Intra and interhost dynamics shaping arbovirus adaptation and evolution

Alexander T. Ciota

University at Albany, State University of New York, aciota@wadsworth.org

The University at Albany community has made this article openly available. Please share how this access benefits you.

Follow this and additional works at: https://scholarsarchive.library.albany.edu/legacy-etd

Part of the Biology Commons, Evolution Commons, and the Virology Commons

Recommended Citation
https://scholarsarchive.library.albany.edu/legacy-etd/524

This Dissertation is brought to you for free and open access by the The Graduate School at Scholars Archive. It has been accepted for inclusion in Legacy Theses & Dissertations (2009 - 2024) by an authorized administrator of Scholars Archive.
Please see Terms of Use. For more information, please contact scholarsarchive@albany.edu.
INTRA AND INTERHOST DYNAMICS SHAPING ARBOVIRUS ADAPTATION AND EVOLUTION

by

Alexander T. Ciota

A Dissertation
Submitted to University of Albany, State University of New York
In Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

College of Arts and Sciences
Department of Biological Sciences
2012
Acknowledgments

I would like to thank my research advisor, mentor, and supervisor Laura Kramer for all her support, guidance, and confidence in me throughout the years. I would also like to thank my on campus advisor, IngNang Wang, as well as other members of my committee including Kurt McKean, Tom Caraco, and Paul Turner, who have provided valuable guidance and perspective for these studies. I also thank the entire SUNY Albany Biology Department for supporting my participation in this program. I am sincerely grateful for the abundant technical assistance I have been fortunate enough to have for all my studies. Dylan Ehrbar, Graham Willsey, Anne Payne, Amy Mattachiero, Greta Van Slyke, and Pamela Chin have all played particularly critical technical roles in these studies and all past and current members of the Arbovirus laboratory have also contributed at various times. I thank Claus Wilke for guidance and review of portions of this work as well as Evan Koch for phylogenetic analyses. Thanks also to my friends and family for all their support throughout this process. This work was supported by funds from the National Institutes of Health (grant number RO1-AI-077669). The construction of the Wadsworth Center Insectary Facility was funded by NIH grant number C06-RR-17715.
Table of Contents

Acknowledgements.................................................................................................ii
Table of Contents......................................................................................................iii
Abstract...................................................................................................................v
Chapter 1: INTRODUCTION AND SYNOPSIS .........................................................1
Chapter 2: TEMPORAL AND SPATIAL ALTERATION IN MUTANT SWARM SIZE OF ST. LOUIS ENCEPHALITIS VIRUS IN MOSQUITO HOSTS ..................................................................................................................9
Chapter 3: QUANTIFICATION OF INTRAHOST BOTTLENECKS OF WEST NILE VIRUS IN CULEX PIPIENS MOSQUITOES USING AN ARTIFICIAL MUTANT SWARM .................................................................38
Chapter 4: COOPERATIVE INTERACTIONS IN THE WEST NILE VIRUS MUTANT SWARM ................................................................................................................70
Chapter 5: THE EVOLUTION OF VIRULENCE OF WEST NILE VIRUS IN A MOSQUITO VECTOR: IMPLICATIONS FOR ARBOVIRUS ADAPTATION AND EVOLUTION .................................................................99
ABSTRACT:

Arthropod-borne viruses (arboviruses), which are predominately mosquito-borne and almost exclusively RNA viruses, are maintained in nature in complex transmission cycles involving blood sucking invertebrates and vertebrate hosts. Although over 120 arboviruses are human pathogens responsible for causing a significant and expanding global health burden, a detailed understanding of the complex interactions between these pathogens and their hosts, particularly invertebrate hosts, is lacking. Defining these interactions is necessary if we are to understand the selective pressures and, therefore, evolutionary, adaptive, and epidemiological potential of arboviruses. This requires experimental infection and evolution studies, particularly in vivo, with natural hosts. The results presented here, with West Nile virus (WNV) and St. Louis encephalitis virus (SLEV; Flaviviridae: Flavivirus), are a compilation of such studies. WNV is the most prevalent arboviral pathogen in the U.S. and the most geographically distributed arbovirus in the world. SLEV is a close relative of WNV, with genetic, antigenic, and ecological similarities, yet both geographic range and levels of activity of SLEV are much more limited than WNV. Despite the evolutionary potential of these RNA viruses, arboviruses including WNV and SLEV have experienced only limited consensus level change. This relative evolutionary stasis has been attributed primarily to the requirement for host cycling, yet previous results demonstrate differential selective pressures may be overstated and results presented here demonstrate that evolutionary constraint may also result from interseasonal and intrahost bottlenecks, cooperative interactions in the viral swarm, and virulence in mosquito vectors. In addition, these studies expand on previous work demonstrating that a complete characterization of the mutant swarm, both its breadth and composition, is required if we are to fully evaluate genetic change and its phenotypic consequences. Lastly, these results demonstrate the specificity and complexity of vector-virus interactions and the need to evaluate individual systems independently. Taken together, these studies expand our knowledge of the complex forces shaping arbovirus evolution and adaptation and establish a baseline for future mechanistic studies.
CHAPTER 1: INTRODUCTION AND SYNOPSIS

Arthropod-borne viruses (arboviruses) are maintained in nature in complex transmission cycles between vertebrate hosts and hematophagous invertebrate vectors. By definition, these viruses require infection, replication, and ultimately transmission from competent arthropod vectors including mosquitoes, ticks, and various biting flies to susceptible hosts that allow amplification of the virus to sufficiently high levels to infect new vectors. There are more than 14,000 species of arthropods identified as having the capacity to harbor and transmit arboviruses. Additionally, there are greater than 120 arboviruses that are associated with human disease and many that are major public health problems with global distributions (Ciota et al., 2010; Crosskey, 1988; Kuno & Chang, 2005). West Nile virus (WNV; Flaviviridae: Flavivirus) is the most geographically widespread arbovirus in the world. Although the majority of WNV infections are thought to be asymptomatic, approximately 20% of infections result in West Nile fever, and ~1% result in central nervous system infection. CNS infections with WNV may result a range of clinical outcomes including encephalitis, meningitis, acute flaccid paralysis and death (Petersen & Marfin, 2002).

WNV and its close relative, St. Louis encephalitis virus (SLEV; Flaviviridae: Flavivirus), are maintained in nature in enzootic transmission cycles between predominately ornithophilic Culex sp. mosquitoes and birds. Since mammals represent dead-end hosts for these viruses, human infection occurs only as a result of spillover from this cycle and therefore does not contribute to the evolutionary pressures shaping these pathogens (Bernard & Kramer, 2001). The flavivirus genome is single-stranded, positive
polarity RNA and is approximately 11kb in length. The genomes consist of a 5’ untranslated region (UTR), a single, long open-reading frame (ORF), and a 3’ UTR. The ORF encodes a polyprotein that is co- and post-translationally cleaved by viral and cellular proteases into three structural proteins (C, premembrane [prM] or M, and E) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5). Structural proteins function primarily in viral particle formation, while nonstructural proteins function in viral replication, virion assembly, and evasion of host innate immune response. WNV infection occurs through receptor-mediated endocytosis. Flaviviral RNA replication is associated with rough endoplasmic reticulum (ER), and is coupled to virion assembly (Khromykh et al., 2001; Monath & Heinz, 1996).

WNV was first introduced into the U.S. in the New York City area in 1999. Following its introduction, WNV steadily increased both its host and geographic ranges throughout the Americas. WNV is now the most widespread and prevalent arbovirus in the U.S. and human cases in 2012 have now exceeded levels recorded in any previous year other than 2003. Outbreaks of the SLEV, in contrast, occur only periodically in the U.S., with levels of activity and host range which are much more restricted than those of WNV (www.cdc.gov/ncidod/dvbid).

The fact that these arboviruses are almost exclusively RNA viruses may be explained by a requirement for plasticity in the face of ever changing host environments (Weaver, 2006). RNA-dependent RNA-polymerase (RdRp) error rates are estimated to range from $10^{-3}$ to $10^{-5}$ errors / nucleotide / round of replication (Domingo & Holland, 1994; Drake & Holland, 1999), which is equivalent to approximately 1 mutation in each new WNV virion produced. This, together with rapid and high levels of viral replication,
allows quick exploration of fitness landscapes. Despite the enormous potential for sequence change inherent in RNA viruses, the consensus sequences of most arboviruses, including WNV and SLEV, have remained highly genetically conserved in nature (Davis et al., 2005; Ebel et al., 2004; Jenkins et al., 2002; Twiddy et al., 2003; Weaver et al., 1992). This evolutionary stasis is generally attributed to the differential selective pressures applied by disparate vertebrate and invertebrate hosts (Scott et al., 1994; Woolhouse et al., 2001). This implies that only mutations which are either beneficial or neutral in both hosts become fixed, resulting in a situation in which sequence changes are much more likely to be purged by purifying selection than in single host systems (Domingo & Holland, 1997; Levins, 1968; Wright, 1931). Indeed, phylogenetic studies of arboviruses analyzing the proportion of nonsynonomous change over time demonstrate that purifying selection is generally the dominant selective force in arbovirus evolution (Woelk & Holmes, 2002). An extension of the concept of genetic constraints is limitation on host-specific adaptation, i.e., fitness trade-offs. The generally accepted theory is that cycling between disparate hosts selects for generalists and, as a consequence, arboviruses sacrifice the ability to be host specialists (Kassen, 2002; Wilson & Yoshimura, 1994). Specifically, arboviruses are hypothesized to lack host specialization as it would result in either positive selection for changes which are advantageous to one host but would be detrimental in the alternate host (antagonistic pleiotropy), or the accumulation of neutral mutations in one host which would be detrimental in the alternate host (mutational accumulation) (Elena & Lenski, 2003). Despite these relatively well accepted concepts, evidence for adaptive trade-offs from experimental studies is inconsistent, and particularly weak for WNV and SLEV, implying
co-adaptation is unattainable without significant host-specific sacrifice (Ciota and Kramer, 2010). In addition, there has been little consideration of the breadth and composition of the mutant swarm when assessing evolutionary constraint. It has become clear that minority variants are not just transient genomes awaiting selection but, rather, often stable members of intrahost populations in equilibrium that contribute significantly to phenotypes (Ciota et al., 2007c; Fitzpatrick et al., 2010; Novella & Ebendick-Corp, 2004). Accurate assessment of genetic change of arboviruses therefore requires assessment of the entire mutant swarm.

The goal of these studies was to evaluate the evolutionary and adaptive potential of these viruses and to better characterize the forces that determine this potential. Specifically, using a combination of in vitro and in vivo experimental evolution and infection studies, together with genetic analyses, evaluation of the genetic correlates and adaptive trade-offs associated with host-specific adaptation, the extent of both intrahost and interseasonal bottlenecks, the potential for strain interactions and complementation, and the relationship between virus adaptation and vector fitness were evaluated.

Chapter 2 has been previously published in Infection, Genetics, and Evolution (Ciota et al., 2011a). In these studies, the breadth and composition of the SLEV mutant swarm were evaluated in single passage mosquito isolates obtained from Texas and California. Previous studies with WNV demonstrate significant intrahost diversity, particularly in mosquitoes, yet no evaluation of SLEV isolates had previously been completed. These data demonstrate important temporal, spatial, and species-specific differences which have implications for SLEV adaptability. Although these data confirm the potential for intrahost variation of SLEV, they also suggest that significant seasonal
bottlenecks exist and that recent isolates may, unlike WNV, be highly homogeneous. These results expand on previous experimental studies identifying differences in the levels of intrahost diversity, adaptability, and host-specificity of WNV and SLEV (Ciota et al., 2007a; Ciota et al., 2007b; Ciota et al., 2009) and demonstrate that even closely related viruses may differ in terms of their evolutionary potential and selective pressures. These differences may be important to understanding the considerably different epidemiological patterns of these two viruses. More complete evaluation of genetic changes, both on the consensus level and within the mutant swarm, are currently underway with WNV and will further characterize the importance of seasonal bottlenecks, temporal genetic variation and its relationship to epidemiological patterns, and differences in selective pressures between WNV and SLEV.

Chapters 3-5 focus exclusively on WNV and present data utilizing both in vitro and in vivo experimental infections to better characterize the selective pressures and evolutionary constraints imposed on WNV during mosquito vector infection. Chapter 3, the bulk of which was also published in Infection, Genetics, and Evolution (Ciota et al., 2012a), utilizes an artificial mutant swarm to characterize the intrahost bottlenecks that WNV is subject to as a result of infection, dissemination, and transmission in Cx. pipiens mosquitoes. Despite high levels of variability, results demonstrate that intrahost bottlenecks, particularly midgut and salivary gland infection, as well as transmission, can significantly decrease the breadth of the WNV swarm. In addition, results demonstrate a stochastic sweep which serves to narrow the swarm with increased time of infection in secondary tissues. Results from similar studies in Cx. quinquefasciatus, which are unpublished, are also shown in Chapter 3. These data confirm previous work with WNV
in this species (Brackney et al., 2011) demonstrating that the extent of intrahost bottlenecks are species-specific, with less constraint in *Cx. quinquefasciatus*, a result which has significant implications for geographical differences in the evolutionary potential of WNV.

Chapter 4, which was previously published in *BMC Evolutionary Biology* (Ciota et al., 2012b) is an expansion of previous characterization of the highly diverse, mosquito-cell adapted strain, WNV CP40 (Ciota et al., 2007a; 2008a). These studies clearly demonstrate that complementation among co-infecting strains can significantly enhance WNV swarm fitness, and that cooperative interactions can slow selection for high fitness variants and serve to maintain genetic and phenotypic diversity within the WNV swarm. This work suggests that cooperation can evolve under high levels of co-infection. The ability to maintain diversity could be advantageous for an arbovirus given the need for infection and replication in diverse tissues and hosts, and these results provide another example of a mechanism which could act to slow arbovirus evolution. The importance of cooperative interactions *in vivo* is unknown, and studies are currently underway investigating the extent of co-infection and potential for strain complementation in mosquito vectors.

