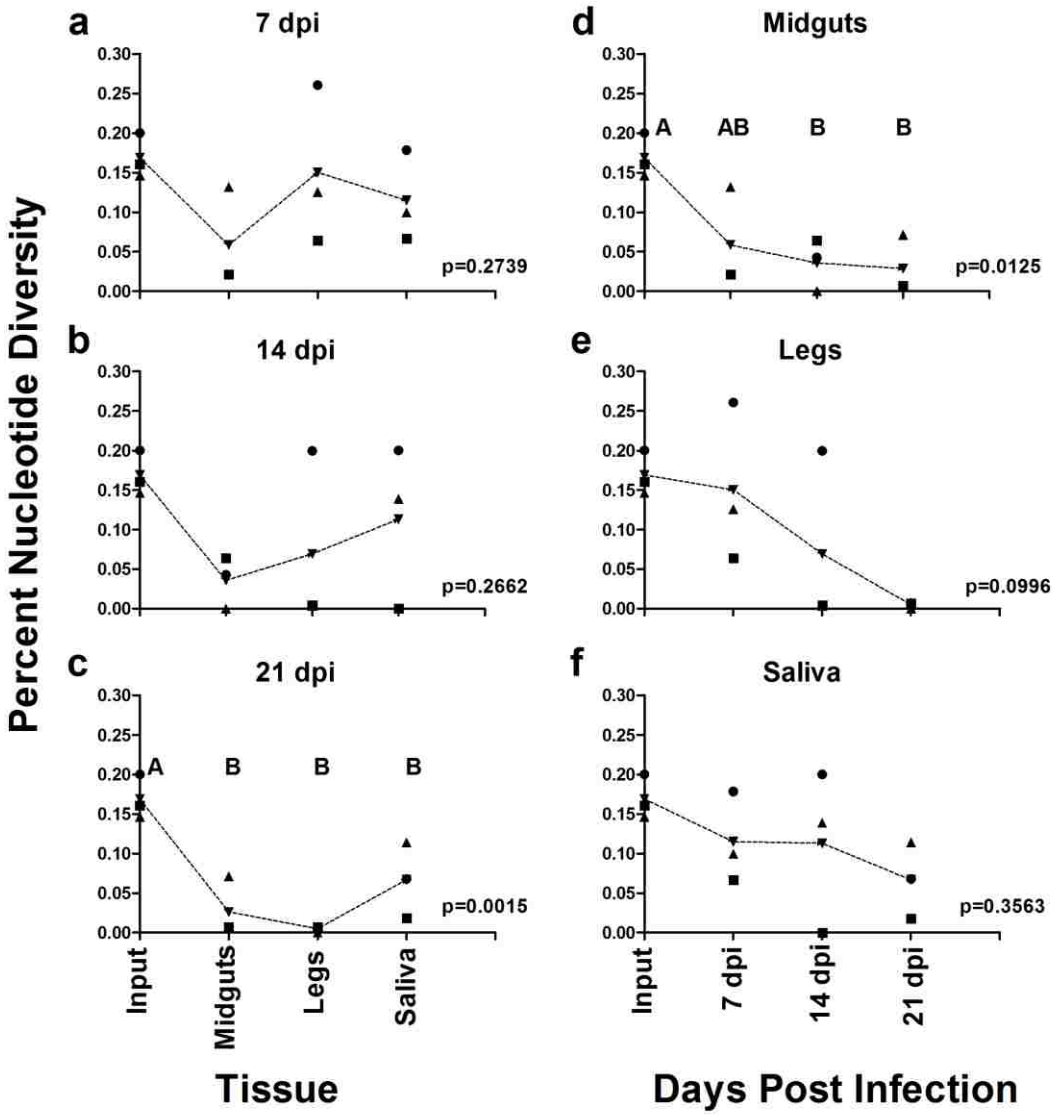


Figure 1: Percent nucleotide diversity by time and tissue. The percent nucleotide diversity was determined for each sample and plotted by either days post infection (7 dpi (A), 14 dpi (B), and 21 dpi (C)) or by tissue type (midguts (D), legs (E), and saliva (F)). Dotted lines connect the means for each sample set. P-values were determined by ANOVA using Tukey's multiple comparison post test. Letters above sample sets represent statistically significant groupings (p-value<0.05). Figures without letters denote that samples were not significantly different from one another.

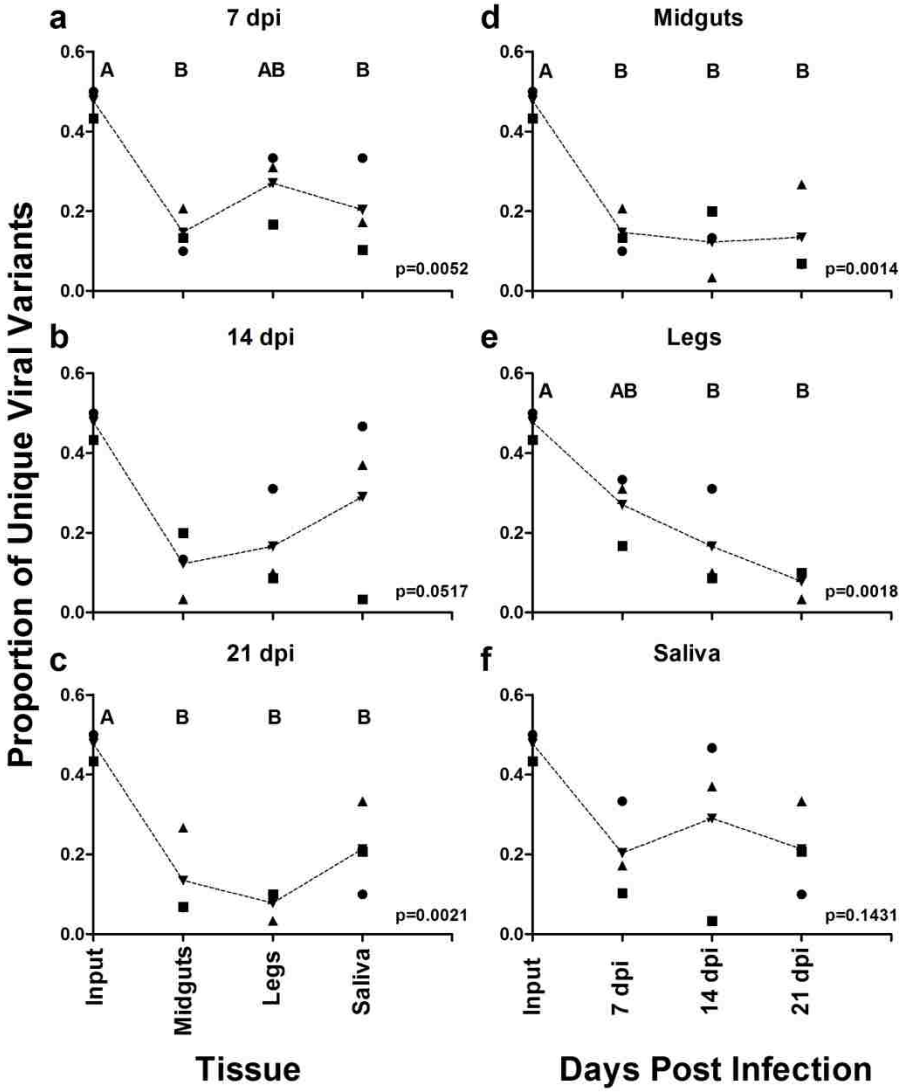


Figure 2: Proportion unique viral variants by time and tissue. The proportion of unique viral variants was determined for each sample and plotted by either days post infection (7 dpi (A), 14 dpi (B), and 21 dpi (C)) or by tissue type (midguts (D), legs (E), and saliva (F)). Dotted lines connect means for each sample set. P-values were determined by ANOVA using Tukey's multiple comparison post test. Letters above sample sets represent statistically significant groupings (p-value<0.05). Figures without letters denote that samples were not significantly different from one another.

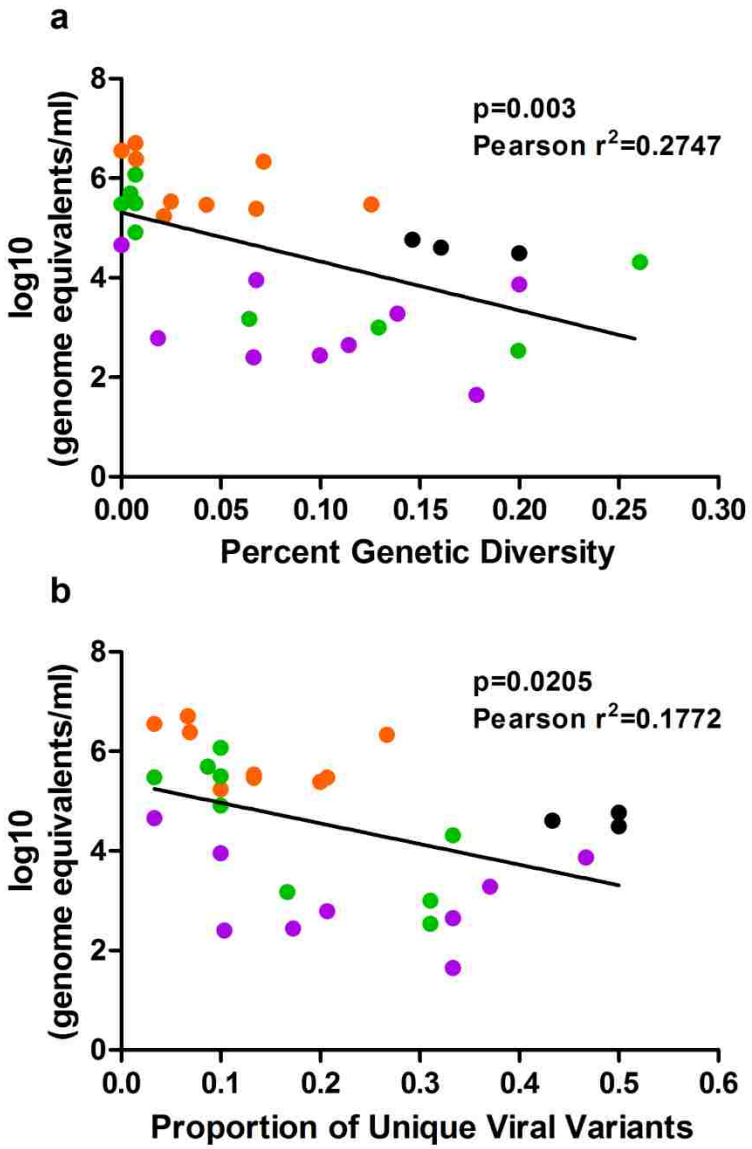


Figure 3: Viral genome equivalents and genetic diversity are inversely correlated**.** Black = 'input', Orange = midguts, Green = legs, and Purple = saliva.(A) Log transformed genome equivalents for each sample plotted against percent nucleotide diversity, n=30, p=0.003, Pearson r^2=0.2747. (B) Log transformed genome equivalents for each sample plotted against the proportion of unique viral variants, n=30, p=0.0205, Pearson r^2=0.1772.

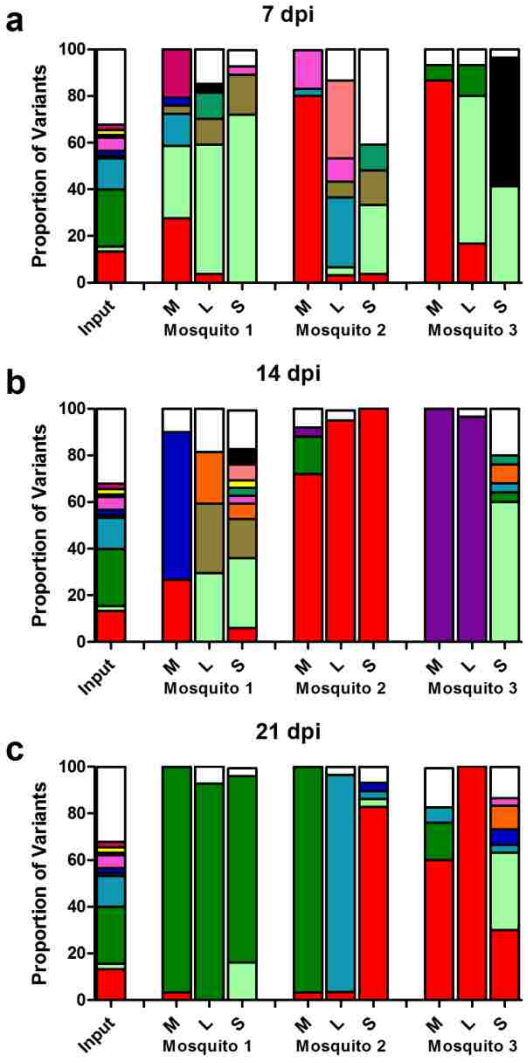


Figure 4: Frequency and location of unique viral variants. There were 883 clones sequenced, of which 78 sequences were unique. The frequency of the fourteen most common viral variants was mapped back to each sample. The column labeled input combines the data from all three 0 hours post infection mosquitoes. Samples are broken down by days post infection (7 dpi (A), 14 dpi (B), and 21 dpi (C)). Each time point includes three mosquitoes (denoted mosquitoes 1-3) and further broken down by tissue

(M (midguts), L (legs), and S (saliva)). The white sections of the histograms represent the remaining 64 uncommon variants and the black sections represent a single nucleotide deletion mutant found in multiple samples.

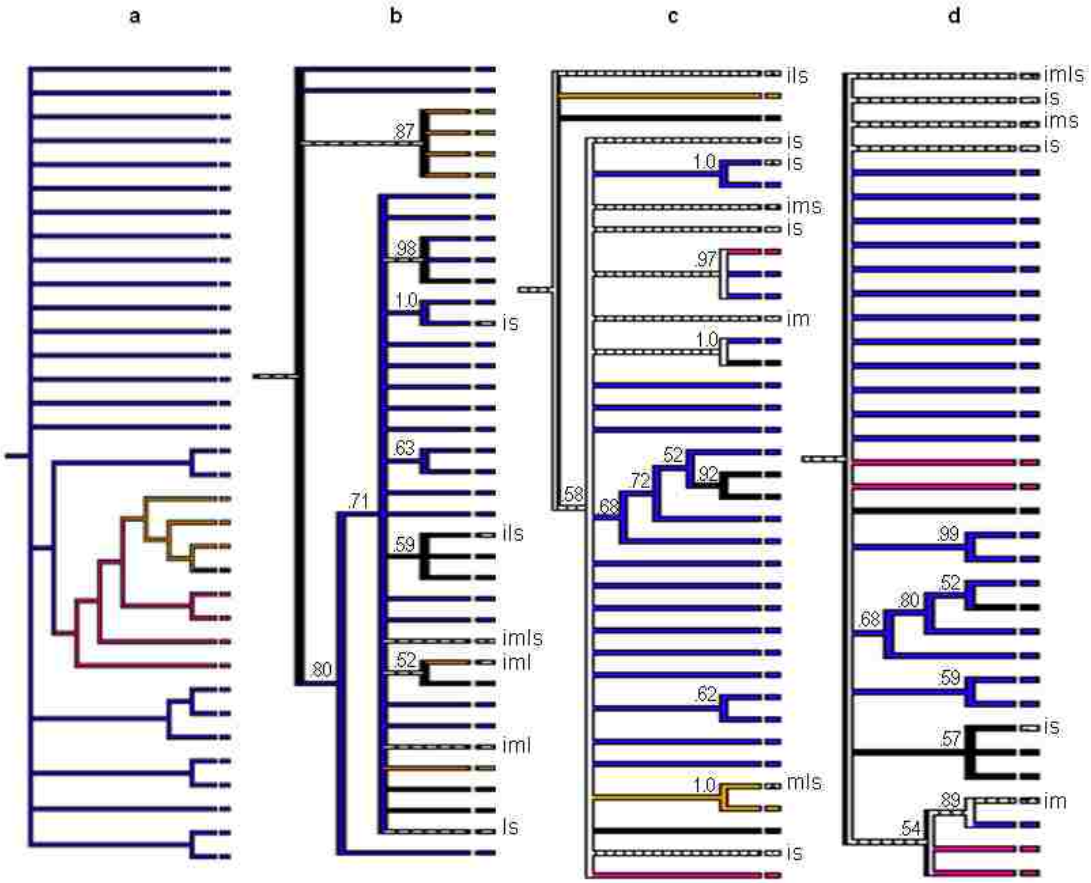


Figure 5: Bayesian trees with the most parsimonious reconstruction of tissue character states for three mosquitoes**.** Blue=input (i), pink=midgut (m), gold=legs (l), black=salivary secretions (s), dotted=multiple tissues, with specific tissues indicated by abbreviations. Hatched branches indicate equivocal reconstruction of character states. Numbers above nodes are the posterior probabilities inferred for each clade. A) Hypothetical tree with predicted outcome assuming the presence of genetic bottlenecks.

The most parsimonious reconstruction of ordered character steps for each tree for three representative mosquitoes is as follows, B) 7 dpi mosquito 2 (22 steps, $p<0.0001$), C) 14 dpi mosquito 1 (30 steps, $p<0.0001$), D) 21 dpi mosquito 3 (23 steps, $p=0.0001$).

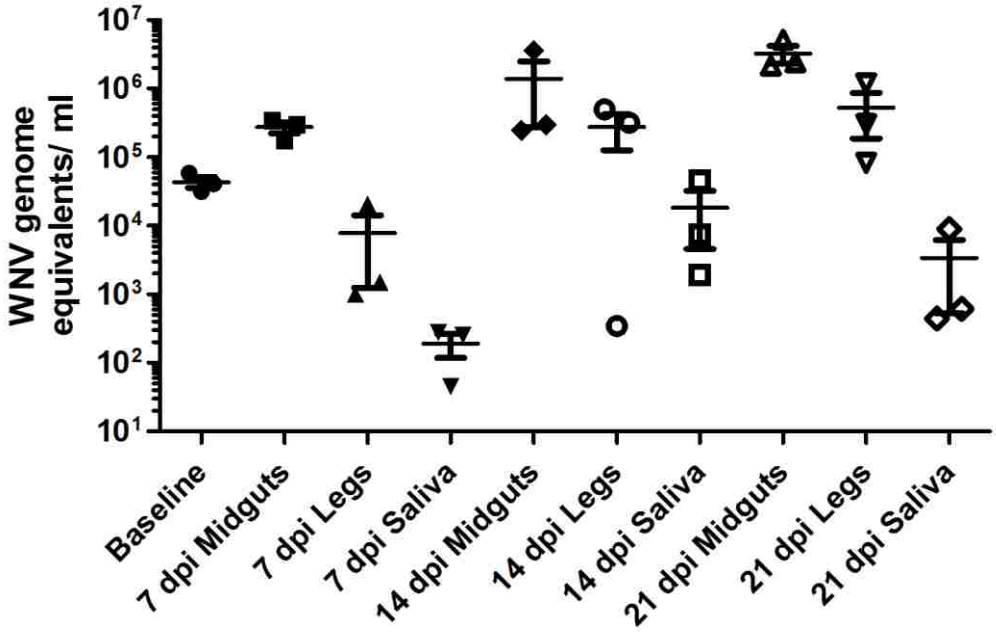


Figure S1: WNV genome equivalents per tissue sample. WNV genome equivalents were determined by Q-RT-PCR for each sample characterized.