Chapter 5 is currently unpublished. This work represents an integration of two previous studies demonstrating (i) that infection with wildtype WNV does not alter life-history traits of *Cx. pipiens* mosquitoes (Ciota et al., 2011b) and (ii) that wildtype WNV has the capacity to evolve increased intrahost fitness in *Cx. pipiens* mosquitoes in the lab (Ciota et al., 2008). These observations led to a question not previously evaluated in an arbovirus-vector system, i.e. is viral intrahost fitness/transmission coupled with virulence
in a mosquito vector and does this constrain the capacity for evolution of high fitness strains in nature? These data demonstrate that, unlike wildtype WNV, mosquito-adapted WNV is indeed moderately virulent in *Cx. pipiens*, and that the changes to life-history traits resulting from exposure to this strain decrease the vectorial capacity (transmissibility) of mosquitoes such that, despite increased intrahost fitness, this strain would in fact be out-competed by wildtype WNV on the population level.

Taken together, these studies demonstrate that there are many forces beyond adaptive trade-offs that contribute to the evolutionary constraint of arboviruses. Intrahost and interseasonal bottlenecks, cooperative interactions among strains, and virulence in the vector may all contribute to dampening selection for high fitness strains. These results reveal the complexity of the selective forces acting on arboviral mutant swarms, particularly in mosquito vectors. In addition, this work adds to a growing body of data demonstrating the specificity of vector-virus interactions and the need to more fully define the mechanisms that determine outcomes of individual infections and, ultimately, the evolutionary trajectory of arboviruses.
References


CHAPTER 2: TEMPORAL AND SPATIAL ALTERATION IN MUTANT SWARM SIZE OF ST. LOUIS ENCEPHALITIS VIRUS IN MOSQUITO HOSTS*


Abstract

*St. Louis encephalitis virus* (SLEV; *Flaviviridae; Flavivirus*) is a member of the Japanese encephalitis serocomplex and a close relative of *West Nile virus* (WNV). Although SLEV remains endemic to the U.S., both levels of activity and geographical dispersal are relatively constrained when compared to the widespread distribution of WNV. In recent years, WNV appears to have displaced SLEV in California, yet both viruses currently coexist in Texas and several other states. It has become clear that viral swarm characterization is required if we are to fully evaluate the relationship between viral genomes, viral evolution, and epidemiology. Mutant swarm size and composition may be particularly important for arboviruses, which require replication not only in diverse tissues but also divergent hosts. In order to evaluate temporal, spatial, and host-specific patterns in the SLEV mutant swarm, we determined the size, composition, and phylogeny of the intrahost swarm within primary mosquito isolates from both Texas and California. Results indicate a general trend of decreasing intrahost diversity over time in both locations, with recent isolates being highly genetically homogeneous. Additionally, phylogenic analyses provide detailed information on the relatedness of minority variants both within and among strains and demonstrate how both geographic isolation and seasonal maintenance have shaped the viral swarm. Overall, these data generally provide insight into how time, space, and unique transmission cycles influence the SLEV mutant
swarm and how understanding these processes can ultimately lead to a better understanding of arbovirus evolution and epidemiology.

**Introduction**

*St. Louis encephalitis virus* (SLEV; *Flaviviridae*; *Flavivirus*) is a member of the Japanese encephalitis serocomplex and a close relative of *West Nile virus* (WNV). The initial isolation of SLEV occurred in St. Louis, MS in 1933, when over 1000 cases of encephalitis were recorded (Lumsden, 1958; Reisen, 2003). SLEV is ecologically similar to WNV in that it is maintained in transmission cycles between generally ornithophilic *Culex* mosquitoes and birds. Humans represent dead-end hosts due to generally insufficient viremia levels for successful transmission (Bernard & Kramer, 2001). The main vectors of SLEV in the western U.S. are *Culex quinquefasciatus*, which is generally restricted to southern urban areas, and *Culex tarsalis*, which is present in agricultural and rural areas (Reisen, 2003; Darsie et al., 1981). Since the last major outbreak in 1975-77, during which time more than 2500 cases were confirmed, approximately 55 SLEV cases on average have been reported per year in a total of 35 states. Both levels of activity and geographical dispersal of SLEV are relatively constrained when compared to the widespread distribution of WNV (www.cdc.gov/ncidod/dvbid/arbor). In addition, results from phylogenetic studies indicate SLEV, more so than WNV, is subject to geographic isolation (Kramer & Chandler, 2001; Baillie et al., 2008; May et al., 2008). These differing genetic and epidemiological patterns could be attributed to generally lower SLEV viremia levels in birds, resulting in a narrower competent host range and therefore, more specific transmission cycles (McLean et al., 1983; Komar et al., 2003; Marra et al.,
WNV is unique in that it can maintain high levels of viremia in a range of vertebrate hosts despite its need to cycle in mosquito vectors (Brault et al., 2007). In addition, over 75 mosquito species have been demonstrated to become infected with WNV, although *Culex* mosquitoes are implicated as the predominant vectors of both viruses (Higgs et al., 2004). In recent years, WNV has appeared to displace SLEV in California (Reisen et al., 2008), yet both viruses currently coexist in Texas (Lillibridge et al., 2004; Bradford et al., 2005; Rios et al., 2006). Previous studies demonstrate that WNV strains exist as highly diverse mutant swarms in nature and that the source of this genetic diversity is likely the mosquito (Jerzak et al., 2005). The maintenance of diversity is one mechanism by which arboviruses could overcome significant fitness costs as a result of their requirement for host cycling. *In vitro* passage studies in mosquito cells suggest that SLEV may have a relatively decreased capacity for producing and maintaining intrahost diversity, and that this reduced diversity may have direct consequences for host range (Ciota et al., 2007). Besides being relevant for adaptability, minority variants in many viruses can play significant roles in viral fitness and pathogenesis (Martinez et al., 1991; Novella & Ebendick-Corp, 2004; Jerzak et al., 2007; Vignuzzi et al., 2006). It has become clear that viral swarm characterization is required if we are to fully evaluate the relationship between viral genomes and both viral evolution and epidemiology. In this study we determined the size of the SLEV mutant swarm in mosquitoes using molecular cloning of primary isolates from both Texas and California in order to (i) generally assess mutant swarm breadth, (ii) determine if geographical and/or temporal differences in mutant swarm size and composition exist, and (iii) elucidate the role of variable transmission cycles in shaping the mutant swarm. Advancing our
understanding of these relationships will provide insight into factors important in influencing arbovirus evolution and epidemiology.

**Materials and Methods**

*Virus isolates*

The main criterion used for the selection of SLEV isolates for this study was the availability of material from the initial amplification, i.e. single passage only (table 1). SLEV mosquito isolates from Texas were kindly provided by R. Tesh (UTMB; Galveston, TX) as lyophilized material following a single round of amplification on C6/36 mosquito cells. Material was reconstituted in 500ul BA-1 and used directly for experimental testing. Original California mosquito pools previously determined to be SLEV positive were obtained from A. Brault (UC Davis; Davis, CA) and amplified once on C6/36 mosquito cells at the Arbovirus laboratory, Wadsworth Center. Original mosquito pools from both TX and CA were of variable size (5-50 mosquitoes), yet minimal infection rates within each state have never exceeded a level which would suggest that each pool contained more than a single, positive mosquito (www.cdc.gov/ncidod/dvbid/arbor). All Texas isolates originated from *Cx. quinquefasciatus* mosquitoes trapped in the Houston area (Harris County) whereas CA isolates originated from *Cx. tarsalis* mosquitoes trapped in various counties including Coachella Valley (COAV), San Bernardino (SAND), Kern (BFS), Los Angeles (SOUE), and Imperial (IMPR) counties.
Molecular cloning

Molecular cloning of an approximately 2kb region (nucleotides 1315-3325), including portions of the envelope (E) and non-structural 1 (NS1) genes, was performed as previously described (Ciota et al., 2009). RNA was extracted from virus isolates using the Qia-amp viral RNA extraction kit (Qiagen, Valencia, CA) and RT-PCR was completed with primers designed to amplify the 3’ 1302 nt of the SLEV genome (E gene) and the 5’ 3325 nt of the SLEV genome (NS1 gene). Reverse transcription was performed with Sensiscript RT (Qiagen) and cDNA was amplified with ‘high-fidelity’ PfuUltra (published error rate = 4.3 X 10^{-7}; Stratagene, La Jolla, CA). PCR products were visualized on a 1.5 % agarose gel and DNA was recovered by using a MinElute Gel Extraction kit (Qiagen). The recovered DNA was ligated into the cloning vector pCR-Blunt II-TOPO (Invitrogen, Carlsbad, CA) and transformed into One Shot TOP10 Electrocomp E.coli cells. Colonies were screened by direct PCR using primers specific for the desired insert, and plasmid DNA was purified by using a QIAprep Spin Miniprep kit (Qiagen). Sequencing was carried out with five pairs of overlapping SLEV primers together with T7 and SP6 primers at the Wadsworth Center Molecular Genetics Core using ABI 3700 and 3100 automated sequencers (Applied Biosystems, Foster City, CA). Seventeen to twenty-four clones per sample were sequenced.

Phylogenetic trees
Four different programs, including BEAST v1.4.8 (Drummond & Rambaut, 2007), RaxML v7.2.3 (Stamatakis, 2006), GARLI v0.951 (Zwickl, 2006), and MrBayes (Ronquist & Huelsenbeck, 2003), were used to construct phylogenetic trees from the 255 amino acid sequences of SLEV. BEAST and MrBayes use Bayesian Markov-Chain-Monte-Carlo (MCMC) analyses to infer phylogenies, while RaxML and GARLI use Maximum-Likelihood-based approaches. BEAST was run once using Texas sequences and once for the California sequences, each time with an MCMC length of $1 \times 10^8$, logging every $1 \times 10^4$. A general time-reversible substitution model and uncorrelated log normal molecular clock model were used. A starting tree was constructed using UPGMA (Unweighted Pair Group Method with Arithmetic Mean) and the tree prior was assumed to have a constant population size. All of the priors for the model parameters and statistics were left at the defaults with the exception of including isolate date. RaxML was used to perform rapid bootstrap analyses with 1000 iterations while also finding a Maximum-Likelihood tree; all other parameters were left at default. GARLI was run in addition to RaxML and MrBayes in addition to BEAST in order to provide additional information on the structure of the phylogeny. GARLI was run using the default parameters except for including 100 independent search replicates. MrBayes was run for $1 \times 10^6$ generations each. RaxML, GARLI, and MrBayes were only run on sequence data from which duplicates had been removed. All phylogenetic reconstructions were rooted using sequences from four other flaviviruses as outgroups. These included Japanese encephalitis (genbank accession number M55506), West Nile (M12294), Kunjin (D00246) and Murray Valley encephalitis (X03467).
Sequence and correlation analyses

Sequences were compiled, edited, and aligned using DNASTAR software package (Madison, WI). At least 2-fold coverage was used for sequence determination. The percentage of nucleotide diversity (total number of differences from consensus divided by total number of bases sequenced), amino acid diversity (total number of amino acid differences divided by total number of amino acids sequenced), and the sequence diversity (percent of clones with at least one difference from consensus) were used as basic indicators of genetic diversity. For nucleotide and amino acid diversity calculations redundant changes were counted as separate differences.

The ratio (\( \omega \)) of nonsynonymous (dN) to synonymous (dS) substitutions per site was calculated in an attempt to evaluate the strength and direction of selection acting on these coding regions. Texas and California consensus sequences for all strains were broken up into their envelope (E) and non-structural (NS1) protein components and analyzed separately. Calculations were performed using HyPhy (Pond et al., 2005) on the Datamonkey webserver (Pond & Frost, 2005). The three methods used were single likelihood ancestor counting (SLAC), fixed effects likelihood (FEL), and random effects likelihood (REL) (Kosakovsky Pond & Frost, 2005).

Correlation analyses between nucleotide diversity and mosquito collection date were conducted with and without corrections for phylogenetic dependence. Correlation coefficients for nonphylogenetic analyses (treating each strain as independent) were calculated using the program R v2.9.2 (R Development Core Team, 2009). For
phylogenetic analyses, independent contrasts were used to correct for phylogeny (Felsenstein, 1985) using the APE package (Paradis et al., 2004) in R v2.9.2.