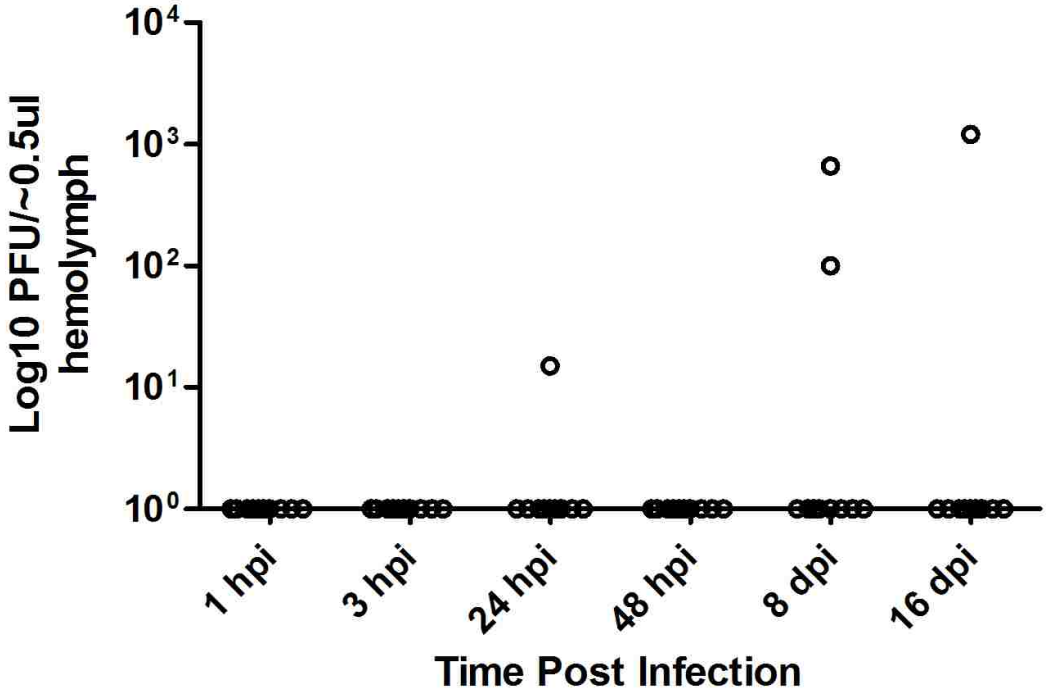


Figure S2: WNV titers in *Culex pipiens quinquefasciatus* hemolymph at early time points. Mosquitoes were offered a WNV Mix24 infectious bloodmeal and hemolymph extracted at multiple time points. WNV titers were determined by plaque assay.

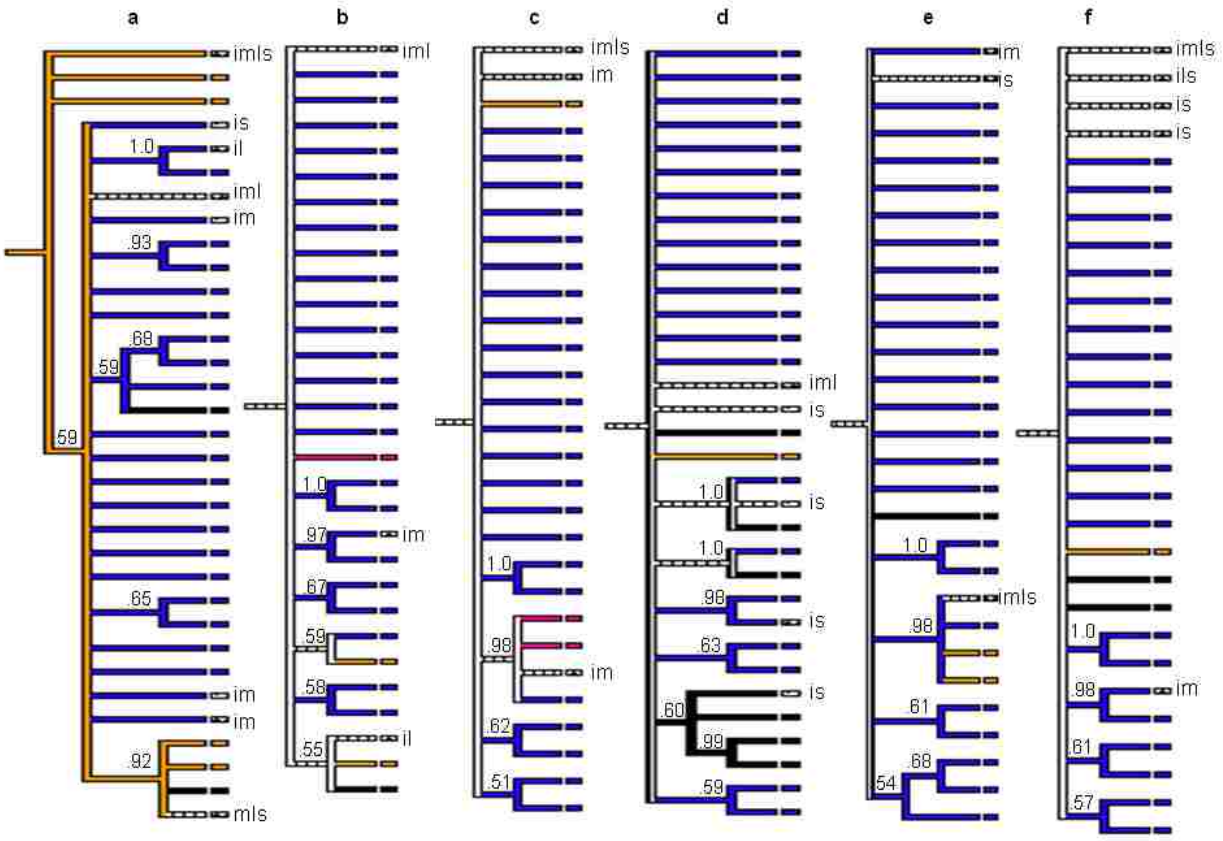


Figure S3: Bayesian trees with the most parsimonious reconstruction of tissue character states from the six remaining mosquitoes. Blue=input (i), pink=midgut (m), gold=legs (l), black=salivary secretions (s), dotted=multiple tissues, with specific tissues indicated by abbreviations. Hatched branches indicate equivocal reconstruction of character states. Numbers above the nodes are the posterior probabilities inferred for each clade. Mosquito analyzed and most parsimonious reconstruction of ordered character steps for each tree is as follows A) 7 dpi mosquito 1 (21 steps, p=0.0002) B) 7 dpi mosquito 3 (10 steps, p=0.0057), C)14 dpi mosquito 2 (8 steps, p=0.087), D) 14 dpi mosquito 3 (23 steps, p=0.0002), E) 21 dpi mosquito 1 (12 steps, p=0.0772), F) 21 dpi mosquito 2 (16 steps, p=0.007).

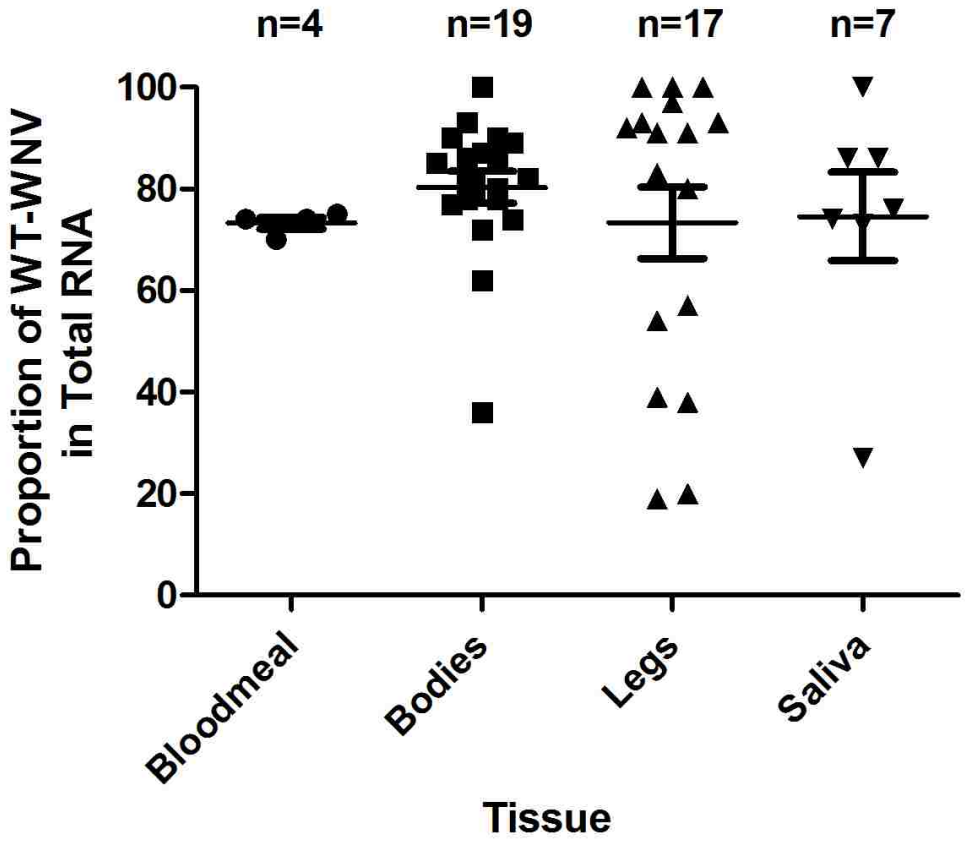


Figure S4: The proportion of wild-type WNV when competed against a marked reference virus does not change as the virus disseminates through the mosquito. *Culex pipiens quinquefasciatus* mosquitoes were fed on live chicks circulating a mixed population of WNV comprised of wild-type (WT) and reference viruses. Tissues were harvested 7 dpi from 20 mosquitoes and the proportion of WT-WNV was determined by RT-PCR followed by SNPS analysis. Samples negative for WNV RNA by RT-PCR were omitted.

References

Aaskov J., Buzacott K., Thu H.M., Lowry K., Holmes E.C. (2006) Long-term transmission of defective RNA viruses in humans and *Aedes* mosquitoes. Science 311: 236-238.

Ali A., Roossinck M.J. (2010) Genetic bottlenecks during systemic movement of cucumber mosaic virus vary in different host plants. Virology 404: 279-283.

Anderson, J.F., Vossbrinck, C.R., Andreadis, T.G., Iton, A., Beckwith, W.H., et al. (2001) A phylogenetic approach to following West Nile virus in Connecticut. Proc Natl Acad Sci USA 98: 12885-12889.

Beasley, D.W.C., Davis, C.T., Guzman, H., Vanlandingham, D.L., da Rosa, A., et al. (2003) Limited evolution of West Nile virus has occurred during its southwesterly spread in the United States. Virology 309: 190-195.

Cadwell, K., Patel, K.K., Maloney, N.S., Liu, T.C., Ng, A.C.Y., et al. (2010) Virus-plus-susceptibility gene interaction determines Crohn's disease gene Atg16L1 phenotypes in intestine. Cell 141: 1135-U1164.

Davis, C.T., Beasley, D.W.C., Guzman, H., Raj, P., D'Anton, M., et al. (2003) Genetic variation among temporally and geographically distinct West Nile virus isolates, United States, 2001, 2002. Emerg Infect Dis 9: 1423-1429.

Domingo, E., Escarmis, C., Sevilla, N., Moya, A., Elena, S.F., et al. (1996) Basic concepts in RNA virus evolution. Faseb J 10: 859-864.

Duarte, E., Clarke, D., Moya, A., Domingo, E., Holland, J. (1992) Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proc Nat Acad Sci USA 89: 6015-6019.

Duarte, E.A., Novella, I.S., Weaver, S.C., Domingo, E., Wainhobson, S., et al. (1994) RNA virus quasispecies - Significance for viral disease and epidemiology. Infect Agent Dis 3: 201-214.

Ebel, G.D., Carricaburu, J., Young, D., Bernard, K.A., Kramer, L.D. (2004) Genetic and phenotypic variation of West Nile virus in New York, 2000-2003. Am J Trop Med Hyg 71: 493-500.

Ebel, G.D., Dupuis, A.P., Ngo, K., Nicholas, D., Kauffman, E., et al. (2001) Partial genetic characterization of West Nile Virus strains, New York State, 2000. Emerg Infect Dis 7: 650-653.

Eigen, M., McCaskill, J., Schuster, P. (1988) Molecular quasispecies. J Phys Chem 92: 6881-6891.

Fitch, W.M. (1971) Toward defining the course of evolution - minimum change for a specific tree topology. System Zoo 20: 406.

Fitzpatrick, K.A., Deardorff, E.R., Pesko, K., Brackney, D.E., Zhang, B., et al. (2010) Population variation of West Nile virus confers a host-specific fitness benefit in mosquitoes. Virology 404: 89-95.

Hall, T.A. (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 95-98.

Hall, G.S. and Little, D.P. (2007). Relative quantitation of virus population size in mixed genotype infections using sequencing chromatograms. Journal of Virological Methods 146, 22-28

Houk, E.J. and Hardy, J.L. (1979) *In vivo* negative staining of the midgut continuous junction in the mosquito, *Culex tarsalis* (*Diptera*, *Culicidae*). Acta Trop 36: 267-275.

Jerzak, G., Bernard, K.A., Kramer, L.D., Ebel, G.D. (2005) Genetic variation in West Nile virus from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection. J Gen Virol 86: 2175-2183.

Jerzak, G.V.S., Bernard, K., Kramer, L.D., Shi, P.Y., Ebel, G.D. (2007) The West Nile virus-mutant spectrum is host-dependant and a determinant of mortality in mice. Virology 360: 469-476.

Kuss, S.K., Etheredge, C.A., Pfeiffer, J.K. (2008) Multiple host barriers restrict poliovirus trafficking in mice. PLoS Path 4: 13.

Lanciotti, R.S., Kerst, A.J., Nasci, R.S., Godsey, M.S., Mitchell, C.J., et al. (2000) Rapid detection of West Nile virus from human clinical specimens, field-collected mosquitoes, and avian samples by a TaqMan reverse transcriptase-PCR assay. J Clin Microbiol 38: 4066-4071.

Monsion, B., Froissart, R., Michalakis, Y., Blanc, S. (2008) Large bottleneck size in cauliflower mosaic virus populations during host plant colonization. PLoS Path 4: 7.

Moudy, R.M., Dupuis, A., Ebel, G.D., Kramer, L.D. (2006) Displacement of the introduced genotype of West Nile virus in New York state. Am J Trop Med Hyg 75: 300-300.

Moudy, R.M., Meola, M.A., Morin, L.L.L., Ebel, G.D., Kramer, L.D. (2007) A newly emergent genotype of west Nile virus is transmitted earlier and more efficiently by *Culex* mosquitoes. Am J Trop Med Hyg 77: 365-370.

Pond, S.L.K., Posada, D., Gravenor, M.B., Woelk, C.H., Frost, S.D.W. (2006) Automated phylogenetic detection of recombination using a genetic algorithm. Mol Biol Evol 23: 1891-1901.

Randolph, S.E. and D.J. Rogers.(2010) The arrival, establishment and spread of exotic diseases: patterns and predictions. Nat Rev Microbiol 8: 361-371.

Ronquist, F. and J.P. Huelsenbeck. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572-1574.

Sauder, C.J., Vandenburgh, K.M., Iskow, R.C., Malik, T., Carbone, K.M., et al. (2006) Changes in mumps virus neurovirulence phenotype associated with quasispecies heterogeneity. Virology 350: 48-57.

Salemi, M., Burkhardt, B.R., Gray, R.R., Ghaffari, G., Sleasman, J.W., et al. (2007) Phylodynamics of HIV-1 in lymphoid and non-lymphoid tissues reveals a central role for the thymus in emergence of CXCR4-using quasispecies. PLoS One 2: 11.

Scholle, F., Girard, Y.A., Zhao, Q.Z., Higgs, S., Mason, P.W. (2004) trans-packaged West Nile virus-like particles: Infectious properties *in vitro* and in infected mosquito vectors. J Virol 78: 11605-11614.

Shi, P.Y., Tilgner, M., Lo, M.K., Kent, K.A., Bernard, K.A. Infectious cDNA clone of the epidemic West Nile virus from New York City. *Journal of Virology* 76, 5847-5856 (2002).

Slatkin, M. and W.P. Maddison .(1989) A cladistic measure of gene flow inferred from the phylogenies of alleles. Genetics 123: 603-613.

Smith, D.R., Adams, A.P., Kenney, J.L., Wang, E., Weaver, S.C. (2008) Venezuelan equine encephalitis virus in the mosquito vector *Aedes taeniorhynchus*: Infection initiated by a small number of susceptible epithelial cells and a population bottleneck. Virology 372: 176-186.

Sullivan, D.G., Bruden, D., Deubner, H., McArdle, S., Chung, M., et al. (2007) Hepatitis C virus dynamics during natural infection are associated with long-term histological outcome of chronic hepatitis C disease. The Journal of Infectious Diseases 196: 239-248.

Tsai, K.N., Tsang, S.F., Huang, C.H., Chang, R.Y. (2007) Defective interfering RNAs of Japanese encephalitis virus found in mosquito cells and correlation with persistent infection. Virus Res 124: 139-150.

Tsetsarkin, K.A., McGee, C.E., Volk, S.M., Vanlandingham, D.L., Weaver, S.C., et al. (2009) Epistatic roles of E2 glycoprotein mutations in adaption of chikungunya virus to *Aedes albopictus* and *Ae. aegypti* mosquitoes. PLoS One 4: 12.

Tsetsarkin, K.A., Vanlandingham, D.L., McGee, C.E., Higgs, S. (2007) A single mutation in chikungunya virus affects vector specificity and epidemic potential. PLoS Path 3: 1895-1906.

Vignuzzi, M., Stone, J.K., Arnold, J.J., Cameron, C.E., Andino, R. (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. Nature 439: 344-348.

Wang, G.P., Sherrill-Mix, S.A., Chang, K.M., Quince, C., Bushman, F.D. (2010) Hepatitis C virus transmission bottlenecks analyzed by deep sequencing. J Virol 84: 6218-6228.

Weaver, S.C., Scott, T.W., Lorenz, L.H., Repik, P.M. (1991) Detection of eastern equine encephalomyeltis virus deposition in *Culiseta melanura* following ingestion of radiolabeled virus in blood meals. Am J Trop Med Hyg 44: 250-259.

Yoon, S.W., Lee, S.Y., Won, S.Y., Park, S.H., Park, S.Y., et al. (2006) Characterization of homologous defective interfering RNA during persistent infection of Vero cells with Japanese encephalitis virus. Mol Cells 21: 112-120.

CHAPTER 4: INTERNALLY DELETED WEST NILE VIRUS GENOMES ISOLATED FROM EXOTIC BIRDS IN NEW MEXICO: FUNCTION IN CELLS, MOSQUITOES, AND MICE

Submitted, Virology, November 8, 2011.

INTERNALLY DELETED WNV GENOMES ISOLATED FROM EXOTIC BIRDS IN NEW MEXICO: FUNCTION IN CELLS, MOSQUITOES, AND MICE

KENDRA N. PESKO[1], KELLY A. FITZPATRICK[1], ELIZABETH M. RYAN[2], PEI-YONG SHI[3], BO ZHANG[4], NIALL J. LENNON[2], RUCHI M. NEWMAN[2], MATTHEW R. HENN[2] AND GREGORY D. EBEL[5]

[1]Department of Pathology, University of New Mexico School of Medicine, Albuquerque, NM

[2]Broad Institute of MIT and Harvard, Cambridge, MA

[3]Wadsworth Center, New York State Department of Health, Albany, NY

[4]State Key laboratory of Virology, Wuhan Institute of Virology, Chinese Academy of Sciences, China

[5]Microbiology, Immunology and Pathology Department, College of Veterinary Medicine and Biomedical Sciences, Colorado State University, Fort Collins, CO

Running head: Defective West Nile viral particles isolated from birds

Address for Correspondence:
Gregory Ebel
Colorado State University
Microbiology, Immunology, and Pathology Department
College of Veterinary Medicine and Biomedical Sciences
Rampart Drive
Fort Collins, CO, 80523
TEL: 970-491-8374
FAX: 970-4918-8323
Gregory.Ebel@colostate.edu

Abstract

RNA viruses exist in their hosts as a heterogeneous population of related variants. Here we characterize three WNV isolates that contain, along with full-length genomes, mutants with large internal deletions to structural and nonstructural protein coding regions. The isolates were obtained from lorikeets that died from WNV encephalitis at the Rio Grande Zoo in Albuquerque, NM. In cell culture, these internally deleted WNV genomes function as defective interfering particles, reducing the production of full length virus when introduced at high multiplicities of infection. In mosquitoes, the presence of these deletion bearing WNV genomes reduced infection and dissemination rates and mosquito body titers overall, although they were not detected in legs or salivary secretions at 14 and 21 days post infection. In mice, inoculation with internally deleted genomes did not attenuate pathogenesis relative to full length or infectious clone derived virus, and subgenomes were not detected in mice at time of death.

Keywords: West Nile virus; genetic variation; defective interfering particles.

Introduction

West Nile virus, (WNV), (*Flaviviridae: Flavivirus*), is a mosquito-borne virus now endemic to the United States. It circulates between avian hosts and various mosquito vectors, mostly in the genus *Culex*. Since its first detection in the United States in 1999, WNV has dispersed throughout the western hemisphere (Artsob et al., 2009; Komar and Clark, 2006; Murray et al., 2010). WNV has a high error rate during replication because the lack of a proofreading mechanism in the virus-encoded RNA dependent RNA polymerase can lead to the production of defective RNA genomes (Holland et al., 1982; Holmes, 2009). WNV and other arthropod-borne viruses (arboviruses) have slower rates of evolution than single-host RNA viruses, which may be a result of evolutionary constraints imposed by the requirement for replication in both vertebrate and invertebrate cells (Jenkins et al., 2002). Supporting this, arbovirus populations in nature are subject to strong purifying selection (Weaver, 2006). Nonetheless, WNV and other arboviruses exist in mosquito and vertebrate hosts as a genetically heterogeneous mixture, with many individual mutants making up a genetically diverse population (Jerzak et al., 2005; Aaskov et al. 2006). Error-prone replication and consequent intrahost viral genetic diversity, including the deleterious variants, are a central feature of the population biology of RNA viruses.

Defective interfering particles (DIPs) are subgenomic viral particles that arise during the course of viral infection, replicate through complementation with full length homologous viruses, and have an inhibitory effect on virus growth (Thompson et al., 2009). DIPs have been detected after serial passage of the mosquito borne flaviviruses, WNV and Japanese encephalitis virus (JEV) in cell culture (M. A. Brinton, 2001;

Debnath et al., 1991; M. A. Brinton, 1983; Tsai et al., 2007; Yoon et al., 2006). Until recently, whether DIPs exist in naturally acquired infections, and their role in shaping the outcome of virus-host interactions *in vivo* has been unclear. Importantly, isolates from patients with acute dengue infections were recently shown to contain deletion mutants made up of the 5' and 3' untranslated regions of the genome (with the entire protein-coding sequence deleted), which behave *in vitro* as DIPs (Li et al., 2011). In another member of *Flaviviridae*, hepatitis C virus, genomes containing large deletions to structural coding regions have been made from chronically infected individuals. Therefore, DIPs may occur in flavivirus infections more frequently than previously appreciated, and their presence within a virus population may significantly impact the outcome of infection from both transmission and pathogenesis standpoints.

Recently, we identified three WNV populations isolated from three lorikeets that died at two separate times in the Rio Grande zoo that contain mutants with large internal deletions. We therefore sought to characterize these mutants and determine whether they function as WNV DIPs. We tested multiple virus doses in Vero cells, hypothesizing that at a low multiplicity of infection (MOI) no coinfection would occur, so virus production would not be impacted, whereas higher MOIs would allow coinfection, thus interference may be observed at higher doses. We then determined the extent to which they influence transmission by mosquitoes and pathogenesis in mice to evaluate the possibility that the mutant genomes may contribute to WNV persistent infection and/or disease outcomes in vertebrates. Our results suggest that naturally occurring WNV deletion mutants function as DIPs, and that these DIPs interfere with virus transmission by mosquitoes. However,

the WNV DIPs we examined have surprisingly little impact on virus pathogenesis in mice.

Results

Identification of deletion mutants. RT-PCR targeting structural coding regions of the WNV genome (table 1: DH1F/DH1R) resulted in bands approximately 2 kb smaller than expected for samples from three birds that died of WNV infection (Fig 1A). All three samples were taken from Rainbow Lorikeets (*Trichoglossus haematodus*) as confirmed by sequencing of the cytochrome oxidase (COI) gene. Isolates were made by inoculating Vero cells with kidney tissue homogenate of birds found dead at the Albuquerque zoo in August of 2005 and 2007. Smaller genomes were confirmed directly using northern analysis of RNAs from one isolate (2774) passaged in Vero cells, at 48 hours post infection and compared to virus derived from an infectious clone (NY99) harvested 48 hours post infection in Vero cells. A probe to the NS-5 coding region annealed to a single band for RNA extracted from infectious clone derived virus infected cells but identified two bands in RNA extracted from cells infected with isolate 2774 (Fig 1B), in contrast, a probe to the envelope coding region resulted in a single band produced for each RNA sample (Fig 1C). Thus, shortened genomes could be detected directly with probes to NS-5 and appeared to make up around half of viral RNA present, but only full length genomes were detected with envelope probe, as expected.

Genomic location of internal deletions. PCR amplicons from bird kidneys were cloned and sequenced to reveal large deletions within the structural and NS-1 coding regions.

Deletions were in-frame and resulted in loss of 3' portions of prM, complete envelope, and 5' portions of NS-1 encoding RNA (Fig 1D). All eight clones sequenced from one isolate (2774) contained the same deletion, whereas the other two isolates (3336, 3337) had two types of deletions present, labeled A and B in figure 1D.

Identification of sequence polymorphism in lorikeet isolates. Full genome sequencing of isolate 2774 revealed a predicted amino acid substitution, 2K-V9M, that had previously been detected in virus that was capable of bypassing superinfection exclusion (Zou et al., 2009b). Sequencing of the original sample to confirm this substitution showed a polymorphism at this site, both A and G appeared to be present at nucleotide (nt) position 6871 (determining 2K-V9 or 2K-M9, respectively). Sequencing of the other two isolates, 3336 and 3337, showed the same polymorphism present. To further characterize the presence of this polymorphism, RNA was extracted from individual plaques of 2774vp1, 2774vp3, 3336 (unpassed), and 3337 (unpassed) and subject to sequencing across the location of this mutation (nt.6871). The 6871G variant was more commonly detected and made up 7/10 plaque from 2774vp1, 2774vp3, and 3336, and 10/10 of the plaques from 3337. The 6871A variant was present as 3/10 plaques from 2774vp1, 2774vp3, and 3336, but not detected in any of the ten plaques from 3337. Nested PCR using primers designed to amplify solely internally deleted WNV genomes showed that 6871A was present for 2774. Bayesian analysis of partial genome sequences of NM isolates 3337 (NM05) and 2774 (NM07) did not group these two isolates together relative to other genomes sampled from the southwestern United States, indicating the mutation leading to this polymorphism may have arisen independently for each isolate (Figure 2).

Isolates inhibit production of full length virus in Vero cells. To assess whether deletion mutants behave as defective interfering particles (DIPs), we evaluated full length virus production from Vero cells infected over a range of multiplicity of infection (MOI) with isolate 2774 and with infectious clone derived WNV (WT). Full length virus production was measured by plaque assay of cell culture supernatants after 3 days of growth in Vero cells. The full length virus yield from each supernatant varied with virus isolate and viral dose, and these two factors had a significant interaction influencing virus production (Figure 3A, F= 28.51, p=0.0014). Supernatants from 2774 infected cells were tested by RNA extraction and quantitative RT-PCR to quantify the presence of deletion mutant genomes at 3 dpi (Fig 3B). No internally deleted genomes were detected in supernatants of 2774 at 3 dpi at the lowest MOI (0.001), and only one of three replicates at MOI 0.01 contained detectable deletion mutant, but the remaining treatments produced comparable amounts of internally deleted genomes (Fig 3B).

Deletion mutants decrease titers and spread of virus in mosquitoes. Mosquitoes were given blood meals containing either 2774 blind passaged two times on Vero cells and thus still containing deletion mutant and full length virus (DM+FL), or a plaque purified isolate from 2774 containing full length virus only (FL), as confirmed by RT-PCR and Northern analysis. Both blood meals contained approximately 2 x 10^7 PFU per ml, as determined by plaque assay. Mosquitoes fed on full length virus alone had significantly higher infection and dissemination rates (Fig 4A, infection rate: X^2=10.44, p=0.0012; dissemination rate: X^2=5.26, p=0.022), and log transformed body titers (Fig 4B, t= 3.189, p=0.0018; difference= 0.69 log10±0.21). Deletion mutants were detected by qRT-PCR in

bodies of 7/44 and 5/26 mosquitoes tested on days 14 and 21 post infection (Table 2), but not in legs or salivary secretions at either time point.

Deletion mutants do not attenuate morbidity and mortality in mice. To assess whether deletion mutants could influence morbidity and mortality in vertebrates, we tested two infecting viral doses in C3H mice, which are susceptible to mortality from infection with WNV, and one dose in C57Bl/6 mice, which are resistant to mortality (Brown et al., 2007). Inoculation with isolates containing internally deleted genomes did not attenuate morbidity or mortality in either mouse strain (Fig 5). Internally deleted genomes were not detected in mice either at time of death from infection, or at 28 days post infection when surviving mice were sacrificed. Mortality rates were not significantly different by inoculating strain, however, more C3H mice inoculated with deletion mutant containing virus died than those inoculated with infectious clone derived or plaque purified full length virus (Fig 5). C3H mice inoculated with plaque purified full length or deletion mutant containing 2774 stocks lost significantly more weight than infectious clone exposed and mock mice in both trials (High dose, $F=57.2$, $p< 0.00001$, Bonferroni multiple comparison t-tests were significant for all comparisons, $p<0.05$; Low dose, $F=55.9$, $p<0.00001$, Bonferroni multiple comparison t-tests were significant for all comparisons at $p<0.05$, except for deletion mutant vs. full length: $t=0.56$, ns). Infection was confirmed in surviving mice by plaque reduction neutralization tests.

Discussion

As part of our ongoing studies of WNV population genomics in North America, we identified a group of virus strains collected in two separate years from the same location and host species that produced atypical RT-PCR products. Cloning and

sequencing these products revealed a group of highly similar, large, in-frame deletions that removed a portion of the prM, all of the E, and a portion of the NS1 coding sequences. RT-PCR analysis of RNA extracted from primary kidney tissue indicated that full length WNV was present in the infecting population along with the shortened genomes. Northern blot analysis of RNA produced by one of these isolates confirmed that both full-length and shortened RNAs are produced during infection of Vero cells. Further, we found that after a single passage in Vero cells, the shortened genomes are retained as a portion of the WNV population. Collectively, these data suggest that genome deletions occur as a portion of the population in WNV in nature, and that these genomes can be complemented by full length genomes and packaged by structural proteins provided in *trans*.

The collection of several apparently similar deletions in WNV from the same place over two years raised the possibility that the deletions were identical by descent, i.e. had a common origin and perpetuated through time by complementation. Phylogenetic and sequence analysis of the deletion events demonstrated that all were slightly different from one another, with deletions beginning at nucleotide positions 529-717 and ending at positions 2627-3093. We reasoned that size differences might be apparent over time if the deletions shared a common origin, with specimens collected later in time possessing larger deletions than those collected earlier. However, no logical relationship was detected between the size of the deletion and when the specimen was collected. In fact, some of the smallest deletions were collected in 2007, two years after the first specimen was collected in 2005. Phylogenetic analysis whole WNV genomes from two isolates (2774 and 3337) indicated they were distantly related compared with other WNV isolates

made nearby in time and space. Taken together, these results strongly suggest that the deletions arose independently. These results differ from previous descriptions of defective dengue genomes, which contain a premature stop codon and were present in isolates from both human and mosquito samples over the course of two years. These dengue virus genomes also had an increased accumulation of nonsynonymous mutations downstream of the stop codon which led researchers to conclude that these truncated genomes are complemented by full length virus and transmitted from vector to host over the course of several years (Aaskov et al., 2006). Since the WNV mutants we describe apparently arose independently and share a common host origin, it may be that some vertebrates (in this case lorikeets) contribute to WNV population biology by selecting common genomic variants.

Comparison of these deletions to previously reported Flavivirus deletion mutants revealed a high degree of similarity to genome variants known to be associated with DIPs found *in vitro* for related viruses JEV and Murray Valley encephalitis virus (MVEV) (Yoon et al., 2006; Lancaster et al., 1998). We therefore sought to determine whether they might interfere with replication of full-length WNV in Vero cells. We found that at higher MOIs, production of full-length WNV is suppressed in comparison to an infectious clone derived WNV, indicating interference can occur when deletion mutants are added at a ratio where coinfection of the same cell may occur. At the lowest MOI, production of full-length WNV did not differ between these two treatments. These observations establish that the deletion mutants detected in this study function as DIPs *in vitro*, similar to deletion mutants isolated from acutely infected dengue patient sera that contained only the untranslated regions of the genome (Li et al. 2011).

To assess the ability of WNV DIPs to interfere with transmission by mosquitoes, we fed mosquitoes with an isolate of WNV that contained internally deleted genomes. Oral exposure to isolates containing defective genomes reduced infection and dissemination rates and body titers in *Culex quinquefasciatus* mosquitoes, relative to rates estimated for mosquitoes fed on a full length only control. For mosquitoes fed on deletion mutants, defective RNAs were detected in mosquito bodies but not legs or salivary secretions, at 14 and 21 days post infection. This could be a result of resource competition occurring in those tissues where coinfection and complementation has occurred, which has previously been shown to decrease replication of superinfecting strains of WNV (Zou et al., 2009b). Dissemination rate has been positively correlated with body titers in WNV infected mosquitoes, so a reduction of body titer through resource competition could have lead to the reduced dissemination rates observed in mosquitoes exposed to virus containing internally deleted genomes, despite the fact that no internally deleted genomes were detected in disseminated tissues (Anderson et al., 2010). Overall these results indicate the internally deleted genomes are able to interfere with the first stage of mosquito infection, which could have a downstream impact on transmission. However, the observed inefficiency of complementation across various mosquito tissues further supports our hypothesis that each mutant arose independently and may not be transmitted efficiently from vector to host.

To assess the influence of deletion mutants on morbidity and mortality in vertebrates, we inoculated two different strains of mice with a virus population containing internally deleted mutants. For both mouse strains, mortality was not attenuated in the mice inoculated with internally deleted mutants, relative to a full length

only control from that isolate, and an infectious clone produced control. For the more susceptible mouse strain, C3H, inoculation with deletion mutant containing and full length virus alone produced significantly more morbidity, as assessed by weight loss, than infectious clone derived WNV. The results of these infections in mice are in contrast to other studies which have used Semliki forest virus DI particles to reduce mortality during concurrent or prior immunization (Barrett and Dimmock 1984; reviewed Barrett and Dimmock 1986). In the infections presented here, internally deleted genomes were not detected in any of the tissues sampled at time of death, suggesting that the lack of interference could be due to a lack of complementation. It may be that inoculating at a higher dose or using a different vertebrate host, such as a bird, would influence the outcome.

Each of the isolates that contained internally deleted genomes as part of the infecting virus population were also polymorphic at a predicted amino acid substitution in the 2K peptide, 2K-V9M. This substitution was present in plaque picks from 2774vp1, 2774vp3, and 3336. It may be that this substitution predisposes the emergence of deletion mutants. Interestingly, the same amino acid substitution evolved in populations of WNV that were able to overcome superinfection exclusion and replicate in WNV replicon bearing cells (Zou et al., 2009b). This mutation has also been associated with West Nile virus resistance to 2'5'-oligoadenylate synthetase 1b, and to the flavivirus specific antiviral lycorine, and appears to enhance replication by increasing the rate of viral RNA synthesis (Mertens et al., 2010; Zou et al., 2009a). This substitution has occurred independently in numerous other North American lineages where it is under positive selection (Armstrong et al., 2011). Whether the presence of 2K-V9M in the population

containing internally deleted genomes was required for the emergence and propagation of these genomes or evolved in response to pressure exerted from these genomes is unclear.

In summary, kidney tissues from three lorikeets that died in the Rio Grande bio park in August of 2005 and 2007 contained as a portion of the virus population genomes with large internal deletions to the entire envelope coding region and parts of pre-membrane and nonstructural protein 1 coding regions. Sequencing revealed deletions of around 2 kb affecting similar regions in each isolate, in frame, and persisting through multiple passages. Subgenomic RNA was directly detected by Northern analysis for one isolate, 2774. This isolate acted as a defective interfering particle when introduced at high doses to Vero cells. Body titers, infection, and dissemination rates in *Culex pipiens quinquefasciatus* fed on stocks containing deletion mutants were lower relative to a full length only plaque pick from the same isolate. We found no evidence for attenuation of infection in a susceptible mouse strain, C3H, inoculated with low or high dose, or in a resistant strain, C57/B6 mice, inoculated with a low dose of this virus. Further studies are needed to determine the prevalence of deletion mutants such as these *in vivo* and elucidate their potential roles in infected birds.

Methods

Cells and viruses. African green monkey kidney (Vero), and baby hamster kidney (Bhk-21) cells were purchased from ATCC. Cells were grown and maintained in Eagle's MEM supplemented by 10% fetal bovine serum (FBS) at 37˚C with 5% CO2. Plaque assays to determine titers of isolates were performed as described (Lindsey et al., 1976). WNV positive samples were provided to our lab from the New Mexico Department of Health. Three samples were identified as containing deletions to structural and NS-1 coding

regions. Two of these (3337, 3336) were submitted to the NM-DOH by the Rio Grande Zoo on 8/30/2005, collection dates are unavailable. The third (2774) was collected and submitted to the NM-DOH on 8/20/2007. All three samples were taken from Rainbow Lorikeets (*Trichoglossus haematodus*) as confirmed by sequencing of cytochrome oxidase (COI) gene (Wright et al., 2008). Virus stocks were propagated by addition of 100 µl of sample to T-25 flasks containing confluent Vero cells. A plaque pick from a plaque assay of 2774, confirmed to contain full length virus only by RT-PCR and Northern analysis was also propagated in Vero cells, and used as a control in multiple experiments, referred to as full length or FL. Infectious clone derived control virus was prepared by transfecting BHK-21 cells using transMessenger kits (Qiagen, Valencia, CA) with RNA transcribed *in vitro* from pFLWNV (Shi et al., 2002), using the mMessage mMachine T7 kit (Ambion, Austin, TX).

Extraction of RNA. RNA was extracted using an RNAeasy Protect Mini-Kit, according to the manufacturer's protocol (Qiagen, Valencia, CA) from 100 µl of viral stock, mouse or mosquito tissue, or cells homogenized in 350 µl buffer RLT, and eluted into 50 µl of Rnase free H_2O.

qRT-PCR, RT-PCR, cloning, and sequencing. Primers used in this study are listed in Table 1. cDNA was reverse transcribed from RNA using Superscript III First Strand synthesis kit with a 3' end primer (Table 1; Invitrogen, Carlsbad, CA). RT-PCR was conducted using Superscript III One-Step Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) with Platinum Taq (Invitrogen, Carlsbad, CA) with the following parameters: 55 °C for 30 min (reverse transcription), 95°C for 15 min (initial denaturation), and 40 cycles of 94 °C for 15 s, 60 °C for 30 s, and 72°C for 30 s, followed

by final extension at 72°C for 10 min. PCR was conducted with Pro(Promega, Madison, WI). PCR products produced for each isolate using Dh1F/Dh1R primers were cloned directly into TOPO vector (PCR4.0), following manufacturer's instructions. Positive colonies were assayed by PCR with M13F/R primers, and products were purified with Stratagene PCR purification kit (Agilent, Santa Clara, CA). PCR products were sequenced by UNM sequencing core or Genewiz (South Plainfield, NJ). Sequences from either end were aligned to the WNV lineage 1, strain NY99 reference genome, and positions given in text correspond to this full length reference genome (NCBI: NC_009942). Quantitative RT-PCR was performed on an ABI Prism 7000 Sequence Detection System (Applied Biosystems), using a double quenched probe across the mutant junction (Integrated DNA technologies, Coralville, IA), and the Brilliant II QRT-PCR kit (Agilent, Santa Clara, CA). RNA copies were estimated using a full length infectious clone construct made with fusion PCR at site BamHI and SphI in pFLWNV (Shi et al. 2002). Plasmid was propagated in HB101 cells, and transcribed *in vitro* using the mMessage mMachine T7 kit (Ambion, Austin, TX).

Virus dose assay. Twelve-well plates containing confluent Vero cells were infected at the noted MOIs with either pFLWNV clone derived WNV, or 2774vp2 (isolate 2774 passaged two times in Vero cells). Virus was added to cells, rocked for 5 min, and allowed to infect at 37˚C for 1 hour before additional media was added. Each MOI was introduced in three replicates. At 72 hours post infection, supernatants were harvested and full length virus was quantified by plaque assay.