Results

Phylogeny

Phylogenetic analysis of Texas isolates yielded relationships in which all variants within individual strains grouped together (Fig. 1A). This pattern occurred in the trees created with all four programs and is illustrated here in the phylogeny reconstructed using RAxML. Duplicate sequences from the individual strains were removed in this analysis. The lack of intersection between any single variant among years demonstrates the uniqueness of consensus sequences from year to year and the general absence of yearly maintenance of past strains within the mutant swarm. Substitution rates on the consensus level were similar from year to year from 1990 to 1996, averaging about 0.9% divergence/year, or, approximately 10 nt substitutions/genome/year. Substantially greater divergence occurred from 1996 to 1998 (Fig. 1A), with almost 10% sequence divergence occurring in the two 1998 strains (V4179 and V3916) when compared to the 1996 isolate (V2679). 87.9% of these changes were synonymous. Interestingly, the 2003 isolate (V0476) does not group with the more recent 1998 isolate but, rather, the more distant isolates. In fact, the sequence most closely related to the 2003 isolate is the 1991 isolate (V3222), in which just 0.5% sequence divergence was measured.
The CA data set was very different from TX both spatially and temporally, with isolates as old as 1950 (BFS 508) and from various counties. In the phylogenetic tree constructed from CA sequences just 4 of 12 strains consistently grouped together (Fig. 1B). This pattern was again consistent among all four programs and is illustrated by the RAxML-created phylogeny in which duplicate sequences from individual strains were removed. In particular, multiple minority variants for isolates after 1993 were identical between strains and this section of the phylogeny exhibited low resolution. Of particular interest, all sequences from a 1993 isolate from San Bernardino (SAND 29) were placed within this highly related clade and one mutant sequence from a 2000 isolate from Coachella valley (COAV 332) was identical to the consensus sequence from San Bernardino (SAND 29). Other than this exception all strains generally grouped both temporally and geographically. Post-1993 strains demonstrated high similarity along with a lack of phylogenetic resolution. Attempts to quantify substitution rates of SLEV in CA are clearly inaccurate when done without consideration of geographic location. For example, just 7 nt differences exist between the consensus sequences of the 1950 and 1960 isolates from BFS in the region analyzed, yet 184 substitutions were identified in this region when comparing the 1984 SOUE isolate (SOUE 20) and a 1993 San Bernardino isolate (SAND 29). When only Coachella Valley isolates are considered independently (2000-2003), the average substitution rates are approximately 12 substitutions / genome/year, which is similar to what was measured with the Texas isolates.

The phylogeny reconstructed using California consensus sequences again showed very little differentiation in the area containing the eight highly related strains (Fig. S1B).
In the best-scoring maximum likelihood tree reconstructed using RAxML this region only had one bipartition that appeared in more than 80 percent of bootstrap replicates. The rest of the bipartitions showed high bootstrap support.

*Genetic diversity and correlation*

Isolation of minority variants by high-fidelity molecular cloning of portions of the SLEV ENV and NS1 genes and subsequent sequencing and analysis was used to assess the level of genetic diversity within individual isolates. The purpose of these analyses was to assess both the size of the mutant swarm in recent isolates and to determine if mutant-swarm size was correlated with both temporal, geographic, and host differences. Initial analysis involved a straightforward assessment of nucleotide and amino acid diversity and overall sequence diversity (Tables 2 and 3).

The most genetically diverse isolate among the Texas strains, V3222 (1991), possessed 28 variable bases within the 24 clones analyzed from the mutant swarm. Assuming this level of genetic diversity exists genome wide, this would equate to an average of 6.4 base differences relative to the consensus sequence per genome. Among the 58 base substitutions identified in the three most diverse isolates (1990, 1991, 1992), 39 were nonsynonomous changes (67.2%). Conversely, a total of just 2 base changes relative to consensus sequences were identified in the three most recent isolates (V0476, V3916, V4179), with not a single change identified in 23 clones of the 2003 isolate (V0476).
Without correcting for phylogeny, a negative correlation between isolation date and nucleotide diversity among the Texas isolates exists \( (R=-0.788, p=0.020; \text{Table 4}) \) (Fig 2A). This relationship, with the most highly homogeneous isolates being the most recent, also exists on the amino-acid level \( (R=-0.823, p=.012; \text{Table 5}) \). However, the isolates are not independent, and thus the correlation is confounded by the phylogenetic structure. The phylogeny that was used to correct for this dependence was created with RAxML and used only the consensus sequences from each strain. When phylogeny of the Texas isolates is corrected for by independent contrasts the correlation between time and intrahost diversity is further strengthened \( (R = -0.891, p = 0.003; \text{Table 4, Fig. 3A}) \). However, this strong correlation is driven by a single outlying contrast and disappears when this point is removed from the set \( (R = -0.341, p = 0.454) \). A Spearman rank correlation, which is less sensitive to outliers, similarly found no correlation, even when the outlying contrast was not excluded \( (\rho = -0.464, p = 0.302; \text{Table 4}) \). The outlying contrast driving the correlation is a result of a strain from 2003 (U0476) with low diversity being closely related to a strain from 1991 (V3222) with a high diversity.

On average, recent California isolates (2000-2003) also displayed high levels of intrahost genetic homogeneity, with all of the isolates having nucleotide diversity of less than 0.016\% (table 2). The most highly homogeneous isolates were again the most recent isolates, with two of the three 2003 CA isolates (both COAV) showing just a single base change in more than 40,000 bases sequenced. In addition, the most highly diverse isolate analyzed was the oldest, BFS 508 (1950), with a nucleotide diversity of 0.078\%, equating to an average of 8.6 substitutions per genome. Unlike the diverse Texas isolates, all but one of the substitutions in this isolate were synonymous. Overall, the
percent of nonsynonymous changes within the CA isolates mutant swarms was
significantly lower than that measured in Texas (32.0% v. 62.3%; Fisher’s exact test, \( p < 0.0001 \)).

The CA isolates showed a negative correlation between nucleotide diversity and
isolation date (\( R = -0.791, p = 0.002; \) Table 4, Fig. 2B), similar to the Texas isolates.
When independent contrasts were performed using a phylogeny of consensus sequences
the correlation strengthened greatly as with the Texas data (\( R = -0.989; p < 0.0001; \) Table
4; Fig. 3B). Unlike for the Texas data, the increased correlation after correcting for
phylogenetic structure cannot easily be attributed to a single outlier although there does
appear to be one (Fig. 3B). The correlations remains significant when the outlier is
removed (\( R = -0.720, p = 0.013; \) Table 4) and when the non-parametric analysis is
performed (\( \rho = -0.791, p = 0.006 \)).

The Texas sequences show a similar relationship between diversity and time on
the amino-acid level without correcting for phylogeny as with nucleotides (\( R = -0.822, p
= 0.012; \) Table 5). The negative relationship seen at the nucleotide level for the California
sequences is absent for amino-acids (\( R = -0.216, p = 0.500; \) Table 5). When independent
contrasts were performed on the Texas data, the correlation disappeared (\( R = -0.483, p =
0.226; \) Table 5) but when this method was applied to the California amino-acid data, the
strength of the correlation greatly increased (\( R = -0.993, p < 0.0001; \) Table 5). Unlike at
the nucleotide level, the correlation between time and amino-acid diversity disappeared
when the outlier was removed (\( R = -0.019, p = 0.956; \) Table 5) and when we calculated a
Spearman correlation (\( \rho = -0.109, p = 0.755 \)).
Selection

There was not enough power to infer selection at individual sites due to the relatively low number of sequences. We used three different methods provided by the DataMonkey web server (Single Likelihood Ancestor Counting, Fixed Effects Likelihood, Random Effects Likelihood), and found no site for which all methods indicated significant selection (data not shown). However, the mean $\omega$ values for the E gene were 0.017736 and 0.020650 in Texas and California, respectively, while the mean $\omega$ values for the NS1 gene were 0.065975 and 0.093021 in Texas and California, respectively (Table 6). This difference suggests that there has been a greater degree of purifying selection acting on the envelope protein than on the non-structural protein in both geographic areas.

Discussion

To some extent, all RNA viruses exist as a compilation of closely related variants, i.e. a mutant swarm (Biebricher & Eigen, 2006). Despite the fact that this swarm has been clearly implicated in viral fitness, pathogenesis, and adaptability, evolutionary studies of natural populations to date rely almost exclusively on consensus sequence analyses. Mutant swarm size and composition may be particularly important for arboviruses, which require replication not only in diverse tissues but also divergent hosts. Here, we evaluated the size and composition of SLEV swarms in single-passage mosquito isolates from Texas and California. Although our data are somewhat limited by the availability
of low passage isolates, significant insight into how time, space, and transmission cycles contribute to shaping the SLEV mutant swarm was gained.

The Texas data set provided an opportunity to monitor genetic changes in a single location over time (Table 1). The Texas phylogeny clearly demonstrates that all variants within each isolate group together (Fig. 1A). The fact that variants within the same strain are more closely related to each other than other strains is not surprising, yet it is interesting that no two isolates from different years share a single variant. This, together with the relationships depicted in both the consensus and variant phylogenetic trees, demonstrates that a unique strain may take hold each successive year in a time-independent fashion (Figs. 1A). For example, the 2003 strain is more closely related to the 1991 strain than it is to any of the more recent strains, while the 1998 strains are relatively distant from all other sequences (Figs 1A and S1A). This suggests that either a new strain was introduced each successive year or that multiple strains coexist locally. Although reintroduction does occasionally occur (Kramer & Chandler, 2001), previous studies demonstrate that the latter (i.e. local maintenance) is more likely (May et al., 2008). Without analysis of multiple strains from each year, it is not entirely clear if significant inter-host variation exists, but comparison of the sequences from the two 1998 strains suggest that this is not the case, and that a single strain likely dominates annually. The variable levels of activity in each successive year in Texas do not necessarily suggest an adaptive process (www.cdc.gov/ncidod/dvbid/arbor), yet individual strains may be better suited for specific environmental conditions present in different years. In addition, the percent of nonsynonomous change both on the consensus level from year to year and within the mutant swarm is not substantially different from what one would
predict given a general lack of selection (Holmes, 2003). It is reasonable to assume that genetic diversity could decrease over time as a virus becomes more specialized to the hosts and ecology unique to a particular region, yet one would expect a dominance of positive selection in this case. The lack of strong statistical support for a negative correlation between intra-host heterogeneity and time when controlling for phylogeny and the outlier (Table 4) agrees with the results of the dN/dS which support a dominance of purifying selection (Table 6). Since the initially observed correlation was confounded by the phylogenetic structure of the sequences, the currently low swarm sizes might possibly be explained by other factors independent of phylogeny.

The source of all isolates from the Houston area is *Cx. quinquefasciatus* mosquitoes, which is the primary vector responsible for urban transmission in the southwest (Reisen, 2003). These mosquitoes are unique among primary vectors of SLEV in that they do not enter diapause (Hayes, 1975). As a result, seasonal maintenance may be achieved through a continuous low level of horizontal transmission, which could equate to significant seasonal bottlenecks during periods of decreased activity (Bellamy, Reeves et al., 1968). In addition, *Cx. quinquefasciatus* have been shown to be less susceptible to SLEV infection than *Cx. tarsalis* (Meyer et al., 1983; Hardy 1990). Little is known about the specifics of within-host bottlenecks in mosquito vectors, but the potential for bottlenecks exists at well-documented barriers including midgut infection and egress, and salivary gland infection and exit (Kramer et al., 1981; Girard et al., 2004). For these reasons, changes in vector competence and even more, in vectorial capacity, could play a significant role in determining the effective viral population size within specific vector populations.
The general phylogenetic grouping of CA strains from the same counties argues against drawing conclusions on the effect of time without consideration of geographic location (Fig. 1B). However, when COAV isolates are considered independently, a negative correlation between time and intrahost diversity is present although it is far less convincing on the amino-acid level (tables 4 and 5). A larger data set would help to clarify if there is any clear temporal pattern with intrahost diversity within individual regions. The other unique characteristic of the CA strains is that, unlike Texas strains, some variants do in fact group within other isolates (Fig. 1B). This is particularly the case with COAV isolates, where there is significant intersection of sequences in all the represented strains. This degree of relatedness agrees with previous finding that yearly strain maintenance is likely common in California (Reisen, 2003). In addition to the specific role of the vector in shaping the viral swarm, a difference on the avian host side may exist. Experimental passage of SLEV in both chickens and Cx. pipiens mosquitoes has demonstrated that the size of the mutant swarm is host-dependent (Ciota et al., 2009). It is reasonable to assume that the size of the mutant swarm could also be species-dependent. Variation in host preference and host abundance has been demonstrated at the sites from which these strains were isolated and could play a major role in shaping individual populations (Reisen, 2003; Lothrop & Reisen, 2001). Studies assessing viremia levels in various bird species following infection with CA strains of SLEV found that very few species produce viremia levels sufficient for infection of mosquitoes (Reisen et al., 2001; Reisen et al., 2003), demonstrating that the availability of competent hosts could be a significant limiting factor on both strain diversity and overall activity.
Despite the difference in the relationship between time and intrahost diversity seen in California and Texas, all of the recent isolates from both locations are highly genetically homogeneous (tables 2 and 3). Although the requirement of all arboviruses to cycle between divergent hosts may restrict genetic change (Holmes, 2003), the extent to which these isolates lack genetic variability is somewhat surprising given the high mutation rate of RNA viruses (Drake & Holland, 1999). A previous study assessing the level of intrahost diversity within the same genomic region of WNV demonstrated that significant genetic heterogeneity exists in isolates from *Cx. pipiens* mosquitoes from New York (Jerzak et al., 2005). Subsequent studies with adapted populations of WNV and SLEV demonstrated that differences in mutant swarm size may have significant implications for host range (Ciota et al., 2007). As with all organisms, a general lack of genetic diversity limits the ability to exploit new and changing ecology, and could therefore be an important factor in predicting exploit of new habitats and epidemiological patterns in changing landscapes. West Nile virus invaded Texas and California in 2002 and 2003, respectively (Reisen et al., 2004; Lillibridge et al., 2004). Interestingly, WNV has appeared to displace SLEV in California yet the two continue to coexist in Texas (Rios et al., 2006; Reisen et al., 2008). The combination of the close antigenic relationship and the sharing of transmission cycles by WNV and SLEV make their interaction inevitable. WNV antibody-positive birds will possess at least partial protection from SLEV and, given the lower viremia and subsequently more limited effective vector range, it is no surprise that WNV could displace SLEV in areas where they coexist. The Alphavirus *western equine encephalomyelitis virus* reemerged in southeastern California in 2005, demonstrating that an antigenically distinct virus which
utilizes the same transmission cycle is capable of coexisting in the region (Hom et. al, 2006). Genetic diversity within a viral swarm could also correspond to antigenic diversity with variable levels of cross-reactivity to closely related viruses, which in turn could have direct effects on the capacity to evade competition. In addition, the existence of WNV could itself drive the purification of the SLEV population, which could potentially contribute to the current lack of intrahost diversity in both California and Texas.