Experimental hosts. Five-week old female C3H/HeN and C57/B6 mice were purchased from Harlan Laboratories (Houston, TX). Mice were housed in groups of 4-6, and

allowed one week to acclimate to BSL-3 conditions. Mice were provided with food and water *ad libitum*. At six weeks of age, mice were inoculated in the left rear footpad with 50 µl of animal inoculation diluent (cation- and endotoxin-free PBS with 1% FBS) containing the indicated dose of West Nile virus, or mock inoculated with inoculation diluent alone. Mice were weighed and assessed daily for clinical signs of disease, and sacrificed at 30% weight loss or when severe clinical signs were observed. At 28 dpi, serum from surviving mice was tested by plaque reduction neutralization test to confirm infection had occurred. Briefly, serum samples were spun down for 10 min at 3,000 rpm, upper aliquots were removed and mixed 1:1 with ~100 units of virus (as determined by plaque assay), incubated at 37°C for 30 min, and added to confluent Vero cells in 6 well plates. Samples were overlayed with agar and media, then agar and media with neutral red at 48 hpi, and plaques were read at 72 hpi. No more than 1-2 plaques were detected in any wells containing serum samples, and the majority contained no plaques, except for mock inoculated, indicating neutralizing antibodies to WNV existed in all surviving mice that had been inoculated with WNV.

Culex pipiens quinquefasciatus mosquitoes were reared in the laboratory at 27˚C, with 16:8 photoperiod (Brackney et al., 2009). At 6-7 days postemergence, female mosquitoes were provided with defribinated goose blood (Rockland, Gilbertsville, PA) and 2 x 10^7 pfu/ml WNV in a Hemotek membrane feeding apparatus (Accrington, UK). After 1 hr of feeding, mosquitoes were cold anesthetized and engorged mosquitoes were sorted into a fresh container. Mosquitoes were held at 27˚C for a 14 or 21 day extrinsic incubation, at which time bodies, legs, and salivary secretions were collected as described

elsewhere (Ebel et al., 2004). Infection and titers in infected tissues were assessed by plaque assay.

Northern hybridization. Northern analysis was performed on RNA extracted directly from cells infected 48 h previously with each isolate passaged one time in Vero cells. cDNA clone derived WNV was used as a control, with virus generated as described elsewhere (Shi et al., 2002). Probes were designed using the DIG Northern Starter Kit, following the manufacturer's instructions (Roche, Mannheim, Germany). Briefly, PCR was performed using primers with T7 promoter sequence on cDNAs derived from WNV infectious clone (pFLWNV). Probes were generated to target the 3' end of the NS-5 region (Table 1, 10296rT7) and the middle of the envelope coding region (Table 1, 2666rT7), and used in conjunction with 9990f and 2326f, respectively. PCR was run for 30 cycles using 2x PCR mastermix (Promega, Madison, WI) after which 4 µl of the resultant amplification product was used as a template for RNA transcription with DIG labeling mix and T7 polymerase. RNA samples were diluted 3:1 in NorthernMax Formaldehyde Load Dye (Ambion, Austin, TX) and run on a 1% agarose gel with 2% formaldehyde, at 60V for 6 hours, in 1X MOPS. RNA was transferred overnight from gel to positively charged nylon membranes (Roche, Mannheim, Germany). RNA was fixed to the membrane by UV-crosslinking on a UV transilluminator for 90 seconds. Membranes were hybridized overnight at 65˚ C in hybridization buffer (5X SSC, 0.1% sarkosyl, 0.02% SDS, 1% block buffer reagent from Northern Starter kit) with 100 ng/ml probe. Immunodetection proceeded according to manufacturer's instructions, except BCIP/NBT phosphatase substrate (KPL, Gaithersburg, MD) was used for detection after final wash step.

Phylogenetic anlaysis. Full genome WNV sequences were downloaded from Genbank on October 14, 2011 (Table 3). Sequences were trimmed to nucleotides 789-10395, because the sequence alignment available for 2774 prior to nt 789 was not clean and aligned manually in BioEdit (Hall, 1999). jModelTest (version 0.1.1) was used to select the best model for nucleotide substitution using Aikake and Bayesian Information Criteria, which was determined to be general time reversible plus gamma distribution of rates (GTR+G)(Posada, 2008). Markov Chain Monte Carlo (MCMC) tree searches of 5 million generations each were run in parallel with sampling one in every 1000 trees using MrBayes version 3.1 (Ronquist and Huelsenbeck, 2003). A 50% majority-rule consensus phylogram is shown based on the last 3750 trees, with posterior probabilities given as values at nodes.

Statistics. Virus production from Vero cells after infection at various MOIs with each treatment was compared using a two way ANOVA in GraphPad (La Jolla, CA). We compared infection, dissemination, and transmission rates in mosquitoes fed full length only or deletion mutant containing virus in a contingency table analysis to $\chi 2$ distribution or using Fisher's exact test. Mosquito titers were log transformed and compared by two tailed t tests. We compared mouse survival rates by log rank tests on Kaplan-Meier survivorship curves, and weight change in mice by group was assessed with repeated measures ANOVA, with Bonferonni post tests to determine comparisons resulting in significant differences.

Acknowledgements

The authors would like to thank Eleanor Deardorff, Doug Brackney, Kathy Hanley, and Lisa Green for technical help and discussions, Adam Aragon from the New Mexico department of health for providing samples. This project has been funded with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health (NIH), Department of Health and Human Services, under grant under grant AI067380 (GE) and contracts HHSN272200900018C (Broad Inst.) and HHSN272200900006C (Broad Inst.) from the U.S. NIH. KP was supported by National Research Service Award Institutional Training Grant 5T32-AI07538-13.

Name	Sequence
DH1F	ACTACTTCGGCTGTGTGAGCT
DH1R	ATGGGCCCTGGTTTTGTGTCT
10296rT7	CTA ATA CGA CTC ACT ATA GGG AGA TCC GAT GAT TGC TCT GAC TT
2666rT7	CTA ATA CGA CTC ACT ATA GGG AGA CGT CCT TCA CTG CTT CCC AGA
3' end	AGATCCTGTGTTCTCGCACCA
qPCR MJ F	ATC CGA GTG CTG GTG AGA CCA AAT
q PCR MJ R	TTC CAA GGG AAG GTG ATG ATG ACG
qPCR MJ probe	/56-FAM/GGA AAG AAC /ZEN/CTA AGC TTA GAA GTG GAG GA/3IABkFQ/

Table 1: Primer and probe sequences

Bloodmeal	Day sampled	n	Number infected	Mutant positive bodies	Number disseminated	Number transmitted
Deletion	14	71	44	7	8	2
Full length	14	41	33	0	12	6
Deletion	21	56	26	5	8	6
Full length	21	24	19	0	10	5

Table 2: *Culex quinquefasciatus* mosquitoes fed on deletion and full length virus

Designation	Strain	Accession #	Host, location, time of sampling
NM05	3337	HM756677	Lorikeet kidney, Albuquerque, August 2005
NM07	2774	N/A	Lorikeet kidney, Albuquerque, August 2007
NY00	WNV3356	EF530047	American crow, Staten Island, NY, October 2000
CA04	COAV689	JF703161	*Culex tarsalis*, California 2004
CA03	Impr1075	DQ080056	*Culex tarsalis*, California 2003
AZ04	N/A	DQ164201	Human, Arizona 2004
TX05	TX5058	JF415929	Blue jay, Texas 2005
TX07	M19433	JF415919	*Aedes albopictus*, Texas 2007
TX09	M20122	JF415928	*Culex quinquefasciatus*, Texas 2009
CO04	04-219CO	DQ431704	Human, Colorado 2004

Table 3: Partial genome sequences used for Bayesian analysis

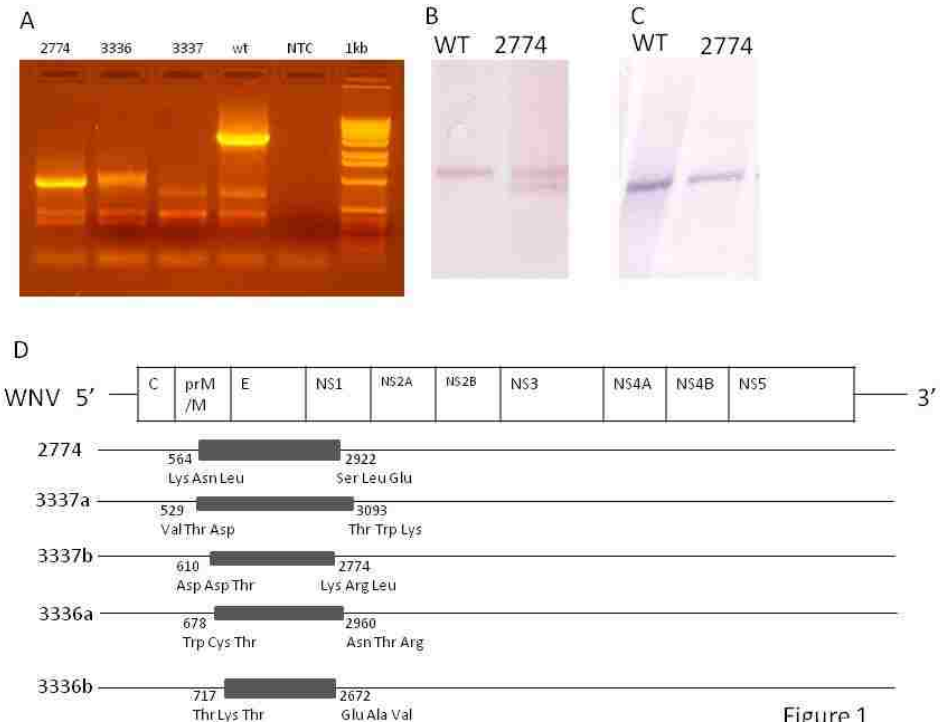


Figure 1

Identification of deletion mutants. A) PCR amplicons produced by primers to the structural coding regions of WNV genome to each of three samples, WT target size is 2.5 kb, actual size observed from isolates is 0.5-0.7 kb, B) Northern blot of RNAs from cells infected with infectious clone derived virus (WT) or isolate 2774 from infected birds, and probe to NS-5. C) Northern analysis of RNAs as in B but with probe to E. D) Genomic locations of deletions found in isolates 2774, 3336, and 3337. At top, diagram of West Nile virus genome, showing structural protein coding regions for a nucleocapsid (C), membrane precursor (prM), and envelope (E) towards the 5' end, and nonstructural proteins (NS1-5) towards the 3' end. Locations of deletions found in isolates 2774, 3337, and 3336 are shown by grey boxes, with numbers at start and finish indicating distance in nucleotide bases from start of genome where deletions begin and end. Three letter codes are given for the resultant amino acid sequence at the deletion sites. For isolates 3337 and 3336 the deletion mutant population was mixed evenly between two variants (a and b).

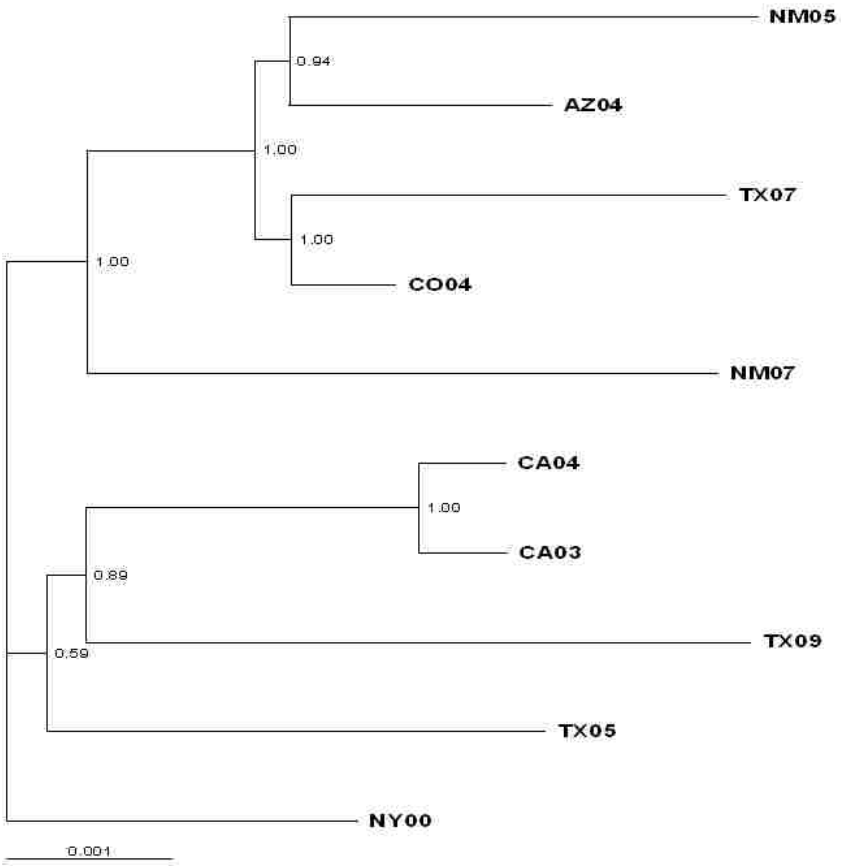


Figure 2

Fifty percent majority-rule consensus phylogram is shown based on partial genome sequences (nt 789-10395) with posterior probabilities given as values at nodes. Strain designations are given in table 3.

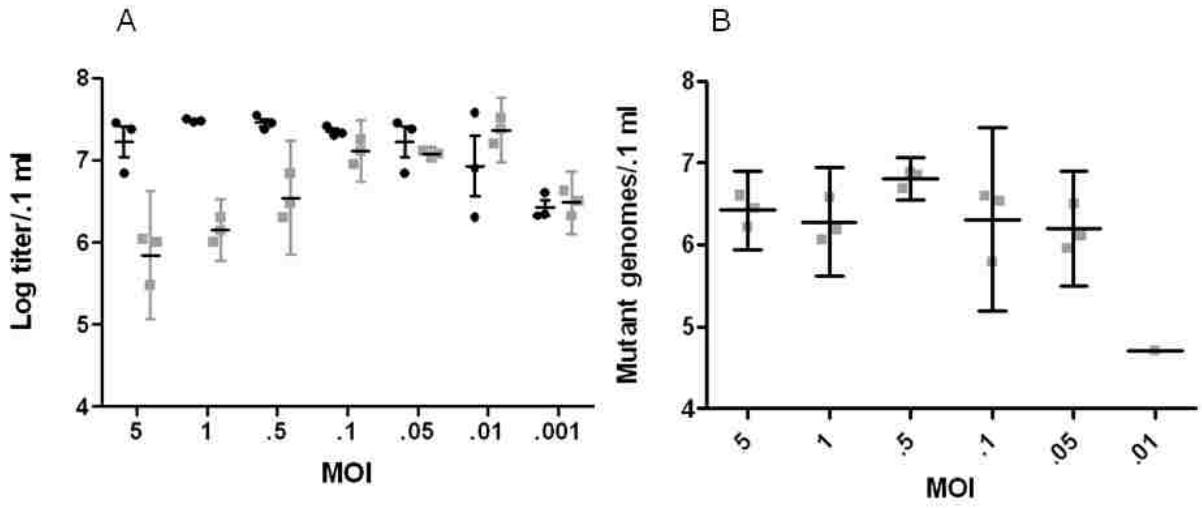


Figure 3
A) Titers produced at 72 h post infection of Vero cells with infectious clone derived virus (WT, black circles) or deletion mutant containing virus (2774, grey squares) at indicated multiplicities of infection. Individual replicates are represented, with the mean and 95% confidence interval given as accompanying bars. B) Deletion mutant genomes detected in supernatants at 72 h post infection of Vero cells with deletion mutant containing virus.

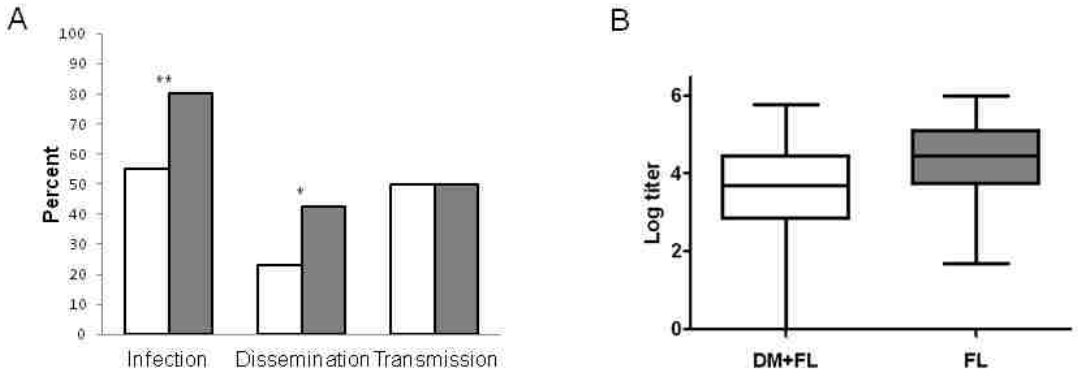


Figure 4
A) Infection, dissemination, and transmission rates for Culex pipiens quinquefasciatus mosquitoes fed on full length only (FL, grey) or deletion mutant containing virus (DM+FL, white). Data from days 14 and 21 are combined, as day of sampling had no significant influence on rates, by chi-squared comparison. **$p<0.005$, *$p<0.05$. Infection was calculated as the percentage of mosquitoes exposed with positive bodies, dissemination rates were calculated as the percentage of mosquitoes with positive bodies containing positive legs, transmission rates were calculated as the percentage of mosquitoes with positive legs containing positive salivary secretions. B) Body titers of infected mosquitoes, as estimated by plaque assay in Vero cells.

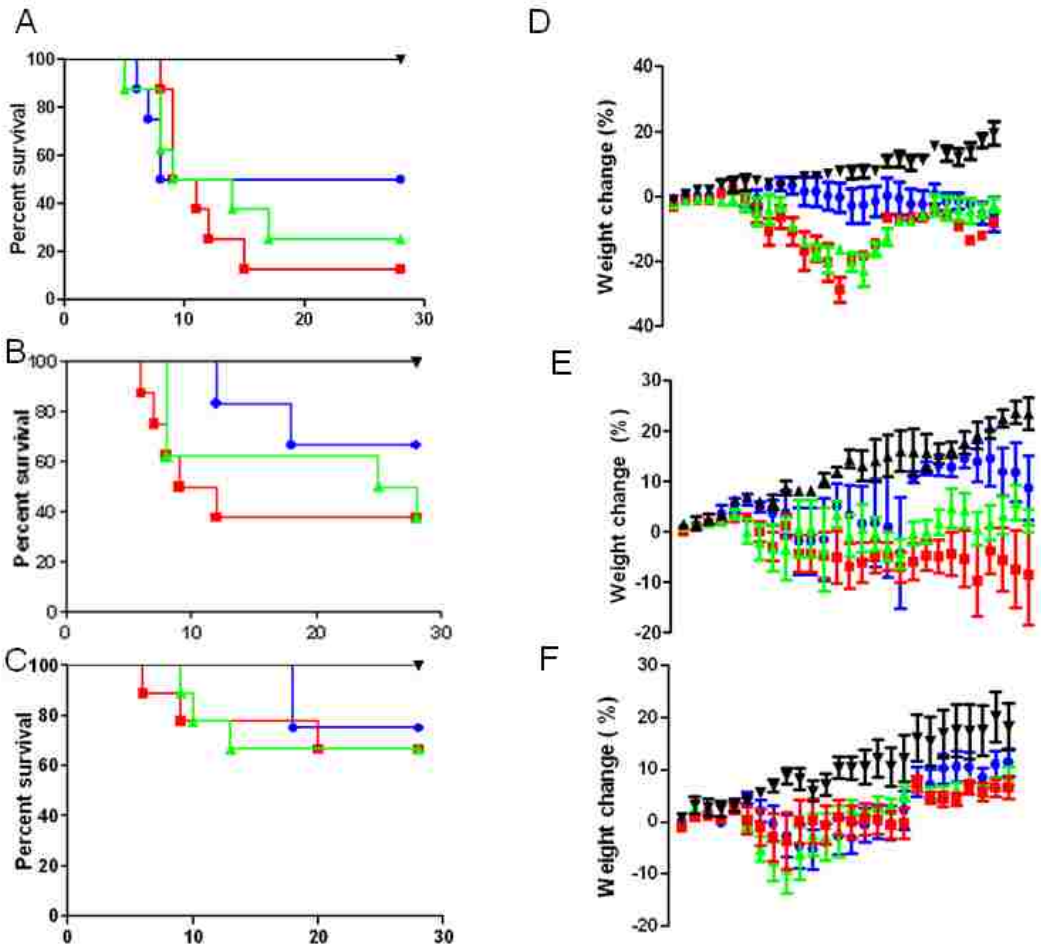


Figure 5

(A-C) Survivorship curves for C3H mice inoculated with 10^2 pfu/mouse (A), C3H mice inoculated with 10^5 pfu/mouse (B), and C57/Bl6 mice inoculated with 10^5 pfu/mouse (C), with deletion mutant containing isolate (red), plaque purified isolate containing full length virus only (green), infectious clone derived full length virus only (blue), or mock inoculated (black). (D-F) Percent weight change in mice measured daily, shown as the average daily percent difference from starting weight for C3H mice inoculated with 10^2 pfu/mouse (D), C3H mice inoculated with 10^5 pfu/mouse (E), and C57/Bl6 mice inoculated with 10^5 pfu/mouse (F) with deletion mutant containing isolate (red), plaque purified isolate containing full length virus only (green), infectious clone derived full length virus only (blue), or mock inoculated (black).

References

Aaskov, J., Buzacott, K., Thu, H.M., Lowry, K., Holmes, E.C. (2006). Long-Term Transmission of Defective RNA Viruses in Humans and Aedes Mosquitoes. Science 311, 236.

Anderson, S.L., Richards, S.L., Tabachnick, W.J., Smartt, C.T. (2010). Effects of West Nile Virus Dose and Extrinsic Incubation Temperature on Temporal Progression of Vector Competence in Culex Pipiens Quinquefasciatus. J.Am.Mosq.Control Assoc. 26, 103.

Armstrong, P.M., Vossbrinck, C.R., Andreadis, T.G., Anderson, J.F., Pesko, K.N., Newman, R.M., et al. (2011). Molecular Evolution of West Nile Virus in a Northern Temperate Region: Connecticut, USA 1999-2008. Virology 417, 203-210.

Artsob, H., Gubler, D., Enria, D., Morales, M., Pupo, M., Bunning, M., et al. (2009). West Nile Virus in the New World: Trends in the Spread and Proliferation of West Nile Virus in the Western Hemisphere. Zoonoses Public Health 56, 357-369.

Brackney, D.E., Beane, J.E., Ebel, G.D. (2009). RNAi Targeting of West Nile Virus in Mosquito Midguts Promotes Virus Diversification. PLoS Pathog 5, e1000502.

Brinton, M.A. (1983). Analysis of Extracellular West Nile Virus Particles Produced by Cell Cultures from Genetically Resistant and Susceptible Mice Indicates Enhanced Amplification of Defective Interfering Particles by Resistant Cultures. J.Virol. 46, 860.

Brinton, M.A. (2001). Host Factors Involved in West Nile Virus Replication. Ann.N.Y.Acad.Sci. 951, 207-219.

Brown, A.N., Kent, K.A., Bennett, C.J., Bernard, K.A. (2007). Tissue Tropism and Neuroinvasion of West Nile Virus do Not Differ for Two Mouse Strains with Different Survival Rates. Virology 368, 422-430.

Debnath, N.C., Tiernery, R., Sil, B.K., Wills, M.R., Barrett, A.D.T. (1991). In Vitro Homotypic and Heterotypic Interference by Defective Interfering Particles of West Nile Virus. J.Gen.Virol. 72, 2705-2711.

Ebel, G.D., Carricaburu, J., Young, D., Bernard, K.A., Kramer, L.D. (2004). Genetic and Phenotypic Variation of West Nile Virus in New York, 2000-2003. Am.J.Trop.Med.Hyg. 71, 493-500.

Hall, T.A. (1999). BioEdit: A User-Friendly Biological Sequence Alignment Editor and Analysis Program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., VandePol, S. (1982). Rapid Evolution of RNA Genomes. Science 215, 1577.

Holmes, E.C. (2009). The Evolutionary Genetics of Emerging Viruses. Annual Review of Ecology, Evolution, and Systematics 40, 353.

Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C. (2002). Rates of Molecular Evolution in RNA Viruses: A Quantitative Phylogenetic Analysis. J.Mol.Evol. 54, 156-165.

Jerzak, G., Bernard, K.A., Kramer, L.D., Ebel, G.D. (2005). Genetic Variation in West Nile Virus from Naturally Infected Mosquitoes and Birds Suggests Quasispecies Structure and Strong Purifying Selection. J Gen Virol 86, 2175-2183.

Komar, N. and G.G. Clark. (2006). West Nile Virus Activity in Latin America and the Caribbean. Revista Panamericana De Salud Pública 19, 112-117.

Lancaster, M.U., Hodgetts, S.I., Mackenzie, J.S., Urosevic, N. (1998). Characterization of Defective Viral RNA Produced during Persistent Infection of Vero Cells with Murray Valley Encephalitis Virus. J. Virol. 72, 2474.

Li, D., Lott, W.B., Lowry, K., Jones, A., Thu, H.M., Aaskov, J. (2011). Defective Interfering Viral Particles in Acute Dengue Infections. PlosONE 6, e19447.

Lindsey, H.S., Calisher, C.H., Mathews, J.H. (1976). Serum Dilution Neutralization Test for California Group Virus Identification and Serology. J.Clin.Microbiol. 4, 503.

Mertens, E., Kajaste-Rudnitski, A., Torres, S., Funk, A., Frenkiel, M.P., Iteman, I., et al. (2010). Viral Determinants in the NS3 Helicase and 2K Peptide that Promote West Nile Virus Resistance to Antiviral Action of 2',5'-Oligoadenylate Synthetase 1b. Virology 399, 176-185.

Murray, K.O., Mertens, E., Desprès, P. (2010). West Nile Virus and its Emergence in the United States of America. Vet.Res. 41, 67.

Posada, D. (2008). JModelTest: Phylogenetic Model Averaging. Mol.Biol.Evol. 25, 1253.

Ronquist, F. and J.P. Huelsenbeck. (2003). MrBayes 3: Bayesian Phylogenetic Inference Under Mixed Models. Bioinformatics 19, 1572.

Shi, P.Y., Tilgner, M., Lo, M.K., Kent, K.A., Bernard, K.A. (2002). Infectious cDNA Clone of the Epidemic West Nile Virus from New York City. J.Virol. 76, 5847.

Thompson, K.A.S., Rempala, G.A., Yin, J. (2009). Multiple-Hit Inhibition of Infection by Defective Interfering Particles. J.Gen.Virol. 90, 888-899.

Tsai, K.N., Tsang, S.F., Huang, C.H., Chang, R.Y. (2007). Defective Interfering RNAs of Japanese Encephalitis Virus found in Mosquito Cells and Correlation with Persistent Infection. Virus Res. 124, 139-150.

Weaver, S.C. (2006). Evolutionary Influences in Arboviral Disease. Curr. Top. Microbiol. Immunol. 299, 285-314.

Wright, T.F., Schirtzinger, E.E., Matsumoto, T., Eberhard, J.R., Graves, G.R., Sanchez, J.J., et al. (2008). A Multilocus Molecular Phylogeny of the Parrots (Psittaciformes): Support for a Gondwanan Origin during the Cretaceous. Mol.Biol.Evol. 25, 2141.

Yoon, S.W., Lee, S.Y., Won, S.Y., Park, S.H., Park, S.Y., Jeong, Y.S. (2006). Characterization of Homologous Defective Interfering RNA during Persistent Infection of Vero Cells with Japanese Encephalitis Virus. Molecules and Cells 21, 112-120.

Zou, G., Puig-Basagoiti, F., Zhang, B., Qing, M., Chen, L., Pankiewicz, K.W., et al. (2009a). A Single-Amino Acid Substitution in West Nile Virus 2K Peptide between NS4A and NS4B Confers Resistance to Lycorine, a Flavivirus Inhibitor. Virology 384, 242-252.

Zou, G., Zhang, B., Lim, P.Y., Yuan, Z., Bernard, K.A., Shi, P.Y. (2009b). Exclusion of West Nile Virus Superinfection through RNA Replication. J.Virol. 83, 11765-11776.

APPENDIX: WEST NILE VIRUS POPULATION GENETICS AND EVOLUTION

Under revision, Infection, Genetics, and Evolution

Kendra N. Pesko and Gregory D. Ebel

Department of Pathology

University of New Mexico School of Medicine

1 University of New Mexico

Albuquerque, NM, USA

87131

Introduction

West Nile virus (WNV, *Flaviviridae: Flavivirus*) has emerged in recent decades as a significant burden to public health in Europe and the Americas. This emergence, in particular the recent invasion of WNV into North America in 1999 and its subsequent spread throughout the new world, has stimulated intense interest in its population genetics and evolution. The dynamics of the WNV epizootic/epidemic in N. America have been of special interest because they provide insight into a longstanding question in evolutionary and invasion biology: What happens when an exotic pathogen is introduced into a naïve environment? Both observational and laboratory studies have therefore been undertaken to determine the modes and direction of virus evolution and examine the evolutionary implications of the host-virus interactions. In this review, we highlight recent advances in research into the population and evolutionary dynamics of WNV and identify key areas for further research.

Molecular biology and replication. WNV is a member of the Japanese Encephalitis virus (JEV) serological complex of the flaviviruses (Calisher et al., 1989). The virion is enveloped, spherical (~40-60nm in diameter) and contains a single copy of the positive-sense RNA genome (Mukhopadhyay et al., 2003, Brinton, 2009). The WNV genome is approximately 11,000nt in length and the translated polyprotein is co- and post-translationally cleaved by viral and host-cell proteases into three structural (capsid C, premembrane prM/M, and envelope E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5). C, M and E are incorporated into the mature virion, while the nonstructural proteins assemble on host cell membranes where they participate

in RNA replication and suppression of the host antivirus response (Brinton, 2009, Brinton, 2001, Westaway et al., 2002, Evans et al., 2011, Avirutnan et al., 2011, Ambrose and Mackenzie, 2011). Overall, the genome organization of WNV, and its protein coding strategy are similar to other flaviviruses.

WNV is believed to enter host cells by receptor-mediated endocytosis that is dependent on an Ig-like fold present in domain III of the E glycoprotein. Virus-containing vesicles enter the endocytic pathways, where acidification leads to a major reorganization of E homodimers into trimers, exposing a hydrophobic peptide (termed the cd loop) contained in the distal portion of domain II of E. Ultimately this reorganization results in fusion of the viral and host cell membranes. Identifying specific host receptors for all flaviviruses has proved difficult and the literature is currently ambiguous on which host-cell molecules are so-called attachment receptors and which, if any, are absolutely required for virus entry. Candidate receptors that have been proposed for WNV include DC-SIGN, DC-SIGNR and $\alpha_v\beta_3$ integrin (Davis et al., 2006, Chu and Ng, 2004b, Chu and Ng, 2004a). In mosquito cells, a c-type lectin is secreted from infected cells and binds virions to enhance uptake involving a phosphatase homolog of human CD45, mosPTP-1 (Cheng et al., 2010). Once viral RNA is released into host cells, it is immediately translated by host machinery. The resulting viral nonstructural proteins assemble on host membranes and replicate the viral genome. Notably, several viral nonstructural proteins are multifunctional and the function of others are poorly defined. Excellent reviews on their roles in flavivirus replication and host cell function have been published recently (Bollati et al., 2009). Mature virions exit cells through the trans-golgi network and are released into the extracellular milieu by exocytosis and/or budding at the

plasma membrane. Thus, the life cycle of WNV within cells is similar to other RNA viruses that replicate cytoplasmically. However, WNV and other arboviruses have evolved the ability to replicate in cells of hosts that are widely taxonomically divergent (i.e. arthropods and vertebrates). This requirement for replication in different host types exerts unique evolutionary and selective pressures on the virus, which are discussed below.

Ecology. Viruses adapt to available ecological niches or they become extinct. A thorough understanding of what constitutes the "niche" for WNV is therefore critical to formulating hypotheses regarding how the virus might evolve in order to maximize its potential to perpetuate. WNV is maintained in nature in an enzootic cycle involving birds and mosquitoes. Although the specific birds and mosquitoes most important for virus perpetuation in any given focus vary locally, they tend to include birds of the order *Passeriformes* and mosquitoes of the genus *Culex*. However, nearly 60 mosquito and 300 bird species have been found infected, and the species of *Culex* mosquito that is most important in a given locality is highly variable. For example, in the Northeastern US, *Cx. pipiens pipiens* is a major vector and appears to be responsible for the vast majority of virus transmission (Bernard et al., 2001). In the central and western US, however, *Cx. tarsalis* is the principal vector, while in southern regions of the US, *Cx. p. quinquefasciatus* is most important (Bell et al., 2006, Molaei et al., 2010, Goldberg et al., 2010, Venkatesan and Rasgon, 2010). In Florida, *Cx. nigripalpus* is the dominant vector (Vitek et al., 2008, Kramer et al., 2008). This pattern is repeated at a global scale, with the dominant Culex mosquitoes in a given locality driving local WNV transmission

(Kramer 2008). Culex species tend to feed mainly on birds in the spring and summer, switching focus to take more mammalian bloodmeals in the fall, when outbreaks of WNV are most likely to occur among humans (Kilpatrick et al. 2006). In addition, several mosquito species not thought to be extremely important in WNV perpetuation, but potentially significant as "bridge" vectors (i.e. species that feed indiscriminantly) have been found infected, including *Ae. albopictus* and *Ae. vexans* (Turell et al., 2002). Several laboratory studies have established the competence of these vectors to transmit WNV (Turell et al., 2005), and field studies have detected both avian and mammalian blood in *Ae. Vexans*, although their relative importance in infecting humans and other hosts is currently unclear (Kilpatrick et al., 2005, Molaei et al. 2006). WNV has also been detected in *Culex pipiens* mosquitoes that have fed on human blood, indicating this mosquito may be the major bridge vector for infecting humans (Hamer et al. 2008). Although WNV may infect taxonomically diverse mosquito species throughout its range, certain Culex species appear to be critically important in WNV perpetuation in each geographic region where it persists.

Similarly, several bird species appear to be capable of generating sufficiently high viremias to infect mosquitoes and contribute to virus perpetuation. American Crow (*Corvus brachyrhynchos*) deaths near the Bronx Zoo in 1999 heralded the arrival of WNV, and these birds have served as useful sentinels since then (Eidson et al., 2001, Kramer and Bernard, 2001). Viremia in Crows reaches extremely high levels ($>10^{10}$ PFU/mL of blood) and mortality is nearly uniform (McLean et al., 2001, Komar et al., 2003). Recently it has become clear that other massively roosting birds, mainly American Robins (*Turdus migratorius*) are important both in enzootic maintenance of WNV in

highly active transmission foci, and in driving a feeding shift in Culex mosquitoes that increases human risk (Kilpatrick et al., 2006). Birds also have been implicated in spreading WNV throughout its distribution. Most importantly, migrating birds have been implicated in transportation of WNV from Africa throughout the Middle East and into Eurasia and within the Americas (Rappole et al., 2006, May et al., 2011, Zehender et al., 2011, Dusek et al., 2009). Clearly a wide variety of birds have been found infected by WNV, but the species most important to virus perpetuation may vary locally.

WNV is capable of being transmitted between a surprisingly large variety of hosts. In contrast, the related Dengue virus (DENV, Flaviviridae, Flavivirus) maintenance is mainly driven by single mosquito and host species (i.e. *Aedes aegypti* and human beings). By comparison, the ability of WNV to act as an ecological generalist is quite clear, and may account, in part, for its dispersal throughout much of the tropical and temperate world. The molecular and/or population mechanisms that form the basis for the relative lack of host-specificity exhibited by WNV are not fully understood, representing a critical area for future research.

Historical perspective

The evolutionary dynamics of WNV are of particular interest because of the emergence of the virus as a significant health burden in the last 20 years. Originally isolated in 1937 from the blood of a patient with fever in the West Nile district of Uganda (Smithburn et al., 1940), the first outbreaks of WNV disease were associated with relatively few cases, mild disease and rural settings (Hayes, 2001). Strikingly, an

outbreak in Romania that occurred in 1996 and 1997 involved over 500 reported cases, with a case-fatality rate of approximately 10% (Tsai et al., 1989, Tsai et al., 1998). This outbreak was also striking in that it occurred in a temperate urban region. Shortly thereafter, epidemics were reported in the south of Romania and in the Volga delta region of Russia. Additional recent epidemics have been reported in Russia, Israel, Greece, France, Hungary, Italy and others (Platonov, 2001, Bin et al., 2001, Papa et al., 2010, Balenghien et al., 2006, Depoortere et al., 2004, Kutasi et al., 2011, Bakonyi et al., 2006, Monaco et al., 2011). Generally, these outbreaks occurred in delta regions of major rivers including the Volga, Rhone and Danube. Comprehensive reviews of WNV in Europe have been published recently (Hubalek and Halouzka, 1999, Zeller and Schuffenecker, 2004).

In 1999, WNV was introduced into North America in the New York City area, resulting in an equine and avian epizootic, and associated human infection, morbidity and mortality (CDC, 1999). The virus rapidly spread throughout the mainland US and into Canada, Mexico, and as far south as Argentina. As has been amply noted, the introduction of WNV at a precisely defined time and place provided a relatively unique opportunity to prospectively observe the adaptation of an exotic RNA virus to an essentially naïve ecosystem. Accordingly, several studies have been conducted to examine the evolution of the virus since its introduction (Anderson et al., 2001, Ebel and Dupuis, 2001, Ebel et al., 2004, Beasley et al., 2003, Davis et al., 2005, Bertolotti et al., 2007, McMullen et al., 2011, Armstrong et al., 2011). Several molecular epidemiologic studies have examined nucleotide sequence data from WNV strains found in birds, mosquitoes and human beings. The most recent of these are discussed in detail below

and others are reviewed elsewhere (Ebel and Kramer, 2009). The ability of WNV to act as an ecological generalist, in combination with recent increases in intercontinental travel and trade, appear to have facilitated its emergence on a global scale.

Taxonomy and Classification

WNV is classified as a member of the Japanese Encephalitis complex of the Flaviviruses on the basis of serological cross-reactivity (Calisher et al. 1989). Within WNV, two major lineages (Lineage I and II) are currently accepted, with several additional lineages that differ from one another by 5-25% recently proposed (Vazquez et al., 2010, Bondre et al., 2007). Lineage I is distributed throughout much of the world, and is further subdivided into several clades, one of which includes NY-99 (clade Ia), the genotype introduced to the US in 1999, another includes Kunjin virus (clade Ib), a variant of West Nile virus endemic to Australia (Ebel and Kramer, 2009, Lanciotti et al., 1999, May et al. 2011). Lineage II was thought to be restricted to sub-Saharan Africa until recently. Since 2004, lineage II has been associated with outbreaks of West Nile virus in Western and Eastern Europe, and appears to have established endemic cycles in Spain and Greece (Papa et al., 2010, Bakonyi et al., 2006, Vazquez et al., 2010, Papa et al., 2011). Lineage III, also known as "Rabensburg virus," is represented by several isolates made from the same region of the Czech Republic in 1997 and 1999 from *Culex pipiens* mosquitoes, and 2006 from a pool of *Ae. rossicus* (Bakonyi et al., 2005, Hubalek et al., 2010). Lineage IV encompasses numerous isolates made in Russia, first detected in 1988 from a *Dermacentor* tick, and since isolated from mosquitoes and frogs in 2002 and 2005 in Russia (May et al. 2011). Lineage V comprises thirteen isolates from India, collected

from humans and *Culex* mosquitoes from the 1950s through 1980, which differ from other West Nile lineages by 20-25% at the nucleotide level (Bondre et al., 2007). Recent publications show these strains as basal to lineage I, comprising an independent cluster, lineage Ic (May et al., 2011). An additional, putative sixth lineage has been isolated in Spain from a pool of *Culex pipiens* mosquitoes, and appears to be most closely related to lineage IV WNV (Vazquez et al., 2010). Additionally, Koutango virus (KOU), a Flavivirus isolated in Senegal, may represent a seventh lineage, as it is ~25% identical to other WNV isolates, although it is currently categorized as a separate species (King et al. 2011). Human infection by KOU has not been reported, and its serological relationships to established WNV strains and transmission cycle are unclear, although partial cross neutralization with WNV and KUN has been shown(Charrel et al., 2003, Calisher et al. 1989). The distribution of all described lineages of WNV is shown in figure 1, by country where isolations have been made. The diversity of proposed WNV lineages worldwide reflects the diversity of the vectors involved in virus perpetuation and suggests that WNV or closely related agents have been introduced, and adapted to local transmission cycles on several occasions.