Overall, this study demonstrates the power of mutant spectrum analysis relative to traditional consensus sequencing. Evaluation of relationships among consensus sequences have historically provided valuable information in the understanding of virus evolution and epidemiology, yet it is now clear that our conclusions to date, particularly with regards to RNA viruses, are drawn from potentially incomplete data sets. Although the methods presented here are currently not feasible on a large scale, recent technological advances in the ability to perform high throughput deep sequencing are beginning to provide researchers with the ability to consider variation in viral swarms on a broader and more precise level. Such analyses are likely to reveal characteristics and relationships among virus variants that have broad implications for viral evolution, fitness, and epidemiology.
Acknowledgments

Evan M. Koch, Graham G. Willsey, Lauren J. Davis, Gretta V.S. Jerzak, Dylan J. Ehrbar, Claus O. Wilke, and Laura D. Kramer are also authors on the originally published manuscript and all contributed to these studies. The authors would like to thank R. Tesh and The World Reference Center for Emerging Viruses and Arboviruses (WRCEVA), through the University of Texas Medical Branch (UTMB) in Galveston, Texas, for providing Texas mosquito isolates. Additionally, we thank A. Brault, and W. Reisen for providing Texas mosquito isolates from California. We also thank J. Maffei for assistance with virus amplification. The authors thank the Wadsworth Center Molecular Genetics Core for sequencing, and the Wadsworth Center Media and Tissue Culture Facility for providing cells and media for this work. This work was supported partially by federal funds from the National Institute of Allergy and Infectious Disease, National Institutes of Health (contract number NO1-AI-25490) and, National Institutes of Health (grant numbers RO1-AI-077669). The construction of the Wadsworth Center Insectary Facility was funded by NIH grant number C06-RR-17715.

References


Chapter 2 tables

Table 1. SLEV mosquito isolates.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Collection date</th>
<th>Location</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>U0476</td>
<td>22 July 2003</td>
<td>Harris Co., TX</td>
<td>Cx. quinquefasciatus</td>
</tr>
<tr>
<td>V3916</td>
<td>29 Sept 1998</td>
<td>Harris Co., TX</td>
<td>Cx. quinquefasciatus</td>
</tr>
<tr>
<td>V4179</td>
<td>6 Oct 1998</td>
<td>Harris Co., TX</td>
<td>Cx. quinquefasciatus</td>
</tr>
<tr>
<td>V2679</td>
<td>8 Aug 1996</td>
<td>Harris Co., TX</td>
<td>Cx. quinquefasciatus</td>
</tr>
<tr>
<td>V1042</td>
<td>28 June 1994</td>
<td>Harris Co., TX</td>
<td>Cx. quinquefasciatus</td>
</tr>
<tr>
<td>V3248</td>
<td>29 July 1992</td>
<td>Harris Co., TX</td>
<td>Cx. quinquefasciatus</td>
</tr>
<tr>
<td>V3222</td>
<td>14 Aug 1991</td>
<td>Harris Co., TX</td>
<td>Cx. quinquefasciatus</td>
</tr>
<tr>
<td>V4209</td>
<td>22 Aug 1990</td>
<td>Harris Co., TX</td>
<td>Cx. quinquefasciatus</td>
</tr>
<tr>
<td>COAV888</td>
<td>2 July 2003</td>
<td>Coachella Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>COAV1257</td>
<td>2 Oct 2003</td>
<td>Coachella Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>IMPR115</td>
<td>16 Sept 2003</td>
<td>Imperial Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>COAV382</td>
<td>11 July 2001</td>
<td>Coachella Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>COAV333</td>
<td>11 July 2001</td>
<td>Coachella Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>COAV332</td>
<td>15 July 2000</td>
<td>Coachella Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>COAV333</td>
<td>15 July 2000</td>
<td>Coachella Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>SAND29</td>
<td>27 Aug 1993</td>
<td>San Bern. Co, CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>SAND39</td>
<td>27 Aug 1993</td>
<td>San Bern. Co, CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>SOUE20</td>
<td>18 Sept 1984</td>
<td>Los Ang. Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>BFS2874</td>
<td>22 Sept 1960</td>
<td>Kern Co., CA</td>
<td>Cx. tarsalis</td>
</tr>
<tr>
<td>BFS508</td>
<td>1 August 1950</td>
<td>Kern Co, CA</td>
<td>Cx. tarsalis</td>
</tr>
</tbody>
</table>
Table 2. Intrahost nucleotide diversity of SLEV isolates for nt 1315-3325.

<table>
<thead>
<tr>
<th>strain</th>
<th>state</th>
<th>year</th>
<th>clones</th>
<th>nt changes</th>
<th>nt div.(^1)</th>
<th>nc var.(^2)</th>
<th>seq div.(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>U0476</td>
<td>TX</td>
<td>2003</td>
<td>23</td>
<td>0</td>
<td>0.000%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>V3916</td>
<td>TX</td>
<td>1998</td>
<td>24</td>
<td>1</td>
<td>0.002%</td>
<td>1</td>
<td>4.17%</td>
</tr>
<tr>
<td>V4179</td>
<td>TX</td>
<td>1998</td>
<td>24</td>
<td>1</td>
<td>0.002%</td>
<td>1</td>
<td>4.17%</td>
</tr>
<tr>
<td>V2679</td>
<td>TX</td>
<td>1996</td>
<td>24</td>
<td>2</td>
<td>0.004%</td>
<td>2</td>
<td>8.33%</td>
</tr>
<tr>
<td>V1042</td>
<td>TX</td>
<td>1994</td>
<td>24</td>
<td>7</td>
<td>0.015%</td>
<td>7</td>
<td>29.17%</td>
</tr>
<tr>
<td>V3248</td>
<td>TX</td>
<td>1992</td>
<td>24</td>
<td>18</td>
<td>0.037%</td>
<td>14</td>
<td>58.33%</td>
</tr>
<tr>
<td>V3222</td>
<td>TX</td>
<td>1991</td>
<td>24</td>
<td>28</td>
<td>0.058%</td>
<td>17</td>
<td>70.83%</td>
</tr>
<tr>
<td>V4209</td>
<td>TX</td>
<td>1990</td>
<td>24</td>
<td>12</td>
<td>0.025%</td>
<td>9</td>
<td>37.50%</td>
</tr>
<tr>
<td>COAV888</td>
<td>CA</td>
<td>2003</td>
<td>19</td>
<td>1</td>
<td>0.008%</td>
<td>1</td>
<td>5.26%</td>
</tr>
<tr>
<td>COAV1257</td>
<td>CA</td>
<td>2003</td>
<td>20</td>
<td>1</td>
<td>0.007%</td>
<td>1</td>
<td>5.00%</td>
</tr>
<tr>
<td>IMPR115</td>
<td>CA</td>
<td>2003</td>
<td>19</td>
<td>5</td>
<td>0.013%</td>
<td>3</td>
<td>15.79%</td>
</tr>
<tr>
<td>COAV382</td>
<td>CA</td>
<td>2001</td>
<td>19</td>
<td>6</td>
<td>0.016%</td>
<td>6</td>
<td>31.58%</td>
</tr>
<tr>
<td>COAV383</td>
<td>CA</td>
<td>2001</td>
<td>19</td>
<td>6</td>
<td>0.016%</td>
<td>3</td>
<td>15.79%</td>
</tr>
<tr>
<td>COAV332</td>
<td>CA</td>
<td>2000</td>
<td>18</td>
<td>3</td>
<td>0.008%</td>
<td>2</td>
<td>11.11%</td>
</tr>
<tr>
<td>COAV333</td>
<td>CA</td>
<td>2000</td>
<td>19</td>
<td>2</td>
<td>0.005%</td>
<td>1</td>
<td>5.26%</td>
</tr>
<tr>
<td>SAND29</td>
<td>CA</td>
<td>1993</td>
<td>19</td>
<td>7</td>
<td>0.021%</td>
<td>1</td>
<td>5.88%</td>
</tr>
<tr>
<td>SAND39</td>
<td>CA</td>
<td>1993</td>
<td>17</td>
<td>11</td>
<td>0.029%</td>
<td>9</td>
<td>47.37%</td>
</tr>
<tr>
<td>SQUE20</td>
<td>CA</td>
<td>1984</td>
<td>20</td>
<td>2</td>
<td>0.005%</td>
<td>1</td>
<td>5.00%</td>
</tr>
<tr>
<td>BFS2874</td>
<td>CA</td>
<td>1960</td>
<td>19</td>
<td>9</td>
<td>0.024%</td>
<td>6</td>
<td>31.58%</td>
</tr>
<tr>
<td>BFS508</td>
<td>CA</td>
<td>1950</td>
<td>21</td>
<td>33</td>
<td>0.078%</td>
<td>7</td>
<td>33.33%</td>
</tr>
</tbody>
</table>

\(^1\) nucleotide diversity = total nt changes (differences from consensus) / total bases sequenced (~45,000)
\(^2\) nonconsensus variants = haplotypes with at least one nt change relative to consensus
\(^3\) sequence diversity = % of clones with at least one nt change
Table 3. Intrahost amino acid diversity of SLEV isolates for aa 439-1109.

<table>
<thead>
<tr>
<th>strain</th>
<th>state</th>
<th>year</th>
<th>clones</th>
<th>aa changes</th>
<th>aa div.</th>
<th>nc var</th>
<th>seq div.</th>
</tr>
</thead>
<tbody>
<tr>
<td>U0476</td>
<td>TX</td>
<td>2003</td>
<td>23</td>
<td>0</td>
<td>0.000%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>V3916</td>
<td>TX</td>
<td>1998</td>
<td>24</td>
<td>1</td>
<td>0.006%</td>
<td>1</td>
<td>4.17%</td>
</tr>
<tr>
<td>V4179</td>
<td>TX</td>
<td>1998</td>
<td>24</td>
<td>0</td>
<td>0.000%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>V2679</td>
<td>TX</td>
<td>1996</td>
<td>24</td>
<td>1</td>
<td>0.006%</td>
<td>1</td>
<td>4.17%</td>
</tr>
<tr>
<td>V1042</td>
<td>TX</td>
<td>1994</td>
<td>24</td>
<td>3</td>
<td>0.019%</td>
<td>3</td>
<td>12.50%</td>
</tr>
<tr>
<td>V3248</td>
<td>TX</td>
<td>1992</td>
<td>24</td>
<td>13</td>
<td>0.081%</td>
<td>11</td>
<td>45.83%</td>
</tr>
<tr>
<td>V3222</td>
<td>TX</td>
<td>1991</td>
<td>24</td>
<td>16</td>
<td>0.099%</td>
<td>13</td>
<td>54.17%</td>
</tr>
<tr>
<td>V4209</td>
<td>TX</td>
<td>1990</td>
<td>24</td>
<td>10</td>
<td>0.062%</td>
<td>8</td>
<td>33.33%</td>
</tr>
<tr>
<td>COAV888</td>
<td>CA</td>
<td>2003</td>
<td>19</td>
<td>1</td>
<td>0.008%</td>
<td>1</td>
<td>5.26%</td>
</tr>
<tr>
<td>COAV1257</td>
<td>CA</td>
<td>2003</td>
<td>20</td>
<td>1</td>
<td>0.007%</td>
<td>1</td>
<td>5.00%</td>
</tr>
<tr>
<td>IMPR115</td>
<td>CA</td>
<td>2003</td>
<td>19</td>
<td>1</td>
<td>0.007%</td>
<td>1</td>
<td>5.00%</td>
</tr>
<tr>
<td>COAV382</td>
<td>CA</td>
<td>2001</td>
<td>19</td>
<td>0</td>
<td>0.000%</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>COAV383</td>
<td>CA</td>
<td>2001</td>
<td>19</td>
<td>3</td>
<td>0.024%</td>
<td>2</td>
<td>10.53%</td>
</tr>
<tr>
<td>COAV332</td>
<td>CA</td>
<td>2000</td>
<td>18</td>
<td>1</td>
<td>0.008%</td>
<td>1</td>
<td>5.55%</td>
</tr>
<tr>
<td>COAV333</td>
<td>CA</td>
<td>2000</td>
<td>19</td>
<td>1</td>
<td>0.007%</td>
<td>1</td>
<td>5.00%</td>
</tr>
<tr>
<td>SAND29</td>
<td>CA</td>
<td>1993</td>
<td>19</td>
<td>6</td>
<td>0.052%</td>
<td>5</td>
<td>26.32%</td>
</tr>
<tr>
<td>SAND39</td>
<td>CA</td>
<td>1993</td>
<td>17</td>
<td>2</td>
<td>0.015%</td>
<td>1</td>
<td>5.88%</td>
</tr>
<tr>
<td>SOUE20</td>
<td>CA</td>
<td>1984</td>
<td>20</td>
<td>1</td>
<td>0.007%</td>
<td>1</td>
<td>5.00%</td>
</tr>
<tr>
<td>BFS2874</td>
<td>CA</td>
<td>1960</td>
<td>19</td>
<td>5</td>
<td>0.039%</td>
<td>3</td>
<td>15.79%</td>
</tr>
<tr>
<td>BFS508</td>
<td>CA</td>
<td>1950</td>
<td>21</td>
<td>1</td>
<td>0.007%</td>
<td>1</td>
<td>4.76%</td>
</tr>
</tbody>
</table>