Taxonomic relationships are not entirely clear, and require reevaluation, especially with the recent proposal of so many new WNV lineages. In terms of nucleotide identity, they may be too disparately related to qualify as part of the same virus species, as the cutoffs proposed by researchers are >84% pairwise sequence identity (Kuno et al., 1998) or >79% for inclusion within a species (Ebel and Kramer, 2009, Charrel et al., 2003), although identity limits for inclusion within a lineage or species are generally arbitrary. According to the first estimate, lineage II WNV would have to be separated

into its own species, as it has between 17-20 % pairwise distance from lineage I, while the second pairwise distance might prevent proposed lineages III-VI from inclusion in WNV, as most show greater than 21% pairwise distance from the first two lineages at whole and partial genome levels (Vazquez et al., 2010, Bondre et al., 2007). These lineages appear to show some cross reactivity (Bondre et al., 2007, Bakonyi et al., 2005), may persist in similar transmission cycles, as most have been isolated from mosquitoes and birds, and appear to form a monophyletic clade when examined alongside Japanese encephalitis virus and Usutu virus (from the same serogroup) but a more systematic examination of relationships between the different lineages is warranted, as they have not been universally accepted (Ebel and Kramer, 2009, Vazquez et al., 2010). The relationships between these different lineages are further elucidated by figure 2, a phylogram based on the complete coding sequences for WNV lineages available in Genbank.

Lineage I has been subject to the most intensive study. It is now distributed worldwide, and includes the genotype introduced to the US in 1999 (NY99). Some genotypes appear to be more pathogenic than others, for example NY99 shows enhanced pathogenesis in birds (see below), whereas Kunjin virus (clade Ib) is associated with attenuated infection and decreased neuroinvasion (Brault et al., 2007), (Daffis et al., 2011). Lineage II is mainly associated with less severe disease, and less frequent neuroinvasion. However, recent reports describe encephalitis produced by infection with lineage II strains in both humans and horses in South Africa (Venter and Swanepoel, 2010, Venter et al., 2009). Lineage III has only been isolated from mosquitoes, and did not produce mortality in adult mice infected subcutaneously, intraperitoneally, or

intracranially (Hubalek et al., 2010). Lineage V viruses from India are also associated with lower virulence (Davis et al., 2005, Bondre et al., 2007). WNV has thus clearly adapted to a wide array of transmission cycles and environments worldwide. This process of migration and adaptation to these environments has produced the currently observed lineages. Additional studies are required in order to define differences in virulence, neuroinvasiveness, natural hosts and vectors, and basic ecology for each putative lineage.

Molecular epidemiology

Upon its introduction to the United States, WNV was initially recognized by sequence comparisons and phylogenetic analysis (Lanciotti et al., 1999). The genotype introduced to the New World, dubbed NY99 for its initial isolation in New York in 1999, is most closely related to isolates made in Israel in 1998 and Hungary in 2003 (Zehender et al., 2011, Lanciotti et al., 1999, Jia et al., 1999). Initial sequence analysis of the WNV strains isolated during the first two years in New England showed a remarkable amount of genetic conservation, indicating a single point of introduction and very little diversification in WNV populations during this time period (Anderson et al., 2001, Ebel and Dupuis, 2001, Lanciotti et al., 1999) reviewed in (Kramer et al., 2008, Ebel and Kramer, 2009). Subsequently, an additional subtype of WNV, WN02, with an amino acid substitution in the envelope protein, A159V, was detected in samples isolated in Texas (Beasley et al. 2003). From 2001 to 2003, WN02 rapidly displaced NY99, becoming the dominant genotype in North America (Ebel et al., 2004, Davis et al., 2005). WN02 strains require a shorter extrinsic incubation period in mosquitoes, which appears

to be the mechanism for its increased fitness relative to NY99 (Ebel et al. 2004, Moudy et al. 2007). Thus, shortly after WNV was introduced into North America, the process of evolution led to increases in the basic reproductive rate of this pathogen.

As WNV became established throughout North America, the genetic diversity present in different types of data sets has led to insights into its emergence and expansion. Studies have found increased genetic diversity in mosquitoes relative to birds (Bertolotti et al., 2007, Amore et al., 2010), perhaps due to different selective pressure from the immune pathways used by these different hosts, which will be discussed in the section on genetic diversity below (Brackney et al., 2009). Genetic diversity and therefore estimated virus population size appeared to initially increase yearly after introduction to the US, although studies suggest this may be leveling off as WNV becomes established endemically (Bertolotti et al., 2007, Amore et al., 2010, Snapinn et al., 2007, Bertolotti et al., 2008).

Several studies (Armstrong et al., 2011)showed a lack of geographical partitioning among sequences, especially those that examined sequence data from isolates sampled immediately after the introduction of WNV to novel environments and relied primarily on envelope sequences (Bertolotti et al., 2007, Bertolotti et al., 2008, Davis et al., 2007). Recent studies, relying on full genome sequences and encompassing samples taken over a number of years after introduction of WNV uncovered more evidence for geographical structure to samples (McMullen et al., 2011, Armstrong et al., 2011, Herring et al., 2007, Grinev et al., 2008). A recent analysis indicates sequences from the envelope coding region may not be the most phylogenetically informative, and suggests NS3 or NS5 may be better partial sequences for reconstructing the phylogenetic

relationships between different isolates, and can provide reconstructions that more closely resemble those resulting from whole genome sequences (Gray et al., 2010). Several distinct genetic variants of WNV have arisen in certain geographical areas, such as Texas (McMullen et al., 2011, Davis et al., 2004). One attenuated genetic lineage seems to have become extinct after being detected over the course of two years (Davis et al., 2005, Ebel and Kramer, 2009, Davis et al., 2004, Ebel, 2010). Another distinctive genotype, SW/WN03, contains several amino acid changes relative to other WN02 and is recently reported to be spreading through numerous states, although the phenotype associated with this new genotype has not been characterized (McMullen et al., 2011). It may be that more genetic changes accumulated in these WNV populations after it adapted to local transmission cycles.

Recent studies using full genome sequences of WNV from isolates made globally have uncovered phylogeographical influences on clade 1a distribution (May et al., 2011, Zehender et al., 2011). This clade seems to have a common ancestor that existed in sub-Saharan Africa in the early 20th century, which had multiple migrations to both Western and Eastern European countries in the 1970s and 1980s, and single introductions to India and Australia around the same time (May et al., 2011, Zehender et al., 2011). The patterns of distribution from Africa to Europe seem to follow white stork migration routes, indicating a possibly important role for this bird species in the spread of WNV into that continent (Zehender et al., 2011). Other mosquito borne Flaviviruses have also apparently originated in Africa, including yellow fever and dengue virus (Bryant et al., 2007, Gaunt et al., 2001, Holmes and Twiddy, 2003).

Arboviruses are unique in that they require replication in taxonomically divergent hosts – vertebrates and invertebrates (Weaver, 2006). This requirement is thought to restrict the amount of mutation that can occur in arboviruses, relative to single host viruses (Jenkins et al., 2002). Experimental studies have shown lower mutation rates in viruses serially passaged in alternating hosts, relative to those passaged in a single host type (Jerzak et al., 2007, Coffey et al., 2008, Coffey and Vignuzzi, 2011). To date, numerous studies have shown purifying or negative selection is dominant in arbovirus populations, including West Nile virus (Bertolotti et al., 2007, McMullen et al., 2011, Armstrong et al., 2011, Amore et al., 2010, Bertolotti et al., 2008, Jerzak et al., 2005). In WNV phylogenetic analyses, only a few genetic changes have been identified that appear to be the subject of positive selective pressure. These include the amino acid residue associated with increased pathogenesis among North American birds, NS3 T249P and a mutation to NS4A, or the 2K protein (V135M or V9M) (Armstrong et al., 2011, Brault et al., 2007). The valine to methionine mutation in NS4A/2K is associated with OAS1b resistance, resistance to the flavivirus specific antiviral lycorine, and ability to overcome superinfection exclusion in replicon containing cell lines, which appears to be related to enhanced viral RNA synthesis (Mertens et al., 2010, Zou et al., 2009a, Zou et al., 2009b). Adaptive evolution has also been detected at amino acid sites: E-V431I, NS2A-A224V/T, NS4A-A85T, NS5-K314R, and NS5-R422K although the functional significances of these sites are unclear (May et al., 2011, McMullen et al., 2011). Thus, several of the encoded WNV proteins are subject to positive selection that may lead to increased transmission efficiency and the likelihood for perpetuation in different transmission cycles.

Synonymous changes to the WNV genome could also impact its pathogenesis and evolution. Numerous synonymous changes were associated with the new genotype, WN02, and although some of these are assumed to have become fixed by association with other mutations that might confer a selective advantage, they could also exert an effect through codon bias or changes to the RNA genomic structure. The 5' and 3' untranslated regions are well conserved and have essential roles during viral replication (Khromykh et al., 2001, Zhang et al., 2008). Additional studies have shown that other RNA genome structures present in the capsid coding region can operate to upregulate flaviviral replication (Tuplin et al., 2011, Clyde and Harris, 2006). Additionally, codon bias in flaviviruses reflect the host usage (vertebrate only and alternating seem to display vertebrate codon biases, invertebrate only have more invertebrate bias), so examination of codon bias for a given virus can provide insight into its evolutionary history (Schubert and Putonti, 2010). Additional studies are required in order to determine the extent to which nonsynonymous variation impact RNA genomic structure in a way that influences WNV phenotype.

Within-host population dynamics

Molecular epidemiologic studies such as those discussed in the preceding section have provided insights into the selective forces that act on WNV and shown clearly that the virus is a dynamic, evolving entity with the capacity to adapt to a wide range of hosts and environments. These findings have stimulated studies aimed at understanding the viral population genetic mechanisms that account for this, and to assess whether the two very different kinds of host required for WNV perpetuation (mosquitoes and birds) influence the WNV population in different ways. Early studies suggested that within hosts, WNV

forms a genetically complex distribution of mutants that vary in their degree of nucleotide divergence from the population consensus sequence. Further, Jerzak et al (2005) showed that whereas WNV populations in naturally infected birds are relatively genetically homogeneous and purifying selection is strong, in field collected WNV infected mosquitoes they are very diverse, and purifying selection seems to be relaxed. The observations were supported by a series of laboratory studies that passed WNV in colonized mosquitoes and chickens (Jerzak et al., 2007), and cultured cells (Ciota et al., 2007). Importantly, the mosquito passed virus was inoculated intrathoracically and whole mosquitoes were triturated to obtain passed WNV, bypassing putative transmission barriers in the midgut and salivary glands (Hardy et al., 1983, Ciota et al., 2008). A highly similar study conducted using virus obtained from mosquito saliva failed to confirm these results raising the possibility that infection of, or escape from salivary glands might constitute a population bottleneck in the WNV system(Ciota et al. 2008, Ciota and Kramer, 2010) . Nonetheless, several studies have clearly established that mosquitoes and birds exert different evolutionary pressures on WNV.

The mechanistic basis for this difference has been addressed from a variety of perspectives. First, vertebrates and invertebrates respond to virus infections differently. In vertebrates, the earliest responses to infection by RNA viruses are dominated by type I interferon (IFNα/β). This response is triggered when RIG-I senses dsRNA in host cell cytosol, initiating signaling cascades that ultimately result in an antiviral state in the cell (reviewed in (Daffis et al., 2009)). Therefore, in vertebrates, WNV may be required to essentially "outrun" the antiviral state in infected individuals. This would result in strong purifying selection that has been observed after virus replication in these hosts (Ding,

2010), (Jerzak et al., 2007), where presumably all or nearly all nonsynonymous mutation results in genomes of diminished fitness.

In contrast, insects respond to virus infection mainly through RNA interference (RNAi), which is also triggered by dsRNA within cells (Reviewed in (Ding, 2010)). Ultimately, virus-derived small-interfering RNAs (viRNAs) are loaded into the RNA induced silencing complex (RISC) to degrade target viral RNA in a sequence-specific manner. Therefore, the antiviral state in mosquito cells seems to drive WNV diversification through a mechanism akin to negative, frequency-dependant selection, wherein rare genotypes (i.e. those that do not match common guide sequences loaded into the RISC) are favored because they are less efficiently degraded (Brackney et al., 2009). The precise relationship between this mechanism and the observed lack of purifying selection in mosquitoes has not been resolved or adequately addressed, and may represent two sides of the same coin. Overall, WNV population biology seems to be dominated by largely opposing forces that exist within its natural transmission cycle. Specifically, WNV undergoes alternating cycles of genetic expansion in mosquitoes that generates novel genotypes, and purification in birds that ensures that high fitness is maintained.

Other forces that influence WNV genetic diversity also have been examined recently. Population bottlenecks can stochastically reduce population diversity and lead to fitness declines through the action of Muller's ratchet (Duarte et al., 1992). Convention holds that in natural transmission cycles, arboviruses undergo population bottlenecks as they pass through mosquitoes, where they seem to sequentially infect the epithelium of the mosquito midgut, peripheral tissues and ultimately the salivary glands, from which they are released into salivary secretions that are inoculated during mosquito feeding (Hardy et

al., 1983). Such population bottlenecks have been described for alphaviruses and flaviviruses. Studies examining early mosquito infection by Venezuelan equine encephalitis virus (VEEV; Togaviridae, Alphavirus) and WNV demonstrated that only a few midgut cells are susceptible to infection, suggesting that anatomical bottlenecks may reduce genetic variability (Smith et al., 2008). Conversely, identical non-consensus WNV genomes have been detected in intrahost populations infecting birds in a single transmission focus, suggesting that population bottlenecks may not be as restrictive as had been assumed (Jerzak et al., 2005), and defective DENV genomes appear to perpetuate in transmission cycles through complementation (Aaskov et al., 2006). Supporting this, Brackney et al. recently failed to document significant population bottlenecks during infection of *Cx. quinquefasciatus* mosquitoes by WNV (Brackney et al., 2011). It may be that the importance of bottlenecks during arbovirus transmission is a function of the specific virus-host system under study, and not consistent across systems.

Genetic correlates of pathogenesis and fitness

Molecular genetic and phenotypic studies of WNV mutants and engineered clones have revealed multiple genetic variations correlated with increased or decreased pathogenicity. WNV was long thought to be a less pathogenic flavivirus, with sporadic epidemics producing little or no mortality in human populations up until the early 1990s (Hayes, 2001). The recent introduction of WNV to the United States was marked by large die offs in bird populations, and a wave of epidemic cases among humans (Murray et al., 2010a). Studies have identified a single amino acid substitution in the NS3 helicase coding region, T249P, that increased morbidity and viral load in American crows, and

appeared to be under selective pressure in areas with multiple genotypes present (Brault et al., 2007, Brault et al., 2004). After establishment of this initial pathogenic strain of WNV across the US, phylogenetic analysis of WNV sequences detected a new genotype, WN02, which displaced the initial strain NY99 in less than 4 years (Ebel et al., 2004, Davis et al., 2005). This new genotype had a single amino acid substitution in the envelope coding protein, V159A, that significantly decreased the extrinsic incubation time from virus infection until transmission by *Culex pipiens* mosquitoes important vectors in the northeastern United States (Moudy et al., 2007, Kilpatrick et al., 2008). This mutation occurs nearby the envelope glycosylation motif for WNV, which is at nucleotide positions 154-156 in the envelope coding sequence.

Envelope protein glycosylation sites are conserved throughout the genus Flavivirus, although natural variation in glycosylation is present in populations of WNV (Adams et al., 1995, Berthet et al., 1997, Shirato et al., 2004, Hanna et al., 2005). An N linked glycosylation site at position 154 in the envelope protein has been associated with increased neuroinvasiveness for WNV in mice, and increased virulence and viremia in young chicks (Shirato et al., 2004, Beasley et al., 2005, Murata et al., 2010). Envelope protein glycosylation is also necessary for efficient transmission by *Culex pipiens*, *Cx tarsalis*, and *Cx quinquefasciatus*, but not *Cx pipiens pallens*, thus it influences vector competence in a species specific way (Murata et al., 2010, Moudy et al., 2009). Glycosylation patterns from virus propagated in insect versus vertebrate cells also seem to influence the ability of envelope protein to modulate innate immune response, and leads to different patterns of infectivity and propagation in different cell types, thus the role of glycosylation is also host specific (Hanna et al., 2005, Arjona et al., 2007).

The mechanism behind envelope protein glycosylation and modulation of WNV activity could be related to a number of different phenomena. The ability of WNV envelope to suppress dsRNA activated innate immune response is dependent on glycosylation status, which leads to increased inflammatory cytokine production for cells infected with virus lacking glycosylation (Arjona et al., 2007). Envelope glycosylation status influences ability to survive in lower pH environments (Beasley et al., 2005, Langevin et al., 2011). Genomes lacking the envelope glycosylation site have decreased replication, which may be related to budding of mature virions from the lumen of the endoplasmic reticulum rather than the plasma membrane (Berthet et al., 1997, Shirato et al., 2004, Li et al., 2006). Abolishing the N linked glycosylation site on WNV envelope also influences receptor interactions, as it decreases greatly the ability of WNV to bind DC-SIGNR (Davis et al., 2006). One or a combination of these mechanisms, or perhaps a mechanism yet to be uncovered may explain the increased virulence of WNV with envelope N-linked glycosylation.

Strains associated with greater neuroinvasiveness and pathogenesis in mice and humans tend to be better controllers of interferon mediated responses (Daffis et al., 2011, Daffis et al., 2009). Numerous WNV proteins may modulate the interferon signaling cascade in vertebrate hosts, including all nonstructural coding proteins (NS1: Wilson et al., 2008, NS2A/B: Liu et al. 2006, NS3: Liu et al. 2005, NS4A/B: **Muñoz-Jordán** et al. 2003 NS5: Laurent-Rolle et al., 2010; Reviewed, Diamond et al., 2009, Samuel and Diamond, 2006, ,)). With the advent of reverse genetics, recent studies have determined distinct amino acid changes in certain residues that are correlated with changed ability to control host immune responses. For example, a single residue at position 653 in NS5

appears to be responsible for the enhanced ability of NY99 like viruses to suppress interferon response. In North American genotype 1a viruses, position 653 of NS5 is a phenylalanine, whereas in the less pathogenic Kunjin virus, this position is a serine, and if these residues are switched through reverse genetics, suppression is enhanced for Kunjin and depleted for NY99 (Laurent-Rolle et al., 2010). Another example comes from studies of the host factor OAS1b which appears to confer natural resistance to WNV (Samuel and Diamond, 2006, Lucas et al., 2003). Virus cultivated in the presence of OAS1b can circumvent this factor by mutating at several residues, including NS3-S365G, which seems to lower the requirement of ATP for the ATPase dependent cleavage activity of this protease, and 2K-V9M, which generally enhances viral RNA synthesis (Mertens et al., 2010). Virus strains with higher rates of replication may be positively selected (Armstrong et al., 2011). Studies like this that correlate genotype with phenotype, and determine the underlying mechanisms, should expand our understanding of virus pathogenesis and the forces shaping the emergence of pathogenic phenotypes.

Attenuated genotypes of WNV emerged during the course of its spread across the United States. In Texas, a number of small plaque variants of WNV that displayed reduced neuroinvasiveness in mice were detected in 2003 (Davis et al., 2004). Comparison to NY99 strain followed by introduction of similar mutations into an infectious clone identified a combination of mutations to NS4, NS5, and the 3' UTR as being necessary for the attenuated phenotype found in one bird sample (Davis et al., 2007, Davis et al., 2004). Further analysis of other samples from the same region found that different amino acid substitutions in these strains appear to confer attenuation, indicating multiple pathways towards attenuated phenotypes (May et al., 2010). A single

amino acid substitution in the central portion of NS4B, C102S, was enough to attenuate neurovirulence in mice (Wicker et al., 2006). Lineage II WNV has been associated with encephalitis and more severe disease only rarely, but comparison of the strains isolated from patients with more severe disease to less virulent strains indicates an enhanced role for NS proteins in determining virulence, relative to structural proteins, similar to findings for lineage I (Botha et al., 2008).

Conclusions and future research directions

In recent decades, WNV has emerged as one of the most intensely studied arthropod borne viruses. The spread of WNV throughout North America and periodic outbreaks in Eastern and Western Europe have enhanced worldwide interest in understanding viral, host, and ecological factors that result in outbreaks. Due to the availability of an increasing number of complete genome sequences (as of October 12, 2011, 479 are currently available on GenBank) major advances have occurred in understanding its molecular epidemiology, particularly in North America. Collectively, these studies have clearly demonstrated that WNV is a dynamic virus population that is able to perpetuate in and adapt to a wide range of ecological settings through mutation and selection. In addition, since the virus transmission cycle is relatively tractable in the laboratory using colonized mosquitoes and captive wild or domestic birds, experimental studies have provided insights into how different components of the transmission cycle influence virus population genetics and evolution. These studies have shown that mosquitoes drive WNV diversification through their innate antivirus response and because purifying selection is weak in these hosts, provide a reservoir of genetic diversity that allows the

observed adaptation. In contrast, infection of birds ensures that the resulting variants are of high fitness through strong purifying selection. The processes that underpin WNV evolution are thus beginning to be understood in greater detail.

Several important areas for future research, however, remain. First, persistent infection of WNV in humans and vertebrates is emerging as a significant health issue in particular regions of the US (Murray et al., 2010b). The extent to which viral genetic and population determinants influence this has not been adequately addressed. In cell culture, establishment of persistent infection with Flaviviruses can occur with subgenomic replicons or the development of defective interfering particles (DIPs) (Zou et al., 2009b, Yoon et al., 2006). A second pressing matter is a reevaluation of the serological relationships within the Japanese encephalitis serogroup. Molecular genetic studies have proposed numerous lineages of WNV beyond the traditional two lineages recognized previously. The basic biology, transmission cycle, host range and pathogenecity of these putative lineages should also be studied further. Additionally, numerous amino acid residues may be under positive selection, but the roles of these residues are still unclear. Reverse genetic studies should be undertaken to determine the influence of these changes on WNV biology. With the advent of new sequencing technologies, our ability to design and conduct experiments into immune responses of hosts to viral infection and viral population biology has greatly increased. The role of RNAi in generating viral diversity, the potential bottlenecks associated with the WNV transmission cycle, and the interaction of individuals within a viral population may all become better understood with deep sequencing approaches. Finally, more collaboration between people studying ecology,

epidemiology, molecular genetics, and pathology of WNV could lead to greater insight into its overall biology.

We have highlighted major advances in WNV biology over the past decade, including understanding of host specific selective pressures on viral populations, genotypic correlates with pathogenic phenotypes, and phylogenetic relationships between different lineages, strains, and genotypes. Molecular epidemiology studies continue to elucidate the spread and evolutionary change that is ongoing in WNV populations.

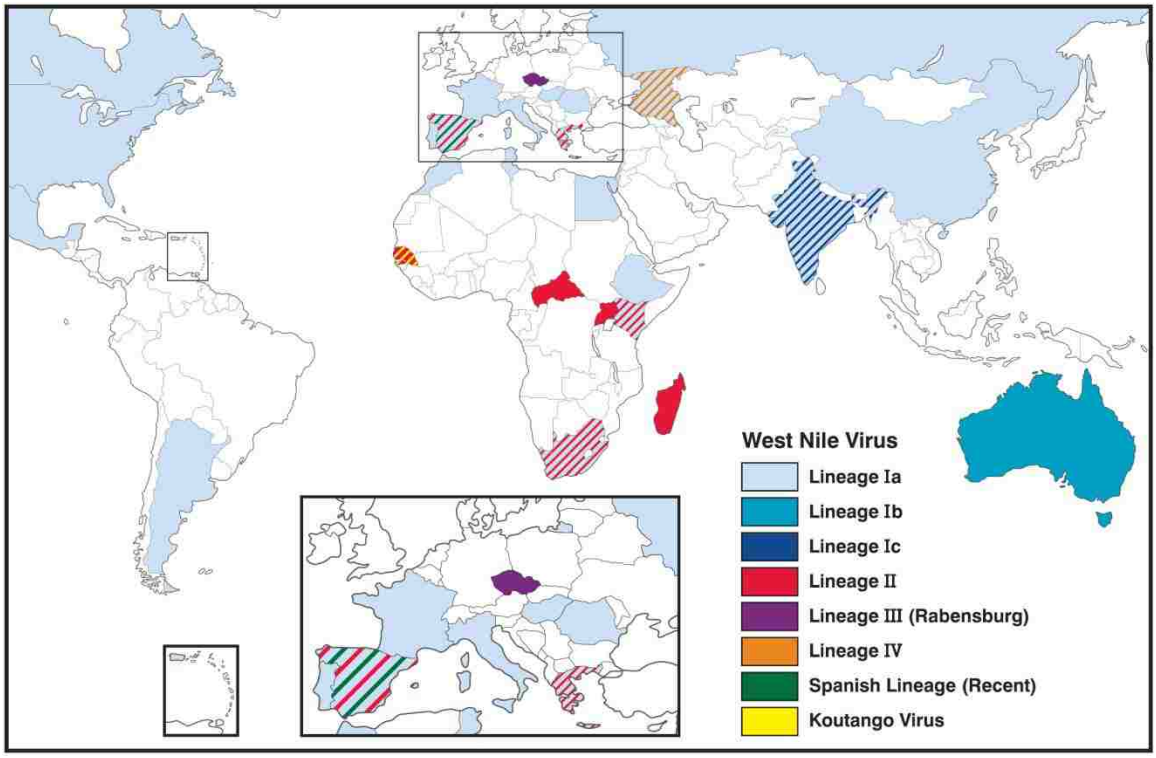


Figure 1. Worldwide map with countries where West Nile virus has been isolated colored as follows: Lineage Ia in light blue; lineage Ib in medium blue; lineage Ic in dark blue, lineage II in red, lineage III or "Rabensburg" virus in purple, lineage IV in orange, recent Spanish lineage (Vazquez et al. 2010) in green, and Koutango virus is colored yellow. Hatched coloring indicates more than one lineage has been isolated from that country. Lineage I distribution is adapted from May et al. 2011, other lineage isolates adapted from Charrel et al. 2003, Vazquez et al. 2010. Blank world map is adapted from: http://edit.freemap.jp/en/trial_version/edit/world.

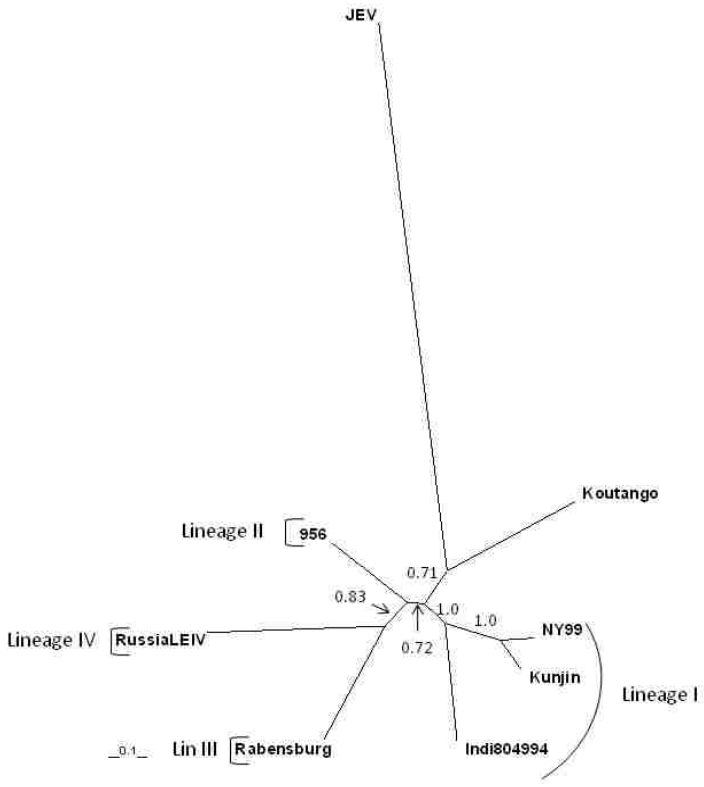


Figure 2. Radial phylogram showing relationships between different lineages of WNV. Complete coding sequences were downloaded from Genbank and aligned manually in BioEdit. Strains used and accession numbers are as follows: JEV, NC_001437; NY99, lineage Ia, NC_009942; Kunjin, paKUN, lineage Ib, AY274505.1; Indi804994, Indian lineage Ic, DQ256376.1; 956, lineage II, NC_001563; Rabensburg, lineage III, AY765264.1; RussianLEIV, lineage IV, strain Krnd88-190, AY277251.1; Koutango virus, EU082200.1. Bayesian phylogeny is shown, generated with MrBayes 3.1.2 run with a general time reversible (GTR) model with gamma shaped rate variation and invariable sites (Ronquist and Hulsenbeck, 2003). Two Markov chain Monte Carlo (MCMC) tree searches of 5,000,000 generations each were run in parallel with sampling one in every 1000 trees. Radial 50% majority-rule consensus tree is shown based on the last 3750 trees. Posterior probabilities are given as numbers at each node.

References

Aaskov, J., Buzacott, K., Thu, H.M., Lowry, K., Holmes, E.C. (2006). Long-term transmission of defective RNA viruses in humans and Aedes mosquitoes. Science 311, 236.

Adams, S., Broom, A., Sammels, L., Hartnett, A., Howard, M., Coelen, R., Mackenzie, J., Hall, R. (1995). Glycosylation and antigenic variation among Kunjin virus isolates. Virology 206, 49-56.

Ambrose, R.L. and J.M. Mackenzie. (2011). West Nile virus differentially modulates the unfolded protein response to facilitate replication and immune evasion. J. Virol. 85, 2723.

Amore, G., Bertolotti, L., Hamer, G.L., Kitron, U.D., Walker, E.D., Ruiz, M.O., Brawn, J.D., Goldberg, T.L. (2010). Multi-year evolutionary dynamics of West Nile virus in suburban Chicago, USA, 2005-2007. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 365, 1871-1878.

Anderson, J.F., Vossbrinck, C.R., Andreadis, T.G., Iton, A., Beckwith, W.H., Mayo, D.R. (2001). A phylogenetic approach to following West Nile virus in Connecticut. Proceedings of the National Academy of Sciences 98, 12885.

Arjona, A., Ledizet, M., Anthony, K., Bonafé, N., Modis, Y., Town, T., Fikrig, E. (2007). West Nile virus envelope protein inhibits dsRNA-induced innate immune responses. The Journal of Immunology 179, 8403.

Armstrong, P.M., Vossbrinck, C.R., Andreadis, T.G., Anderson, J.F., Pesko, K.N., Newman, R.M., Lennon, N.J., Birren, B.W., Ebel, G.D., Henn, M.R. (2011). Molecular evolution of West Nile virus in a northern temperate region: Connecticut, USA 1999-2008. Virology 417, 203-210.

Avirutnan, P., Hauhart, R.E., Somnuke, P., Blom, A.M., Diamond, M.S., Atkinson, J.P. (2011). Binding of Flavivirus Nonstructural Protein NS1 to C4b Binding Protein Modulates Complement Activation. The Journal of Immunology 187, 424.

Bakonyi, T., Hubálek, Z., Rudolf, I., Nowotny, N. (2005). Novel flavivirus or new lineage of West Nile virus, central Europe. Emerging Infect. Dis. 11, 225-231.

Bakonyi, T., Ivanics, E., Erdelyi, K., Ursu, K., Ferenczi, E., Weissenbock, H., Nowotny, N. (2006). Lineage 1 and 2 strains of encephalitic West Nile virus, central Europe. Emerg. Infect. Dis. 12, 618-623.

Balenghien, T., Fouque, F., Sabatier, P., Bicout, D. (2006). Horse-, bird-, and human-seeking behavior and seasonal abundance of mosquitoes in a West Nile virus focus of southern France. J. Med. Entomol. 43, 936-946.

Beasley, D.W.C., Davis, C.T., Guzman, H., Vanlandingham, D.L., Travassos da Rosa, A., Parsons, R.E., Higgs, S., Tesh, R.B., Barrett, A.D.T. (2003). Limited evolution of West Nile virus has occurred during its southwesterly spread in the United States. Virol 309, 190-195.

Beasley, D.W.C., Whiteman, M.C., Zhang, S., Huang, C.Y.H., Schneider, B.S., Smith, D.R., Gromowski, G.D., Higgs, S., Kinney, R.M., Barrett, A.D.T. (2005). Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. J. Virol. 79, 8339.

Bell, J.A., Brewer, C.M., Mickelson, N.J., Garman, G.W., Vaughan, J.A. (2006). West Nile virus epizootiology, central Red River Valley, North Dakota and Minnesota, 2002-2005. Emerg. Infect. Dis. 12, 1245-1247.

Bernard, K., Maffei, J., Jones, S., Kauffman, E., Ebel, G., Dupuis, A. (2001). West Nile virus infection in birds and mosquitoes, New York State, 2000. Emerging Infectious Diseases 7, 679.

Berthet, F., Zeller, H., Drouet, M., Rauzier, J., Digoutte, J., Deubel, V. (1997). Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses. J. Gen. Virol. 78, 2293.

Bertolotti, L., Kitron, U., Goldberg, T.L. (2007). Diversity and evolution of West Nile virus in Illinois and the United States, 2002-2005. Virology 360, 143-149.

Bertolotti, L., Kitron, U.D., Walker, E.D., Ruiz, M.O., Brawn, J.D., Loss, S.R., Hamer, G.L., Goldberg, T.L. (2008). Fine-scale genetic variation and evolution of West Nile Virus in a transmission "hot spot" in suburban Chicago, USA. Virology 374, 381-389.

Bin, H., Grossman, Z., Pokamunski, S., Malkinson, M., Weiss, L., Duvdevani, P., Banet, C., Weisman, Y., Annis, E., Gandaku, D. (2001). West Nile Fever in Israel 1999-2000. Ann. N. Y. Acad. Sci. 951, 127-142.

Bollati, M., Alvarez, K., Assenberg, R., Baronti, C., Canard, B., Cook, S., Coutard, B., Decroly, E., de Lamballerie, X., Gould, E.A. (2010). Structure and functionality in flavivirus NS-proteins: Perspectives for drug design. Antiviral Res. 87, 125-18.

Bondre, V.P., Jadi, R., Mishra, A., Yergolkar, P., Arankalle, V. (2007). West Nile virus isolates from India: evidence for a distinct genetic lineage. J. Gen. Virol. 88, 875.

Botha, E.M., Markotter, W., Wolfaardt, M., Paweska, J.T., Swanepoel, R., Palacios, G., Nel, L.H., Venter, M. (2008). Genetic determinants of virulence in pathogenic lineage 2 West Nile virus strains. Emerging Infectious Diseases 14, 222.

Brackney, D.E., Beane, J.E., Ebel, G.D. (2009). RNAi targeting of West Nile virus in mosquito midguts promotes virus diversification. PLoS Pathog 5, e1000502.

Brackney, D.E., Pesko, K.N., Brown, I.K., Deardorff, E.R., Kawatachi, J., Ebel, G.D. (2011). West Nile virus genetic diversity is maintained during transmission by *Culex pipiens quinquefasciatus* mosquitoes. PLoS ONE 6, e24466.

Brault, A.C., Huang, C.Y.H., Langevin, S.A., Kinney, R.M., Bowen, R.A., Ramey, W.N., Panella, N.A., Holmes, E.C., Powers, A.M., Miller, B.R. (2007). A single positively selected West Nile viral mutation confers increased virogenesis in American crows. Nat. Genet. 39, 1162-1166.

Brault, A.C., Langevin, S.A., Bowen, R.A., Panella, N.A., Biggerstaff, B.J., Miller, B.R., Komar, N. (2004). Differential virulence of West Nile strains for American crows. Emerging infectious diseases 10, 2161.

Brinton, M.A. (2009). Molecular Biology of West Nile Virus. West Nile Encephalitis Virus Infection , 97-136.

Brinton, M.A. (2001). Host factors involved in West Nile virus replication. Annals of the New York Academy of Sciences 951, 207-219.

Bryant, J.E., Holmes, E.C., Barrett, A.D.T. (2007). Out of Africa: a molecular perspective on the introduction of yellow fever virus into the Americas. PLoS Pathog 3, e75.

Calisher, C.H., Karabatsos, N., Dalrymple, J.M., Shope, R.E., Porterfield, J.S., Westaway, E.G., Brandt, W.E. (1989). Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. J. Gen. Virol. 70, 37.

CDC. (1999). Outbreak of West Nile-like viral encephalitis—New York, 1999. Morbidity and Mortality Weekly Report 48, 845-849.

Charrel, R., Brault, A., Gallian, P., Lemasson, J.J., Murgue, B., Murri, S., Pastorino, B., Zeller, H., De Chesse, R., De Micco, P. (2003). Evolutionary relationship between Old World West Nile virus strains:: Evidence for viral gene flow between africa, the middle east, and europe. Virology 315, 381-388.

Cheng, G., Cox, J., Wang, P., Krishnan, M.N., Dai, J., Qian, F., Anderson, J.F., Fikrig, E. (2010). A C-type lectin collaborates with a CD45 phosphatase homolog to facilitate West Nile virus infection of mosquitoes. Cell 142, 714-725.

Chu, J.J., Ng, M.L. (2004a). Infectious entry of West Nile virus occurs through a clathrin-mediated endocytic pathway. Journal of virology 78, 10543-10555.

Chu, J.J., Ng, M.L. (2004b). Interaction of West Nile virus with alpha v beta 3 integrin mediates virus entry into cells. Journal of Biological Chemistry 279, 54533-54541.

Ciota, A.T. and L.D. Kramer. (2010). Insights into arbovirus evolution and adaptation from experimental studies.Viruses. 12, 2594-617.

Ciota, A.T., Lovelace, A.O., Jia, Y., Davis, L.J., Young, D.S., Kramer, L.D. (2008). Characterization of mosquito-adapted West Nile virus. J. Gen. Virol. 89, 1633.

Ciota, A.T., Ngo, K.A., Lovelace, A.O., Payne, A.F., Zhou, Y., Shi, P.Y., Kramer, L.D. (2007). Role of the mutant spectrum in adaptation and replication of West Nile virus. J. Gen. Virol. 88, 865-874.

Clyde, K. and E. Harris. (2006). RNA secondary structure in the coding region of dengue virus type 2 directs translation start codon selection and is required for viral replication. Journal of virology 80, 2170-2182.

Coffey, L.L., Vasilakis, N., Brault, A.C., Powers, A.M., Tripet, F., Weaver, S.C. (2008). Arbovirus evolution in vivo is constrained by host alternation. Proceedings of the National Academy of Sciences 105, 6970.

Coffey, L.L. and M. Vignuzzi. (2011). Host alternation of chikungunya virus increases fitness while restricting population diversity and adaptability to novel selective pressures. J. Virol. 85, 1025.

Daffis, S., Lazear, H.M., Liu, W.J., Audsley, M., Engle, M., Khromykh, A.A., Diamond, M.S. (2011). The naturally attenuated Kunjin strain of West Nile virus shows enhanced sensitivity to the host type I interferon response. J. Virol. 85, 5664.

Daffis, S., Suthar, M.S., Gale Jr, M., Diamond, M.S. (2009). Measure and Countermeasure: Type I IFN (IFN-α/β) Antiviral Response against West Nile Virus. Journal of Innate Immunity 1, 435.

Davis, C.T., Beasley, D.W.C., Guzman, H., Siirin, M., Parsons, R.E., Tesh, R.B., Barrett, A.D.T. (2004). Emergence of attenuated West Nile virus variants in Texas, 2003. Virology 330, 342-350.

Davis, C.T., Li, L., May, F.J., Bueno Jr, R., Dennett, J.A., Bala, A.A., Guzman, H., Elizondo-Quiroga, D., Tesh, R.B., Barrett, A.D. (2007). Genetic stasis of dominant West Nile virus genotype, Houston, Texas. Emerging Infectious Diseases 13, 601.

Davis, C.W., Nguyen, H.Y., Hanna, S.L., Sánchez, M.D., Doms, R.W., Pierson, T.C. (2006). West Nile virus discriminates between DC-SIGN and DC-SIGNR for cellular attachment and infection. J. Virol. 80, 1290.

Davis, C.T., Ebel, G.D., Lanciotti, R.S., Brault, A.C., Guzman, H., Siirin, M., Lambert, A., Parsons, R.E., Beasley, D.W., Novak, R.J., Elizondo-Quiroga, D., Green, E.N., Young, D.S., Stark, L.M., Drebot, M.A., Artsob, H., Tesh, R.B., Kramer, L.D., Barrett, A.D. (2005). Phylogenetic analysis of North American West Nile virus isolates, 2001-2004: evidence for the emergence of a dominant genotype. Virology 342, 252-265.

Depoortere, E., Kavle, J., Keus, K., Zeller, H., Murri, S., Legros, D. (2004). Outbreak of West Nile virus causing severe neurological involvement in children, Nuba Mountains, Sudan, 2002. Tropical Medicine and International Health 9, 730-736.