1 amino acid diversity = total aa changes (differences from consensus) / total sequenced (~15,000)
2 nonconsensus variants = aa haplotypes with at least one aa change relative to consensus
3 sequence diversity = % of clones with at least one aa change
Table 4. Correlation analysis between year and intrahost nucleotide diversity

<table>
<thead>
<tr>
<th>data set</th>
<th>analysis</th>
<th>R</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texas</td>
<td>raw data</td>
<td>-0.788</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>PIC(^1)</td>
<td>-0.891</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>PIC w/o outlier</td>
<td>-0.341</td>
<td>0.454</td>
</tr>
<tr>
<td>California</td>
<td>raw data</td>
<td>-0.791</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>PIC(^1)</td>
<td>-0.989</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>PIC w/o outlier</td>
<td>-0.720</td>
<td>0.013</td>
</tr>
</tbody>
</table>

\(^1\) PIC refers to phylogenetic independent contrasts

Table 5. Correlation analysis between year and intrahost amino-acid diversity

<table>
<thead>
<tr>
<th>data set</th>
<th>analysis</th>
<th>R</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texas</td>
<td>raw data</td>
<td>-0.822</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>PIC(^1)</td>
<td>-0.483</td>
<td>0.226</td>
</tr>
<tr>
<td></td>
<td>PIC w/o outlier</td>
<td>-0.341</td>
<td>0.454</td>
</tr>
<tr>
<td>California</td>
<td>raw data</td>
<td>-0.216</td>
<td>0.500</td>
</tr>
<tr>
<td></td>
<td>PIC(^1)</td>
<td>-0.993</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>PIC w/o outlier</td>
<td>-0.019</td>
<td>0.956</td>
</tr>
</tbody>
</table>

\(^1\) PIC refers to phylogenetic independent contrasts

Table 6. dN/dS analysis

<table>
<thead>
<tr>
<th>data set</th>
<th>protein</th>
<th>length (aa)</th>
<th>mean ω</th>
</tr>
</thead>
<tbody>
<tr>
<td>Texas</td>
<td>E</td>
<td>382</td>
<td>0.017736</td>
</tr>
<tr>
<td></td>
<td>NS1</td>
<td>285</td>
<td>0.065975</td>
</tr>
<tr>
<td>California</td>
<td>E</td>
<td>380</td>
<td>0.020650</td>
</tr>
<tr>
<td></td>
<td>NS1</td>
<td>284</td>
<td>0.093021</td>
</tr>
</tbody>
</table>
Chapter 2 Figures
Figure 1. Phylogenetic-tree reconstruction of individual isolates, using RAxML. Duplicate sequences have been removed for this analysis. (A) Texas isolates from the same sample cluster together. Numbers on the right indicate the year of sample collection. (B) California isolates from different samples (obtained at different time points and different locations) grouped together. Multiples following the year of sample collection indicate how many independent samples are part of the same group; i.e., the top group consists of three samples taken in 2003, two taken in 2001, two in 2000, and one in 1993.

Figure 2. Nucleotide diversity vs. isolation year, not corrected for phylogeny. (A) Texas. (B) California.
CHAPTER 3: QUANTIFICATION OF INTRAHOST BOTTLENECKS OF *WEST NILE VIRUS* IN *CULEX PIPiens* MOSQUITOES USING AN ARTIFICIAL MUTANT SWARM*

* The majority of this material has appeared previously in Ciota et al. 2012, Infect Genet Evol. Apr;12(3):557-64.

Abstract

Mosquito-borne viruses are predominantly RNA viruses which exist within hosts as diverse mutant swarms. Defining the way in which stochastic forces within mosquito vectors shape these swarms is critical to advancing our understanding of the evolutionary and adaptive potential of these pathogens. There are multiple barriers within a mosquito which a viral swarm must traverse in order to ultimately be transmitted. Here, using artificial mutant swarms composed of neutral variants of West Nile virus (WNV), we tracked changes to swarm breadth over time and space in *Culex* mosquitoes. Our results
demonstrate that all variants have the potential to survive intrahost bottlenecks, yet mean swarm breadth decreases during both midgut infection and transmission in *Cx. pipiens* when starting populations contain higher levels of minority variants. In addition, WNV swarms are subject to temporal sweeps which act to significantly decrease intrahost diversity in *Cx. pipiens* over time. Lastly, similar experiments with *Cx. quinquefasciatus* demonstrate a lack of significant intrahost bottlenecks. Taken together, these data demonstrate the profound effects that stochastic forces can have in shaping arboviral mutant swarms, but also clearly show that the extent of intrahost bottlenecks is species-specific.

**Introduction**

Mosquito-borne viruses are predominantly RNA viruses which exist within hosts as diverse mutant swarms. A number of studies to date have established the significant role of this swarm in viral fitness, adaptation, and pathogenesis (Ciota and Kramer, 2010), yet the specific dynamics altering swarm composition and breadth within hosts remain largely uncharacterized. There are multiple barriers a viral swarm must traverse within a mosquito in order to ultimately be transmitted, each of which could act as a significant bottleneck of genetic diversity. Viral infection in competent mosquitoes begins with acquisition of infectious virus during bloodfeeding and proceeds through infection of, replication within, and dissemination from the mesenteron (midgut). Secondary amplification occurs in parenteral tissues, including fat body cells and nerve tissue,
allowing infection of critical target organs such as the salivary glands and, ultimately, virus transmission during subsequent bloodfeeding (Girard et al., 2004; Kramer and Ebel, 2003). *West Nile virus* (WNV, *Flaviviridae, Flavivirus*), the most widespread and medically important arthropod-borne virus in the United States, is transmitted primarily by *Culex* spp. mosquitoes (Kramer et al., 2007). Previous studies demonstrate that WNV exists as a diverse viral swarm in nature and that the mosquito is the primary source of this diversity (Jerzak et al., 2005). This mosquito derived diversity likely results from a combination of relaxed purifying selection and density-dependent selection needed to overcome the innate RNAi immune response (Jerzak et al., 2008; Brackney et al., 2009). Despite these findings, passage studies of WNV in *Cx. pipiens* using virus secreted in saliva resulted in a highly homogeneous viral swarm, suggesting that the diversity generated in the mosquito may often be purged by within-host bottlenecks prior to transmission (Ciota et al., 2008). Although it has been shown that *Culex* spp. mosquitoes are capable of transmitting up to $10^6 \log_{10}$ plaque forming units of WNV, the genetic diversity within this transmitted population has not been fully characterized (Styer et al., 2007). A recent study with *Culex quinquefasciatus* demonstrates that diversity may be maintained during transmission by this species (Brackney et al., 2011), yet *Culex* species often differ significantly in their vector competence (Kilpatrick et al., 2010), suggesting the extent of within-host bottlenecking is also likely species-specific. Here, we characterized time and tissue-dependent changes in mutant swarm breadth in *Culex pipiens* mosquitoes using an artificial swarm of neutral WNV variants that could be tracked over time and space. Our results demonstrate the extent to which the breadth of
intrahost viral diversity may be limited by both temporal and spatial variation in mosquito hosts.

**Materials and Methods**

**Mosquitoes**

_Cx. pipiens_ egg rafts were originally collected in Pennsylvania in 2004 (courtesy of M. Hutchinson) and subsequently colonized at the Arbovirus laboratory, Wadsworth Center. _Cx. quinquefasciatus_ colony mosquitoes were derived from a laboratory colony provided by D. Fonseca (Rutgers Univ.) and initially derived from egg rafts from Benzon Research Inc. (Carlisle, PA). Mosquitoes were reared and maintained in 30.5 cm$^3$ cages in an environmental chamber at 27°C, 50-65% relative humidity with a photoperiod of 16:8 (light:dark) hours. 1500-2000 adult female mosquitoes to be used for experimental infections were collected upon emergence and held in mesh top 3.8 L paper cartons and provided cotton pads with 10% sucrose ad libitum. Mosquitoes were held for 7-10 days prior to bloodfeeding.

**Virus strains**

The _West Nile virus_ infectious clone (FL-WNV) virus was generated from an infectious cDNA clone based on New York strain 3356, isolated from an American Crow (_Corvus brachyrhynchos_) in Staten Island in 2000 (AF404756). Methods for clone manipulation and rescue of infectious WNV are as described previously (Shi et al., 2002). Insertions of silent changes in the NS1 gene of WNV, which has no known role in viral fitness in the mosquito vector, were chosen as markers for individual variants. WNV mutants were
generated by site-directed mutagenesis (SDM) of the FL-WNV using the QuikChange XLII SDM kit (Stratagene, La Jolla, CA) as per the manufacturer’s protocol. Mutant FL-WNV DNA was then amplified in *E. coli* and plasmid harvested by Highspeed Midiprep (Qiagen, Valencia, CA). Sequencing of NS1 mutant WNV plasmids confirmed that desired changes were engineered. All purified mutant and control FL-WNV plasmids were linearized with XbaI and transcribed using the MEGAscript kit (Invitrogen, Carlsbad, CA) supplemented with Anti-reverse cap analog (Invitrogen) and assembled as per manufacturer’s protocol. Transcription reactions were incubated at 37 °C for 4 hrs.

Resulting RNA was purified with the MEGAclear kit (Invitrogen) and quantified on a Nanodrop 2000 (Thermo Fisher Scientific, Waltham, MA). RNA was stored in 10µg aliquots at -80 °C.

Wild-type FL-WNV RNA (variant 1) and mutant RNA including: WNV T2700C (variant 2), T2913C (variant 3), T2940C (variant 4), A2844G (variant 5), T3117C (variant 6), A2847T (variant 7) and T2910C (variant 8) were electroporated into 0.8 x 10^7 C6/36 mosquito cells in PBS using a GenePulser (BioRad). Transfected cells were seeded into T75 flasks and supernatants were collected from day 3 to 7 post-transfection, aliquoted and stored at -80 °C. WNV titers were quantified by plaque assay on Vero cells as previously described (Payne et al., 2006) and levels of WNV RNA were quantified by Taqman qRT-PCR (Applied Biosystems, Foster City, CA). WN S5 was created by mixing equivalent amounts (based on RNA copies) of stock of variants 1-5. Mixtures of variants 2 and 4 were used for proportion experiments and proportions were again based on RNA copy numbers. WN S8 was created by mixing equivalent amounts of stock of variants 1-8 followed by a single amplification of this mixture on mosquito cell culture (C6/36, ATCC #CRL-1660) in order to increase infectivity for feeding
experiments. Exact proportions of variants used for infectious feedings were determined by molecular cloning and sequencing of experimental bloodmeals as described below.

*Individual strain characterization*

All individual variants were assessed for vector competence and growth kinetics, as well as genetic stability in *Cx. pipiens* mosquitoes following feeding on an infectious bloodmeal. Experimental infections were carried out as previously described (Ciota et al., 2009). Briefly, individual WNV variants were diluted to equivalent titers (~9.0 log\(_{10}\) pfu/ml), mixed 1:5, virus: defibrinated goose blood (Hema Resources Inc, Aurora, OR) + 2.5% sucrose, and offered to ~500 female mosquitoes using a Hemotek membrane feeding system (Discovery Workshops, Accrington, UK). Following 1-hour, mosquitoes were anesthetized using CO\(_2\) and fully engorged mosquitoes were saved and housed at 27°C for subsequent testing. Twenty-five to 50 mosquitoes per variant were tested for vector competence at days 5, 7, 10, 14 and 21 days post-feeding. Mosquitoes were incapacitated and legs were removed and placed in 1 ml mosquito diluent [MD; 20% heat-inactivated fetal bovine serum (FBS) in Dulbecco’s phosphate-buffered saline (PBS) plus 50 μg/ml penicillin/streptomycin, 50 μg/ml gentamicin, and 2.5 μg/ml Fungizone]. Capillaries charged with FBS plus 50% sucrose (1:1) were used to collect salivary secretions for approximately 30 minutes followed by ejection into 0.3ml MD. Mosquitoes were then placed in individual tubes with 1.0 ml MD. All samples were held at -80°C until tested. Bodies and legs were processed separately as previously described and all samples were screened and titrated by plaque assay in duplicate on Vero cell culture. Infection rates were determined by the number of WNV positive mosquito bodies per fed mosquito, dissemination rates by the proportion of WNV positive
mosquitoes with WNV positive legs, and transmission rates by the proportion of WNV positive mosquitoes with WNV positive salivary secretions. Growth kinetics were determined using viral titers of WNV positive mosquitoes on individual days. Replicative fitness refers to the slope of the best-fit log-linear line of mosquito body titers as a function of time post-feeding.

*Mixed infections and sample collections*

Three separate experiments utilizing artificial swarms composed of WNV variants were performed in *Cx. pipiens*. In experiments 1 and 2, WN S5 (swarm of 5 variants) and WNV S8 (swarm of 8 variants), respectively, were diluted 1:5 in defibrinated blood plus sucrose and offered to ~1000 female *Cx. pipiens* mosquitoes as described above. Tissues samples were collected on days 5, 7, 10, 14, and 21 post-feeding from 12-20 mosquitoes. Collected samples included salivary secretions (SS), legs (LEG), midguts (MG), salivary glands (SG), and body remnants (REM). Dissected samples, i.e., MG, SG, and REM, were washed by passing twice through separate clean drops of MD prior to collection. All samples were collected in MD, stored and processed as described above, and subsequently tested for WNV by plaque screen on Vero cell culture. In experiment 3, five different bloodmeals were created using proportions of variants 2 and 4 including 1:1, 2:1, 4:1, 8:1 and 16:1 and separately offered to ~500 mosquitoes. In addition to above sample collections, 5 mosquitoes/proportion were saved for testing immediately after feeding. A modified experiment was performed in *Cx. quinquefasciatus* with WN S8. This study was identical to experiment 2, yet just 4-8 tissues/timepoint were harvested for analyses.
Molecular cloning and sequencing

In order to confirm proportion of variants in WNV positive bloodmeals and determine proportions in harvested tissues, molecular cloning and sequencing were performed on WNV positive samples. Generation and analysis of clones were performed using a modification of previous studies (Ciota et al., 2007; Ciota et al., 2009). RNA was extracted from infected specimens with QIAamp viral RNA extraction kit (Qiagen) and RT-PCR was conducted using primers designed to amplify the 3' 2201 nt of the WNV envelope (E) coding region and the 5' 3248 nt of the WNV non-structural protein 1 (NS1) coding region. RT was performed with Sensiscript RT (Qiagen) at 45 °C for 40 min followed by heat inactivation at 95 °C for 5 min. The resulting cDNA was used as a template for PCR amplification. WNV cDNA was then amplified with Taq polymerase (New England Biolabs, Ipswich, MA) according to the manufacturer's specifications. PCR products were visualized on a 1.5 % agarose gel and DNA was recovered by using a MinElute Gel Extraction kit (Qiagen) as specified by the manufacturer. The recovered DNA was ligated into a TOPO-TA cloning vector (Invitrogen) and transformed into One Shot TOP10 Electro-competent E.coli cells according to the manufacturer's protocol. Kanamycin resistant colonies were screened by direct PCR using primers specific for the desired insert and plasmid DNA was purified using a QIaprep Spin Miniprep kit (Qiagen) as specified by the manufacturer. Sequencing was performed at the Wadsworth Center Applied Genomics Technology Core using ABI 3700 and 3100 automated sequencers (Applied Biosystems). Seventeen to twenty-two clones were sequenced per sample.

Sequence and data analyses
WNV sequences were compiled, edited, and aligned using the SeqMan module of the DNASTAR software package and a minimum of two-fold redundancy throughout each clone was required for individual sequence data. As used in previous studies to quantify mutant swarm breadth, normalized Shannon entropy ($S_n$) was calculated based on the frequency of genotypes in populations as follows: $S_n = \sum -i \ln Pi / \ln N$, where $Pi =$ frequency of individual genotype and $N =$ number of clones sequenced (Ciota et al., 2007). $S_n$ values can range from 0 (completely homogeneous) to 1 (completely heterogeneous) yet in these experiments only input haplotypes were considered and, therefore, heterogenetity could not exceed the input value in any given tissue. Statistical analyses were performed using both Microsoft Excel 2003 and GraphPad Prism version 4.00.

**Results**

*Characterization of WNV variants in Cx. pipiens mosquitoes*

The goal of these studies was to create a WNV swarm composed of variants of neutral fitness which could be tracked over time and space *in vivo* in Cx. pipiens mosquitoes. For this reason, all variants were individually characterized *in vivo* to confirm that no fitness changes existed relative to the FL-WNV strain (variant 1) from which they were derived. This characterization occurred over multiple experiments and included quantification of vector competence (infection, dissemination, and transmission rates), WNV titers on individual days and overall growth rates (replicative fitness, table 7). The mean replicative fitness among variants was $0.15 \log_{10}$ pfu/mosquito/day for days 5-14 post-feeding and no statistical difference occurred among variants (linear regression analysis of log-linear slopes, $p=0.18$). Generational fluctuations in mosquito colonies are
not uncommon and, in some cases, differences between individual experiments were measured in terms of both vector competence and WNV titers in mosquitoes. In all experiments, variant 1 was used as an internal control to correct for this variation. No significant differences in WNV titers on individual days were detected (t-tests, p>0.05), nor were significant differences in vector competence detected among variants used (Chi-squared, p>0.05; data not shown) within individual experiments.

In addition to confirmation of neutral fitness among variants, genetic stability was assessed by sequencing of the NS1 gene from WNV positive legs and/or salivary secretions from individual mosquitoes at days 14-21 post-feeding. All variants used for mixed feedings (variants 1-8) were confirmed to be genetically stable. A total of 3 NS1 variants created were disqualified from this study due to identification of fitness changes and/or lack of genetic stability.

_Tissue-specific diversity in Cx. pipiens following feeding on artificial mutant swarms_

WN S5 was used for the first experimental mixed feeding. Molecular cloning and sequencing confirmed that variants 1-5 were present at relatively high proportions ranging from 0.15 to 0.3 (table 7). The WNV input titer for this feeding was 8.1 log_{10} pfu/ml. Overall infection rates were relatively low, 45.0%, with 27 of 60 mosquito midguts WNV positive. Of these, 15 of 27 (56.0%) had disseminated infections (WNV positive legs and/or remnants) and 6 of 27 (22.0%) had infected salivary glands. Three of 6 mosquitoes with WNV positive salivary glands also had virus in their salivary secretions, yet no salivary secretions were successfully cloned in experiment 1. WNV RNA isolated from a total of 21 tissues, including 13 midguts, 6 leg samples and 6
salivary gland samples, was cloned and sequenced from the WN S5 feeding. There were a minimum of 2 WNV positive tissues of each type from each timepoint, with the exception of day 5 post-feeding when no WNV positive legs or salivary glands were identified. Comparison of $S_n$ values for the WN S5 bloodmeal to days 5, 7, and 10 midguts demonstrated that no substantial decrease in swarm breadth occurred with midgut infection and replication (t-test, p>0.05; Fig. 4). Despite this, 3 of the 7 midgut samples from days 5, 7 and 10 had lost 1 or more variants originally present in the bloodmeal, demonstrating a modest narrowing of the swarm at the population level. In addition, no significant difference in mutant swarm breadth was measured between tissue types at any individual timepoint, despite a downward trend in $S_n$ from midguts to legs to salivary glands on day 10 (Fig. 4; t-test, p>0.05). Although no tissue-specific trend was measured in experiment 1, a distinct temporal trend was apparent, with a significant decrease in swarm breadth occurring from days 7-21 (Fig. 4; linear regression analysis, $r^2=0.74$, p<0.001; Pearson correlation, $r=-0.92$, p=0.04). The mean number of variants identified in day 21 tissues was 1.33 and 4 of 6 tissues contained only a single variant. No consistent trend was observed in terms of the variant which ultimately dominated, with all but variant 3 represented as a master sequence in one of the 6 day 21 tissues. This result supported the conclusion of neutral fitness among variants.

WN S8 was used for the second experimental mixed feeding. The purpose of using this strain in the follow-up experiment was (i) to more accurately represent proportions one would expect to find in a natural strain and (ii) to characterize the effects of intrahost bottlenecks on minority variants. Unlike the WN S5 population in which variants existed in roughly equivalent proportions and had no variant below 15% in the population,
individual variants in the WN S8 population ranged from 52.5% (variant 1) to <2.5% (variants 6 and 7; table 7). The bloodmeal titer for this feeding was $8.7 \log_{10}$ pfu/ml. Twenty-one of 41 midguts tested (51.2%) were WNV positive. Of these, 14 (67.0%) had disseminated infections, 6 (29.0%) had WNV positive salivary glands, and 3 (14.0%) were capable of virus transmission. To attain more WNV positive salivary secretions in this experiment, secretions were collected from additional experimental but undissected mosquitoes at each timepoint. This provided an additional 4 WNV positive secretion samples. A total of 37 samples were successfully cloned and sequenced following the WN S8 feeding, including 13 MG, 12 LG/RM, 6 SG and 6 SS. There was some concern with the combining of legs and remnant tissues to represent disseminated virus in this experiment given the larger diversity of tissues infected in mosquito remnants, yet successful cloning of these tissues from the same mosquitoes demonstrated similar levels of swarm breadth. Unlike experiment 1, significant tissue-specific differences were identified following feeding on WN S8 (Fig. 5). Specifically, when tissues from all timepoints were combined for analyses, a significant decrease in mutant swarm breadth ($S_n$) was measured with WNV infection of the midgut (t-test, $p<0.001$) and the combination of WNV infection and release from the salivary glands (Fig. 5a, SS relative to MG or LEG/RM, t-test, $p<0.05$). Narrowing of the swarm was confirmed by quantifying the proportion of the population composed of minority sequences (Fig. 5b). Specifically, a decreasing proportion of minority sequences were identified with each successive tissue from BM to SS and individual differences were significant from BM to MG and SG to SS (chi-squared, $p<0.05$, Fig. 5b). To assess if these decreases in swarm breadth in MG and SS were due to tissue-specific bottlenecks rather than temporal
changes (i.e. downstream tissues are on average infected later), breadth in these tissues was analyzed separately on days 7, 14, and 21 (Fig. 6) and additional two-way ANOVA analyses were performed to separate the effects of tissue and time (table 8). Although statistical significance could not be achieved with data from individual days due to low sample sizes (2-4 tissues/timepoint), these results clearly demonstrate consistently narrower swarm breadth in MG samples relative to input and in SS samples relative to MG without significant temporal changes. In addition, results from ANOVA analyses demonstrate that tissue-specific differences in swarm breadth did occur independent of time both when all tissues are analyzed and when BM, MG, and SS data are considered independently (two-way ANOVA, p<0.05; table 8). Despite this, a temporal decrease in mutant swarm breadth was observed in tissues responsible for secondary virus amplification (LG, RM, SG) following feeding on WNS8 (Fig. 7; linear regression analysis, $r^2 = 0.18$, p=0.046; Pearson correlation, r=-0.96, p=0.04; two-way ANOVA, p<0.05; table 8). Viral titers in individual tissues were highly variable, ranging from 1.0 to 6.5 log$_{10}$ pfu WNV/tissue (Fig. 8). The relationship between viral titers and swarm breadth was evaluated in order to test for true correlation and/or cloning bias. WNV titers did not correlate with swarm breadth (Fig. 8; Pearson correlation, r=0.0003). In fact, both the highest and lowest titer samples (day 21 SG and day 11 LEG, respectively) were completely homogeneous and the most diverse sample (day 14 REM) was among the lowest titer samples.

In order to better determine the threshold input proportion for infection, experiment 3 evaluated day 5 midgut-derived WNV following feeding on five proportions of variant 2 and variant 4 ranging from 1:1 to 32:1. In addition, five time 0 samples, taken
immediately after bloodfeeding, were successfully cloned and sequenced, confirming that starting proportions were maintained with ingestion and that proportions of minority variants were above the threshold for detection (data not shown). Bloodmeal titers for these feedings were comparable for each group, ranging from 8.2 to 8.4 log_{10} pfu/ml blood. Analysis of five WNV MG isolates/group demonstrated that proportions of variants on the population level are generally retained following infection and early replication (Fig. 9A). Despite this, the minority variant (variant 4) was not detected within the swarm in any of the five WNV MG isolates from group 5 (32:1), suggesting this may be a threshold at which infection is highly unlikely. Similar to WN S5 and S8 feedings, variation was seen on the individual level and the proportion of individual midgut samples which had completely lost variant 4 increased with a decreasing input BM proportion (Fig. 9B).

In addition to the midgut derived samples sequenced from experiment 3, at least 1 salivary secretion isolate from each group was successfully cloned and sequenced. These results were combined with results from experiments 1 and 2 in order to generate the probabilities of infection, transmission, and dominance given various starting proportions in the infectious bloodmeal (Fig. 10). Results for infection depict a saturation curve in which the increases in the probability of infection generally lessen with increasing proportions and the probability of infection is high (greater than 80.0%) when the variant proportion is greater than ~ 25.0% (Fig. 9A). Results for transmission were more variable, potentially due to the effects of smaller sample sizes. Despite this, a similar association was seen with the probability of transmission increasing with increased input proportions, yet with smaller effects below starting proportions of 0.25 and larger effects...
above that threshold relative to infection (Fig. 10A). Results for dominance demonstrate that the probability of becoming the master sequence is directly correlated with a variant’s starting proportion and that, even with neutral variants, no low-end threshold was identified that would disqualify rare mutants from dominating in some mosquitoes (Fig. 10B). For example, despite being at a proportion below the level of detection (2.5%) in the WN S8 bloodmeal, variant 6 was found to dominate in two tissues tested from different mosquitoes.

For modified *Cx. quinquefasciatus* experiments with WN S8, just 2-4 samples/tissue type were analyzed. These numbers did not allow adequate assessment of specific temporal and spatial variation, but did suggest that, unlike *Cx. pipiens*, swarm heterogeneity is generally retained (Fig. 11). Specifically, proportions of minority sequences in infected *Cx. quinquefasciatus* midguts and in transmitted virus did not differ significantly from the input virus (Chi-squared, p>0.05), and both had significantly greater proportions of minority variants then was measured in *Cx. pipiens* (Chi-squared, p<0.05).

**Discussion**

Stochastic events, such as intrahost bottlenecks and nonselective sweeps imposed on arboviruses in mosquito vectors could have significant effects on viral fitness and plasticity. Here, we characterized how these events shape the *West Nile virus* mutant swarm in *Cx. pipiens* mosquitoes. Specifically, using artificial mutant swarms composed of neutral fitness variants with stable mutations, we tracked swarm breadth of input
variants over known barriers within the mosquito including: (i) midgut infection, (ii) midgut escape, (iii) salivary gland infection and (iv) virus transmission; and measured changes in breadth in individual tissues over time. Our results demonstrate that both time and space can contribute to significant narrowing of the WNV swarm, yet the extent of this narrowing is highly variable and largely dependent on variant proportions in infectious bloodmeals.