Diamond, M.S., Mehlhop, E., Oliphant, T., Samuel, M.A. (2009). The host immunologic response to West Nile encephalitis virus. Front. Biosci. 14, 3024-3034.

Ding, S.W. (2010). RNA-based antiviral immunity. Nature Reviews Immunology 10, 632-644.

Duarte, E., Clarke, D., Moya, A., Domingo, E., Holland, J. (1992). Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. Proceedings of the National Academy of Sciences 89, 6015.

Dusek, R.J., McLean, R.G., Kramer, L.D., Ubico, S.R., Dupuis, A.P., Ebel, G.D., Guptill, S.C. (2009). Prevalence of West Nile virus in migratory birds during spring and fall migration. Am. J. Trop. Med. Hyg. 81, 1151.

Ebel, G.D. (2010). Update on Powassan virus: emergence of a North American tick-borne flavivirus. Annu. Rev. Entomol. 55, 95-110.

Ebel, G.D. and L.D. Kramer. (2009). West Nile Virus: Molecular Epidemiology and Diversity, in Diamond, M.S. (Ed.), West Nile Encephalitis Virus Infection. Springer, pp. 25-43.

Ebel, G., and A. Dupuis. (2001). Partial genetic characterization of West Nile virus strains, New York State, 2000. Emerging Infectious Diseases 7, 650.

Ebel, G.D., Carricaburu, J., Young, D., Bernard, K.A., Kramer, L.D. (2004). Genetic and phenotypic variation of West Nile virus in New York, 2000-2003. American Journal of Tropical Medicine and Hygiene 71, 493-500.

Eidson, M., Komar, N., Sorhage, F., Nelson, R., Talbot, T., Mostashari, F., McLean, R. (2001). Crow deaths as a sentinel surveillance system for West Nile virus in the northeastern United States, 1999. Emerging Infectious Diseases 7, 615.

Evans, J.D., Crown, R.A., Sohn, J.A., Seeger, C. (2011). West Nile Virus Infection Induces Depletion of IFNAR1 Protein Levels. Viral Immunol. 24, 253-263.

Gaunt, M.W., Sall, A.A., Lamballerie, X.d., Falconar, A.K.I., Dzhivanian, T.I., Gould, E.A. (2001). Phylogenetic relationships of flaviviruses correlate with their epidemiology, disease association and biogeography. Journal of General Virology 82, 1867-1876.

Goldberg, T.L., Anderson, T.K., Hamer, G.L. (2010). West Nile virus may have hitched a ride across the Western United States on *Culex tarsalis* mosquitoes. Mol. Ecol. 19, 1518-1519.

Gray, R., Veras, N., Santos, L., Salemi, M. (2010). Evolutionary characterization of the West Nile Virus complete genome. Mol. Phylogenet. Evol. 56, 195-200.

Grinev, A., Daniel, S., Stramer, S., Rossmann, S., Caglioti, S., Rios, M. (2008). Genetic variability of West Nile virus in US blood donors, 2002–2005. Emerg Infect Dis 14, 436.

Hamer, G.L., Kitron, U.D., Brawn, J.D., Loss, S.R., Ruiz, M.O., Goldberg, T.L., et al. (2008). *Culex Pipiens* (Diptera: Culicidae): A Bridge Vector of West Nile Virus to Humans. J.Med.Entomol. 45, 125-128.

Hanna, S.L., Pierson, T.C., Sanchez, M.D., Ahmed, A.A., Murtadha, M.M., Doms, R.W. (2005). N-linked glycosylation of West Nile virus envelope proteins influences particle assembly and infectivity. J. Virol. 79, 13262.

Hardy, J.L., Houk, E.J., Kramer, L.D., Reeves, W.C. (1983). Intrinsic factors affecting vector competence of mosquitoes for arboviruses. Annual Review of Entomology 28, 229-262.

Hayes, C.G. (2001). West Nile virus: Uganda, 1937, to New York City, 1999. Ann. N. Y. Acad. Sci. 951, 25-37.

Herring, B.L., Bernardin, F., Caglioti, S., Stramer, S., Tobler, L., Andrews, W., Cheng, L., Rampersad, S., Cameron, C., Saldanha, J. (2007). Phylogenetic analysis of WNV in North American blood donors during the 2003-2004 epidemic seasons. Virology 363, 220-228.

Holmes, E.C. and S.S. Twiddy. (2003). The origin, emergence and evolutionary genetics of dengue virus. Infection, Genetics and Evolution 3, 19-28.

Hubalek, Z. and J. Halouzka.(1999). West Nile fever--a reemerging mosquito-borne viral disease in Europe. Emerging Infectious Diseases 5, 643.

Hubalek, Z., Rudolf, I., Bakonyi, T., Kazdova, K., Halouzka, J., Sebesta, O., Sikutova, S., Juricova, Z., Nowotny, N. (2010). Mosquito (Diptera: Culicidae) surveillance for arboviruses in an area endemic for West Nile (Lineage Rabensburg) and Tahyna viruses in Central Europe. J. Med. Entomol. 47, 466-472.

Jenkins, G.M., Rambaut, A., Pybus, O.G., Holmes, E.C. (2002). Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. J. Mol. Evol. 54, 156-165.

Jerzak, G.V., Bernard, K., Kramer, L.D., Shi, P.Y., Ebel, G.D. (2007). The West Nile virus mutant spectrum is host-dependant and a determinant of mortality in mice. Virology 360, 469-476.

Jerzak, G., Bernard, K.A., Kramer, L.D., Ebel, G.D. (2005). Genetic variation in West Nile virus from naturally infected mosquitoes and birds suggests quasispecies structure and strong purifying selection. J Gen Virol 86, 2175-2183.

Jia, X.Y., Briese, T., Jordan, I., Rambaut, A., Chang Chi, H., Mackenzie, J.S., Hall, R.A., Scherret, J., Lipkin, W.I. (1999). Genetic analysis of West Nile New York 1999 encephalitis virus. The Lancet 354, 1971-1972.

Khromykh, A.A., Meka, H., Guyatt, K.J., Westaway, E.G. (2001). Essential role of cyclization sequences in flavivirus RNA replication. J Virol 75, 6719-6728.

Kilpatrick, A.M., Meola, M.A., Moudy, R.M., Kramer, L.D. (2008). Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. PLoS pathogens 4, e1000092.

Kilpatrick, A.M., Kramer, L.D., Campbell, S.R., Alleyne, E.O., Dobson, A.P., Daszak, P. (2005). West Nile virus risk assessment and the bridge vector paradigm. Emerging infectious diseases 11, 425-429.

Kilpatrick, A.M., Kramer, L.D., Jones, M.J., Marra, P.P., Daszak, P. (2006). West Nile virus epidemics in North America are driven by shifts in mosquito feeding behavior. PLoS Biology 4, e82.

King, A.M.Q., Lefkowitz, E., Adams, M.J., Carstens, E.B. (2011). Virus Taxonomy: 9th Report of the International Committee on Taxonomy of Viruses, 1st edition, Elsevier, San Diego, CA.

Komar, N., Langevin, S., Hinten, S., Nemeth, N., Edwards, E., Hettler, D., Davis, B., Bowen, R., Bunning, M. (2003). Experimental infection of North American birds with the New York 1999 strain of West Nile virus. Emerging Infectious Diseases 9, 311.

Kramer, L.D., Styer, L.M., Ebel, G.D. (2008). A global perspective on the epidemiology of West Nile virus. Annu. Rev. Entomol. 63, 81.

Kramer, L.D. and K.A. Bernard. (2001). West Nile virus infection in birds and mammals. Ann NY Acad Sci 951, 84-93.

Kuno, G., Chang, G.J.J., Tsuchiya , K.R., Karabatsos, N., Cropp, C.B. (1998). Phylogeny of the genus Flavivirus. J Virol 72, 73-83.

Kutasi, O., Bakonyi, T., Lecollinet, S., Biksi, I., Ferenczi, E., Bahuon, C., Sardi, S., Zientara, S., Szenci, O. (2011). Equine Encephalomyelitis Outbreak Caused by a Genetic Lineage 2 West Nile Virus in Hungary. Journal of Veterinary Internal Medicine .

Lanciotti, R., Roehrig, J., Deubel, V., Smith, J., Parker, M., Steele, K., Crise, B., Volpe, K., Crabtree, M., Scherret, J. (1999). Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. Science 286, 2333.

Langevin, S., Bowen, R., Ramey, W., Sanders, T., Maharaj, P., Fang, Y., Cornelius, J., Barker, C., Reisen, W., Beasley, D. (2011). Envelope and pre-membrane structural amino acid mutations mediate diminished avian growth and virulence of a Mexican West Nile virus isolate. J. Gen. Virol. .

Laurent-Rolle, M., Boer, E.F., Lubick, K.J., Wolfinbarger, J.B., Carmody, A.B., Rockx, B., Liu, W., Ashour, J., Shupert, W.L., Holbrook, M.R., Barrett, A.D., Mason, P.W., Bloom, M.E., Garcia-Sastre, A., Khromykh, A.A., Best, S.M. (2010). 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, 3503-3515.

Li, J., Bhuvanakantham, R., Howe, J., Ng, M.L. (2006). The glycosylation site in the envelope protein of West Nile virus (Sarafend) plays an important role in replication and maturation processes. Journal of general virology 87, 613-622.

Liu, W. J., X. J. Wang, D. C. Clark, M. Lobigs, R. A. Hall, Khromykh, A.A. (2006). 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, 2396-2404.

Liu, W. J., X. J. Wang, V. V. Mokhonov, P.-Y. Shi, R. Randall, Khromykh, A.A. (2005). 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, 1934-1942.

Lucas, M., Mashimo, T., Frenkiel, M.P., Simon-Chazottes, D., Montagutelli, X., Ceccaldi, P.E., Guénet, J.L., Desprès, P. (2003). Infection of mouse neurones by West Nile virus is modulated by the interferon-inducible 2′-5′ oligoadenylate synthetase 1b protein. Immunol. Cell Biol. 81, 230-236.

May, F.J., Davis, C.T., Tesh, R.B., Barrett, A.D.T. (2011). Phylogeography of West Nile virus: from the cradle of evolution in Africa to Eurasia, Australia, and the Americas. J. Virol. 85, 2964.

May, F.J., Li, L., Davis, C.T., Galbraith, S.E., Barrett, A.D.T. (2010). Multiple pathways to the attenuation of West Nile virus in south-east Texas in 2003. Virology 405, 8-14.

McLean, R., Ubico, S., Docherty, D., Hansen, W., Sileo, L., McNamara, T. (2001). West Nile virus transmission and ecology in birds. Ann. N. Y. Acad. Sci. 951, 54.

McMullen, A.R., May, F.J., Li, L., Guzman, H., Bueno Jr, R., Dennett, J.A., Tesh, R.B., Barrett, A.D.T. (2011). Evolution of New Genotype of West Nile Virus in North America. Strain 17.

Mertens, E., Kajaste-Rudnitski, A., Torres, S., Funk, A., Frenkiel, M.P., Iteman, I., Khromykh, A.A., Despres, P. (2010). Viral determinants in the NS3 helicase and 2K peptide that promote West Nile virus resistance to antiviral action of 2',5'-oligoadenylate synthetase 1b. Virology 399, 176-185.

Molaei, G., Cummings, R.F., Su, T., Armstrong, P.M., Williams, G.A., Cheng, M.L., Webb, J.P., Andreadis, T.G. (2010). Vector-host interactions governing epidemiology of West Nile virus in Southern California. Am. J. Trop. Med. Hyg. 83, 1269.

Molaei, G. and T.G. Andreadis. (2006). Identification of Avian-and Mammalian-Derived Bloodmeals in *Aedes Vexans* and *Culiseta Melanura* (Diptera: Culicidae) and its Implication for West Nile Virus Transmission in Connecticut, USA. J.Med.Entomol. 43, 1088-1093.

Monaco, F., Savini, G., Calistri, P., Polci, A., Pinoni, C., Bruno, R., Lelli, R. (2011). 2009 West Nile disease epidemic in Italy: First evidence of overwintering in Western Europe? Res. Vet. Sci. 91, 321-326.

Moudy, R.M., Meola, M.A., Morin, L.L.L., Ebel, G.D., Kramer, L.D. (2007). A newly emergent genotype of West Nile virus is transmitted earlier and more efficiently by Culex mosquitoes. Am. J. Trop. Med. Hyg. 77, 365.

Moudy, R.M., Zhang, B., Shi, P.Y., Kramer, L.D. (2009). West Nile virus envelope protein glycosylation is required for efficient viral transmission by Culex vectors. Virology 387, 222-228.

Mukhopadhyay, S., Kim, B.S., Chipman, P.R., Rossmann, M.G., Kuhn, R.J. (2003). Structure of West Nile virus. Science 302, 248.

Muñoz-Jordán, J. L., Sanchez-Burgos, G. G., Laurent-Rolle, M., García-Sastre, A. (2003). Inhibition of interferon signaling by dengue virus. Proc. Natl. Acad. Sci. USA. 100, 14333-14338.

Murata, R., Eshita, Y., Maeda, A., Maeda, J., Akita, S., Tanaka, T., Yoshii, K., Kariwa, H., Umemura, T., Takashima, I. (2010). 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.

Murray, K.O., Mertens, E., Desprès, P. (2010a). West Nile virus and its emergence in the United States of America. Vet. Res. 41.

Murray, K., Walker, C., Herrington, E., Lewis, J.A., McCormick, J., Beasley, D.W., Tesh, R.B., Fisher-Hoch, S. (2010b). Persistent infection with West Nile virus years after initial infection. J. Infect. Dis. 201, 2-4.

Papa, A., Danis, K., Baka, A., Bakas, A., Dougas, G., Lytras, T., Theocharopoulos, G., Chrysagis, D., Vassiliadou, E., Kamaria, F. (2010). Ongoing outbreak of West Nile virus infections in humans in Greece, July–August 2010. Age 20, 20-29.

Papa, A., Xanthopoulou, K., Gewehr, S., Mourelatos, S. (2011). Detection of West Nile virus lineage 2 in mosquitoes during a human outbreak in Greece. Clinical Microbiology and Infection .

Platonov, A.E. (2001). West Nile Encephalitis in Russia 1999-2001. Ann. N. Y. Acad. Sci. 951, 102-116.

Rappole, J.H., Compton, B.W., Leimgruber, P., Robertson, J., King, D.I., Renner, S.C. (2006). Modeling movement of West Nile virus in the Western hemisphere. Vector-borne and zoonotic diseases 6, 128-139.

Ronquist, F. and J. P. Huelsenbeck. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Samuel, M.A. and M.S. Diamond. (2006). Pathogenesis of West Nile virus infection: a balance between virulence, innate and adaptive immunity, and viral evasion. J. Virol. 80, 9349.

Schubert, A.M. and C. Putonti. (2010). Evolution of the sequence composition of Flaviviruses. Infection, Genetics and Evolution 10, 129-136.

Shirato, K., Miyoshi, H., Goto, A., Ako, Y., Ueki, T., Kariwa, H., Takashima, I. (2004). 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.

Smith, D.R., Adams, A.P., Kenney, J.L., Wang, E., Weaver, S.C. (2008). Venezuelan equine encephalitis virus in the mosquito vector *Aedes taeniorhynchus*: Infection initiated by a small number of susceptible epithelial cells and a population bottleneck. Virology 372, 176-186.

Smithburn, K., Hughes, T., Burke, A., Paul, J. (1940). A neurotropic virus isolated from the blood of a native of Uganda. American Journal of Tropical Medicine and Hygiene 20, 471-473.

Snapinn, K.W., Holmes, E.C., Young, D.S., Bernard, K.A., Kramer, L.D., Ebel, G.D. (2007). Declining growth rate of West Nile virus in North America. J. Virol. 81, 2531.

Tsai, T., Popovici, F., Cernescu, C., Campbell, G., Nedelcu, N. (1998). West Nile encephalitis epidemic in southeastern Romania. The Lancet 352, 767-771.

Tsai, T.F., Mitchell, C.J., Monath, T.P. (1989). St. Louis encephalitis, in Anonymous (Ed.), The Arboviruses: Epidemiology and Ecology. CRC, Boca Raton, FL, pp. 113-143.

Tuplin, A., Evans, D., Buckley, A., Jones, I., Gould, E., Gritsun, T. (2011). Replication enhancer elements within the open reading frame of tick-borne encephalitis virus and their evolution within the Flavivirus genus. Nucleic Acids Res. 39, 7034-7048.

Turell, M.J., Dohm, D.J., Sardelis, M.R., Oguinn, M.L., Andreadis, T.G., Blow, J.A. (2005). An update on the potential of North American mosquitoes (Diptera: Culicidae) to transmit West Nile virus. J Med Entomol 42, 57-62.

Turell, M.J., Sardelis, M.R., O'Guinn, M.L., Dohm, D.J. (2002). Potential vectors of West Nile virus in North America. Current topics in microbiology and immunology 267, 241-252.

Vazquez, A., Sanchez-Seco, M.P., Ruiz, S., Molero, F., Hernandez, L., Moreno, J., Magallanes, A., Tejedor, C.G., Tenorio, A. (2010). Putative new lineage of west nile virus, Spain. Emerg. Infect. Dis. 16, 549-552.

Venkatesan, M.and J.L. Rasgon. (2010). Population genetic data suggest a role for mosquito-mediated dispersal of West Nile virus across the western United States. Mol. Ecol. 19, 1573-1584.

Venter, M., Human, S., Zaayman, D., Gerdes, G.H., Williams, J., Steyl, J., Leman, P.A., Paweska, J.T., Setzkorn, H., Rous, G. (2009). Lineage 2 West Nile virus as cause of fatal neurologic disease in horses, South Africa. Emerging infectious diseases 15, 877.

Venter, M., and R. Swanepoel. (2010). West Nile virus lineage 2 as a cause of zoonotic neurological disease in humans and horses in southern Africa. Vector-Borne and Zoonotic Diseases 10, 659-664.

Vitek, C.J., Richards, S.L., Mores, C.N., Day, J.F., Lord, C.C. (2008). Arbovirus transmission by *Culex nigripalpus* in Florida, 2005. J. Med. Entomol. 45, 483.

Weaver, S.C. (2006). Evolutionary influences in arboviral disease. Current Topics in Microbiology and Immunology 299, 285-314.

Westaway, E.G., Mackenzie, J.M., Khromykh, A.A. (2002). Replication and gene function in Kunjin virus. Current topics in microbiology and immunology 267, 323-351.

Wicker, J.A., Whiteman, M.C., Beasley, D.W.C., Davis, C.T., Zhang, S., Schneider, B.S., Higgs, S., Kinney, R.M., Barrett, A.D.T. (2006). 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, 245-253.

Wilson, J.R., de Sessions, P.F., Leon, M.A., Scholle, F. (2008). West Nile virus nonstructural protein 1 inhibits TLR3 signal transduction. J. Virol. 82, 8262-8271.

Yoon, S.W., Lee, S.Y., Won, S.Y., Park, S.H., Park, S.Y., Jeong, Y.S. (2006). Characterization of homologous defective interfering RNA during persistent infection of Vero cells with Japanese encephalitis virus. Molecules and Cells 21, 112-120.

Zehender, G., Ebranati, E., Bernini, F., Presti, A.L., Rezza, G., Delogu, M., Galli, M., Ciccozzi, M. (2011). Phylogeography and epidemiological history of West Nile virus genotype 1a in Europe and the Mediterranean basin. Infection, Genetics and Evolution .

Zeller, H. and I. Schuffenecker. (2004). West Nile virus: an overview of its spread in Europe and the Mediterranean basin in contrast to its spread in the Americas. European Journal of Clinical Microbiology and Infectious Diseases 23, 147-156.

Zhang, B., Dong, H., Stein, D.A., Iversen, P.L., Shi, P.Y. (2008). West Nile virus genome cyclization and RNA replication require two pairs of long-distance RNA interactions. Virology 373, 1-13.

Zou, G., Puig-Basagoiti, F., Zhang, B., Qing, M., Chen, L., Pankiewicz, K.W., Felczak, K., Yuan, Z., Shi, P.Y. (2009a). A single-amino acid substitution in West Nile virus 2K peptide between NS4A and NS4B confers resistance to lycorine, a flavivirus inhibitor. Virology 384, 242-252.

Zou, G., Zhang, B., Lim, P.Y., Yuan, Z., Bernard, K.A., Shi, P.Y. (2009b). Exclusion of West Nile virus superinfection through RNA replication. J. Virol. 83, 11765-11776.