Although some narrowing of the WNV swarm was seen with midgut infection, no significant tissue-specific bottlenecking was measured following feeding on WN S5, demonstrating that most WNV variants which exist at relatively high proportions (≥ 15.0%, table 7) are likely to survive the anatomical barriers within Cx. pipiens. Despite this, significant decreases in swarm breadth were measured from day 7 to 21 post-feeding (Fig. 4). The mean number of variants in all tissues 21 days after feeding on WN S5 was 1.2. This sweep was stochastic in that the variant which would ultimately dominate could not be predicted, yet selection could still play a role here if one variant were to acquire a beneficial mutation over the course of mosquito infection. Alternatively, this could be a neutral event in which the innate immune response of the mosquito overcomes infection in an increasing number of cells with time, leaving a relatively small number of cells producing virus and therefore narrowing the swarm indiscriminately. Future sequencing and phenotypic characterization of output viruses will help to clarify the mechanism at work in these temporal sweeps. A similar time-dependent decrease was seen following feeding on the WN S8 strain, yet in this case the decrease was due exclusively to temporal changes in tissues responsible for secondary virus amplification (LEG, REM, SG; Fig. 7; table 8). These results are comparable to a recent study measuring intrahost
WNV nucleotide diversity in Cx. *quinquefasciatus* mosquitoes in which a general trend of decreasing nucleotide diversity was also measured from day 7 to 21 post-feeding (Brackney et al., 2011).

In contrast to these results, here we also identified significant tissue-specific bottlenecking following feeding on the WN S8 population (Figs. 5 and 6; table 8). Specifically, a significant decrease in swarm breadth was seen with midgut infection and virus transmission and, although lacking in statistical significance, a narrowing of the swarm was also seen with salivary gland infection. Although *Cx. pipiens* and *Cx. quinquefasciatus* are sibling species which readily hybridize (Huang et al., 2011), significant variation in vector competence has been measured with these and other closely related *Culex* species (Reisen et al., 2008; Moudy et al. 2011). Such variation could explain differences in the extent of spatial bottlenecking. Additionally, what is measured in these two studies is also quite different. In the Brackney et al study (2011), overall nucleotide diversity, both that which is produced and is maintained, is measured following infection with an attenuated, highly diverse strain composed of multiple variants with potentially altered phenotypes (Fitzpatrick et al., 2010). Here, we’ve tracked proportions of marked input strains confirmed to have equivalent fitness. Additionally, in terms of sequence diversity, the WN S8 strain is a valid representative of what a mosquito might encounter in nature, i.e. a swarm with one dominant genotype surrounded by minority genotypes in variable proportions. This difference in methodology between these two studies could also explain differences observed in the relationship between viral titer and swarm diversity, yet modified experiments performed here with WN S8 in *Cx. quinquefasciatus* do in fact support the idea that significant
species-specific differences exist (Fig. 11). As was shown by Brackney et al, we would anticipate that those tissues in which the level of replication was higher would also experience more mutation and higher levels of genetic diversity, yet since what we have tracked in this study are only the mutations of the variants initially present in the bloodmeal, we wouldn’t necessarily expect a similar relationship with our measure of swarm breadth. Instead, the lack of a correlation between swarm breadth and WNV titer that we measured demonstrated that there was no bias for heterogeneity with high titer samples caused by our methodology (or vice versa), and that the capacity for individual tissues to produce virus had no influence on quantifying the size of the preceding bottleneck (Fig. 8). The identification of tissue-specific bottlenecks following feeding on WN S8 also contrasted with the WN S5 feeding, demonstrating, not surprisingly, that the presence of more minority variants in input strains increased the likelihood of such variants being lost throughout infection.

The most significant bottleneck identified was at the level of midgut infection (Figs. 5 and 6). The host and virus-derived factors governing WNV infection of midgut epithelial cells are not well defined, yet a previous study utilizing WNV virus-like particles in Cx. quinquefasciatus demonstrated that a relatively small number of cells \((\leq 15)\) are initially infected in the posterior midgut (Scholle et al., 2004). Studies with epizootic Venezuelan equine encephalitis virus also demonstrated a low proportion of midgut cells infected in Aedes taeniorhynchus (mean of 28), yet also demonstrated that levels of susceptibility are likely both virus and vector-specific (Smith et al., 2008). Although the extent of co-infection of midgut cells is not known for WNV or any flavivirus, the availability of susceptible cells is likely to be the limiting factor determining the degree of genetic
bottlenecking occurring with initial infection. Feeding on different proportions of variants 2 and 4 demonstrated that, although proportions of neutral variants in infectious bloodmeals are generally retained on the population level (Fig. 9A), even a variant comprising 50% of the viral swarm can often be lost with initial infection (Fig. 9B). In addition, these studies demonstrate that minority variants existing at proportions less than 3% rarely infect *Cx. pipiens* midguts, even at relatively high input titers. Despite this, results from the WN S8 feeding demonstrate that variants which were extremely rare in the input virus can at times persist and even dominate so input proportions are likely to be highly predictive of proportions on the population level (Fig. 10).

Following midgut infection and replication, arboviruses need to disseminate into the hemocoel to infect secondary tissues. This barrier has historically been viewed as among the most important in determining vector competence (Hardy et al., 1983). Surprisingly, neither artificial swarm used in this study identified a significant midgut escape bottleneck for mosquitoes developing disseminated infections, as the breadth of the swarm in both midguts and parenteral tissues (LEG/REM) was similar (Figs. 4 and 5). It is clear that such a dissemination barrier exists, as many infected mosquitoes do not develop disseminated infections, yet what this result suggests is an all or nothing phenomenon rather than a case of a subset of the population reaching these secondary tissues. This implies that it is more of an immunological rather than anatomical barrier in which all variants in a neutral viral swarm that successfully infect the midgut cells and replicate are equally capable of midgut egress, yet does not rule out the possibility of strong selective pressure for individual variants in phenotypically diverse swarms.
A further narrowing of the swarm was observed with salivary gland infection and a significant decrease in swarm breadth was measured with virus transmission, supporting the idea that the infection of and egress from the salivary glands also acts as a significant anatomical barrier to WNV infection [(Chamberlain and Sudia, 1961);(Fig. 5)]. Although salivary secretion titers were generally among the lowest of tested samples, one of the six salivary secretion samples tested from the WN S8 feeding had a viral titer of $\sim 5.0 \log_{10} \text{pfu/ml}$ yet was completely genetically homogeneous. This suggests again that the measured swarm narrowing is a real phenomenon rather than the consequence of a methodology imposed bottleneck resulting from using \textit{in vitro} transmission assays. What remains unclear is if all variants infecting the salivary glands are capable of entering the saliva and the bottleneck is during transmission or, if some of the variants getting into the glands never get out. Since all mosquitoes were dissected for these experiments, salivary secretions from individual mosquitoes could not be monitored through time, but such experiments would help to clarify this question.

Although trends are apparent with our data set, there is a large amount of variability among individuals with regard to mutant swarm breadth. For example, some midgut samples tested from early in infection contained just a single variant and some salivary gland samples contained as many as 5 variants. What we can conclude from this is that there is not a level of intrahost bottleneck that is necessarily inevitable, but that the probability of reaching each stage of infection and, ultimately, transmission certainly decreases with decreasing starting proportions (Fig. 10). Despite this, the fact that rare mutants, even with neutral fitness can occasionally dominate, stresses the need to identify minority variants in natural isolates. Incorporating selection coefficients in our
predictions could help us understand the likelihood of strains with variable fitness levels surviving to transmission.

One caveat of this work is that all experimental feedings were performed with high titers of WNV in infectious bloodmeals. Although these titers are well within the range of viremia levels found in many naturally infected birds, particularly Passeriformes (Komar et al., 2003), Culex mosquitoes are capable of acquiring infections from lower titer bloodmeals (Anderson et al., 2010). Such infections would most likely increase the probability of tissue-specific bottlenecks; yet this may not be the case if only a subset of midgut cells are susceptible and consequently, the same limited number of cells initiate mosquito infection. Future studies infecting mosquitoes at variable doses will help to clarify this. In addition, both feeding on artificial bloodmeals and in vitro transmission assays, although useful for accurate quantification of input and output swarm breadth, may not be perfect representations of natural infection and transmission. Lastly, as suggested by preliminary results in Cx. quinquefasciatus (Fig. 11) both species- and population-specific differences together with strain variation could substantially alter the dynamics of WNV infection in Culex mosquitoes. Despite the need to clarify such variables, these results demonstrate the profound effects that stochastic forces across time and space in mosquito vectors can have in shaping arboviral mutant swarms.
Acknowledgements. Dylan J. Ehrbar, Graham G. Willsey, Greta A. Van Slyke, Anne F. Payne, Rachael E. Viscio, and Laura D. Kramer are also authors on the originally published manuscript and all contributed to these studies. The authors would like to thank Pei-Yong Shi for providing the FL-WNV infectious clone for these studies. We also thank Pamela Chin for mosquito rearing and maintenance and Jean Demarco for assistance with molecular cloning. Cells and media for these studies were provided by the Wadsworth Center Tissue Culture and Media facility and sequencing was completed by the Wadsworth Center Applied Genomics Technology Core. This work was supported by federal funds from the National Institute of Health (grant number RO1-AI-077669). The construction of the Wadsworth Center Insectary Facility was funded by NIH grant number C06-RR-17715.
Table 7. Replicative fitness of individual WNV NS1 variants and proportions in artificial mutant swarms.

<table>
<thead>
<tr>
<th>WNV strain</th>
<th>mutation</th>
<th>replicative fitness$^1$</th>
<th>WN S5 proportion$^2$</th>
<th>WN S8 proportion$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WN variant 1</td>
<td>FL-WNV</td>
<td>0.15</td>
<td>0.30</td>
<td>0.525</td>
</tr>
<tr>
<td>WN variant 2</td>
<td>T2700C</td>
<td>0.17</td>
<td>0.15</td>
<td>0.075</td>
</tr>
<tr>
<td>WN variant 3</td>
<td>T2913C</td>
<td>0.19</td>
<td>0.15</td>
<td>0.150</td>
</tr>
<tr>
<td>WN variant 4</td>
<td>T2940C</td>
<td>0.17</td>
<td>0.20</td>
<td>0.200</td>
</tr>
<tr>
<td>WN variant 5</td>
<td>A2844G</td>
<td>0.18</td>
<td>0.20</td>
<td>0.025</td>
</tr>
<tr>
<td>WN variant 6</td>
<td>T3117C</td>
<td>0.11</td>
<td>-</td>
<td>&lt;0.025*</td>
</tr>
<tr>
<td>WN variant 7</td>
<td>A2847T</td>
<td>0.11</td>
<td>-</td>
<td>&lt;0.025*</td>
</tr>
<tr>
<td>WN variant 8</td>
<td>T2910C</td>
<td>0.14</td>
<td>-</td>
<td>0.025</td>
</tr>
</tbody>
</table>

$^1$ Replicative fitness refers to the slope of the best fit line generated by plotting geometric means of WNV titers in 30-45 *Cx. pipiens* bodies per variant on days 5, 7, 10, and 14 post feeding on ~8.2 log$_{10}$ PFU virus.

$^2$ Proportions were determined by molecular cloning and sequencing of 40 clones generated from input virus.

* Variants were included in the original mixture but not identified in the input virus. Both variants were identified in individual tissues following feeding.
Table 8. Results from two-way ANOVA analyses evaluating the effects of tissue and time on WNV mutant swarm breadth ($S_n$) following *Cx. pipiens* feeding on WN S8.

<table>
<thead>
<tr>
<th>Group</th>
<th>Source of variation</th>
<th>df</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>all tissues(^1)</td>
<td>time</td>
<td>3.0</td>
<td>0.047</td>
<td>0.016</td>
<td>0.94</td>
<td>0.434</td>
</tr>
<tr>
<td></td>
<td>tissue</td>
<td>4.0</td>
<td>0.328</td>
<td>0.082</td>
<td>4.93</td>
<td>0.004</td>
</tr>
<tr>
<td></td>
<td>interaction</td>
<td>12.0</td>
<td>0.174</td>
<td>0.014</td>
<td>0.87</td>
<td>0.587</td>
</tr>
<tr>
<td>LG/RM, SG</td>
<td>time</td>
<td>3.0</td>
<td>0.199</td>
<td>0.067</td>
<td>3.43</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>tissue</td>
<td>1.0</td>
<td>0.035</td>
<td>0.035</td>
<td>1.79</td>
<td>0.202</td>
</tr>
<tr>
<td></td>
<td>interaction</td>
<td>3.0</td>
<td>0.031</td>
<td>0.010</td>
<td>0.53</td>
<td>0.667</td>
</tr>
<tr>
<td>BM, MG, SS(^2)</td>
<td>time</td>
<td>3.0</td>
<td>0.009</td>
<td>0.003</td>
<td>0.24</td>
<td>0.870</td>
</tr>
<tr>
<td></td>
<td>tissue</td>
<td>2.0</td>
<td>0.282</td>
<td>0.141</td>
<td>10.36</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td>interaction</td>
<td>6.0</td>
<td>0.016</td>
<td>0.014</td>
<td>0.19</td>
<td>0.972</td>
</tr>
</tbody>
</table>

\(^1\) Significant tissue-specific variation in mutant swarm breadth was measured independent of time.

\(^2\) Time-dependent variation in mutant swarm breadth independent of tissue type is measured when legs/remnants and salivary glands are analyzed separately.

\(^3\) Tissue-dependent variation in mutant swarm breadth independent of time is measured when bloodmeal, midguts, and salivary secretions are analyzed separately.
Figure 4. Changes in mutant swarm breadth ($S_n$) over time and space following bloodfeeding on WN S5. Each bar represents the mean of 2-3 tissues +/- SEM. No significant differences were measured between tissues on individual days (t-test, p>0.05). A significant negative correlation between time and $S_n$ was measured for all tissues from days 7 through 21 (linear regression analysis, $r^2=0.74$, p<0.001; Pearson correlation, r=-0.92, p=0.04).
Figure 5. Tissue-specific differences in mutant swarm breadth in *Cx. pipiens* following feeding on WN S8. MG=midguts (n=13), LG/RM=legs or remnants (n=12), SG=salivary glands (n=6), and SS=salivary secretions (n=6). (A) Mean tissue-specific entropy (S_n +/- SEM). Significant decreases in swarm breadth (*) were measured with WN S8 bloodmeal (BM) relative to MG (t-test, p<0.001) and with MG relative to SS (t-test, p=0.03). (B) Proportion of all sequences from bloodmeal and individual tissues identified as minority variants. Significant differences identified by Chi-squared analysis (*) were measured between BM and MG (X^2 = 11.34, p=0.02) and SG and SS (X^2 = 10.19, p=0.03).

Figure 6. Mutant swarm breadth in *Cx. pipiens* midguts and salivary secretions on individual days following feeding on WN S8. Bars represent mean tissue-specific entropy (S_n +/- SEM, n=2-4/timepoint) of midgut (MG) and salivary secretions (SS).
Figure 7. Temporal change in WNV mutant swarm breadth ($S_n$) in secondary tissues following bloodfeeding on WN S8. Data points represent mean $S_n$ of 3-7 LG (leg), RM (remnant), and SG (salivary gland) samples +/- SD at 7, 11, 14, and 21 days post feeding. A significant negative correlation was measured between $S_n$ and time (linear regression analysis, $r^2 = 0.18$, $p=0.046$; Pearson correlation, $r=-0.96$, $p=0.04$).
Figure 8. Mutant swarm breadth ($S_n$) and viral titers of individual *Cx. pipiens* tissues at days 5-21 following feeding on WN S8. No correlation between viral titer and $S_n$ was measured for individual tissue types or combined data (Pearson correlation, Pearson $r = 0.0003$).
Figure 9. Proportions of two WNV variants in *Culex pipiens* midguts (MG) following feeding on various proportions in bloodmeals (BM). A. Proportions of variant 2 and 4 clones identified from five day 5 MG samples (n=40 clones sequenced/BM and 125-200 clones sequenced/5 MG). No statistically significant differences in proportions were identified between corresponding BM and MG groups (chi-squared, p>0.05), yet no variant 4 clones were identified in group 5 midguts (MG5). B. Proportion of MG samples (n=5) without variant 4 (minority sequence).
Figure 10. Predictions from combined experiments. A. The probability of infection and transmission of WNV in *Cx. pipiens* as a function of WNV variant proportion in the input mutant swarm. The probability of infection was defined as the proportion of midguts in which a variant was identified given multiple starting proportions and the probability of transmission refers to the (probability of infection)(the proportion of individuals transmitting a variant) at each starting proportion. Lines represent best-fit non-linear relationships of association ($R^2=0.87$ and 0.83 for infection and transmission, respectively). B. The likelihood of a particular WNV variant becoming the master (dominant) sequence in all *Cx. pipiens* tissues is directly related to its starting proportion in the infectious bloodmeal (linear regression analysis, slope=1.01, $R^2=0.94$).
Figure 11. Tissue-specific differences in mutant swarm breadth in *Cx. quinquefasciatus* relative to *Cx. pipiens* following feeding on WN S8. Proportion of all sequences from bloodmeal and individual tissues identified as minority variants. Significant differences identified by Chi-squared analysis (*) were measured between species for MG and SS samples (p<0.05).
References


CHAPTER 4: COOPERATIVE INTERACTIONS IN THE WEST NILE VIRUS MUTANT SWARM*

* this material has appeared previously in Ciotla et al. 2012, BMC Evol. Biol. May 22;12:58

Abstract

The rapid replication rates, inherently high error rates, and large population sizes of RNA viruses often result in highly genetically diverse mutant swarms within individual hosts. A more complete understanding of the phenotypic consequences of these diverse swarms is needed in order to equate RNA swarm breadth and composition to specific adaptive and evolutionary outcomes. Here, we determined clonal fitness landscapes of mosquito cell adapted West Nile virus (WNV) and assessed how altering the capacity for cooperative interactions among variants affects mutant swarm dynamics. Our results demonstrate that although there is significant mutational robustness in the WNV swarm, genetic diversity corresponds to substantial phenotypic diversity in terms of relative fitness in vitro. In addition, our data demonstrate that increasing levels of co-infection can lead to widespread strain complementation, which acts to maintain high levels of phenotypic and genetic diversity and potentially slow selection for individual variants. Lastly, we show that these cooperative interactions may lead to swarm fitness levels which exceed the relative fitness levels of any individual genotype. Overall, these studies demonstrate the profound effects cooperative interactions can have on arbovirus evolution and adaptation, and provide a baseline by which to study the impact of this phenomenon in natural systems.
Introduction

The rapid replication rates, inherently high error rates, and large population sizes of RNA viruses often result in highly genetically diverse mutant swarms within individual hosts. A number of studies with arthropod-borne viruses (arboviruses) have demonstrated that the precise composition and breadth of these swarms may be directly responsible for alterations in viral fitness, pathogenesis, and host breadth (Ciota et al., 2010). Taken together, these studies demonstrate that a more comprehensive understanding of mutant swarm dynamics is required for adequate characterization of arbovirus evolution and adaptation. If the maintenance of highly diverse swarms equates to increased phenotypic plasticity, then swarm size could be particularly important for arboviruses given their requirement for replication in disparate hosts and tissues. The fact that arboviruses almost exclusively possess RNA genomes supports the idea that such genomic flexibility may provide a considerable advantage for host cycling. Not surprisingly, many studies have identified significant intrahost diversity of arboviruses both in laboratory-derived and naturally occurring strains (Chen et al., 2003; Jerzak et al., 2005; Jerzak et al., 2007; Wang et al., 2002), yet the phenotypic diversity that correlates to these mutant genotypes has not been adequately characterized. Quasispecies theory, although often invoked as a synonym for genetic diversity, argues against the maintenance of substantial phenotypic diversity, as the inevitable coupling of mutational neighbors should ultimately favor selection for canalized populations which exhibit phenotypic stability in the face of mutational change (Eigen & Biebricher, 1988; van Nimwegen et al., 1999; Wilke et al., 2001). As such, a robust viral swarm would
theoretically lack phenotypic plasticity and therefore tend to negate the adaptive advantage one would normally associate with genetic diversity. A virus could overcome this need for robustness by readily utilizing cooperative interactions via strain complementation in co-infected cells. Although the capacity for such interactions has been demonstrated by previous studies with *Vesicular stomatitis virus* (VSV; [Novella et al., 2004]), consideration of the importance of this phenomenon in the evolution of other medically important arboviruses is lacking. Such interactions could permit arboviruses to maintain genotypes which are host specialized, therefore diminishing the potential for compromises in host adaptation as a result of host cycling. Indeed, a number of studies have now clearly demonstrated that although host-specific adaptation may at times result in a fitness cost, adaptive trade-offs in individual hosts are not inevitable for arboviruses (Ciota et al., 2007b; Ciota et al., 2008; Ciota et al., 2009; Cooper & Scott, 2001; Greene et al., 2005; Remold et al., 2008; Weaver et al., 1999; Zarate & Novella, 2004).

A previous experimental evolution study conducted in our laboratory generated a strain of *West Nile virus* by 40 sequential passages (WNV CP40) which was both highly adapted to mosquito cells and highly genetically diverse. The size of the WNV mutant swarm was shown to be directly correlated with virus adaptation, and reverse genetics studies demonstrated that consensus mutations were not solely responsible for the adaptive phenotype, indicating the importance of mutant swarm variants when considering host adaptation and overall viral fitness (Ciota et al., 2007b; Ciota et al., 2007c). In the current studies, we sought to evaluate the level of robustness in this strain by quantifying the phenotypic diversity and relative fitness values of biological clones derived from this WNV population. This strain provided a unique opportunity to
characterize non-consensus variants of an adapted and genetically diverse arbovirus which had sufficient time to reach an equilibrium state. In addition, we tested the hypothesis that strain complementation plays a significant role in the maintenance of both genetic and phenotypic diversity in the WNV mutant swarm. Lastly, we performed preliminary evaluation of the capacity for strain complementation in vivo in Culex tarsalis mosquitoes.

Material and Methods

Virus strains and isolation of biological clones.

The biological clone of WNV used as the parental strain for passage studies was isolated from WNV NY003356 by three rounds of plaque purification on Vero cells (ATCC #CCL-81) and WNV CP40 was obtained by 40 sequential passages on C6/36 mosquito cells (ATCC #CRL-1660) as previously described (Ciota et al., 2007b). Both the biological clone and strains derived from the initial passage on mosquito cell culture were previously found to be highly genetically homogeneous (Fig. 12; [Ciota et. al., 2007b]). The WNV monoclonal antibody resistant mutant (WNV MARM) was isolated in the presence of WNV MAb 5H10 (BioReliance Invitrogen Bioservices #81-003) as previously described, and retained both equivalent growth kinetics and fitness relative to both the unpassed parental strain and the initial mosquito cell passage (Ciota et al., 2007b). All plaque titrations were completed in duplicate on Vero cells as described elsewhere (Payne et al., 2006). Biological clones for phenotypic characterization were
isolated at random from two 6-well plates following plaque titration of WNV strains to 10-20 pfu/well. Forty single plaques per strain were re-suspended in 100ul BA-1 and immediately inoculated onto confluent monolayers of C6/36 cells for amplification. Following 72 hours of growth, cloned strains were harvested and stored at -80°C for subsequent titration and competition assays.

Infectivity

In order to assess the level of co-infection occurring in C6/36 cells infected with WNV CP40, a combination of quantitative RT-PCR and flow cytometry was used. C6/36 cells were infected with WNV CP40 at an MOI of 10 on 6-well plates for 12 hours. WNV RNA particles were quantified before and after incubation using TaqMan qRT-PCR (Applied Biosystems). Specifically, RNA was extracted using the QIAamp viral RNA extraction kit (Qiagen) from both initial inoculum (time 0) and incubated inoculum following washing (1 hour post infection). The difference in these measures was used to estimate the number of WNV particles entering the cells. Cells were then trypsinized, placed into a flow tube to the concentration of 1x10⁶, and fixed with 4% paraformaldehyde (PFA) at 4°C for 20 minutes. PFA was removed by centrifugation and then re-suspended in 500 ul of saponin and again incubated at 4°C for 20 minutes. After centrifugation to remove saponin, WNV mouse hyperimmune ascites fluid (CDC) polyclonal antibodies (1:100) was added and incubated in the dark at 4°C for 30 minutes. Following additional washes with saponin, goat anti-Fitc (KPL) IgG (1:50) was added and incubated for 30 minutes at 4C. After 3 washes with saponin, cells were resuspended
in 250 PBS +1% FBS and fixed overnight with 250ul 4% PFA. Cells were then analyzed on Becton Dickinson Facscan (BD bioscience).

In vitro competition assays

Relative fitness of WNV was evaluated by competition assays on C6/36 mosquito cells using a modification of previous studies (Holland et al., 1991). Briefly, confluent cell monolayers in six-well plates were infected in duplicate with a 1:1 mixture of control (WNV MARM) to test virus at the desired MOI (0.01 PFU/cell unless otherwise denoted), based on Vero cell titer. WNV MARM has been shown in multiple assays to have equivalent fitness to both the unpassed WNV biological clone and the initial passage in mosquito cell culture (Ciota et al., 2007b). After a 60-min absorption period at 28°C the infected monolayers were washed three times and overlaid with 3 ml of maintenance medium. Medium from the infected cultures was harvested at 24, 48, and 72 hours post infection, diluted 1:10 in growth medium supplemented with 20% FBS, and frozen at -80 °C for subsequent titration. The quantity of control and test virus was obtained by plaque titration in the presence (MARM control titer) or absence of Mab (total titer), and the reported relative fitness refers to the mean output ratio of test:control from duplicate assays.

Molecular cloning and sequencing
Production and analysis of clones was performed basically as previously described (Ciota et al., 2008). RNA was extracted from infected specimens with QIAamp viral RNA extraction kit (Qiagen) and RT-PCR was conducted using primers designed to amplify the 3' 1311 nt of the WNV envelope (E) coding region and the 5' 3248 nt of the WNV non-structural protein 1 (NS1) coding region. RT was performed with Sensiscript RT (Qiagen) at 45 °C for 40 min followed by heat inactivation at 95 °C for 5 min. The resulting cDNA was used as a template for PCR amplification. WNV cDNA was then amplified with a ‘high-fidelity’ protocol using PfuUltra (published error rate = 4.3 X 10^-7; Stratagene), according to the manufacturer's specifications. PCR products were visualized on a 1.5 % agarose gel and DNA was recovered by using a MinElute Gel Extraction kit (Qiagen) as specified by the manufacturer. The recovered DNA was ligated into the cloning vector pCR-Blunt II-TOPO (Invitrogen) and transformed into One Shot TOP10 Electro-competent E.coli cells according to the manufacturer's protocol. Kanamycin resistant colonies were screened by direct PCR using primers specific for the desired insert and plasmid DNA was purified using a QIAprep Spin Miniprep kit (Qiagen) as specified by the manufacturer. Sequencing was performed at the Wadsworth Center Applied Genomics Technology Core using ABI 3700 and 3100 automated sequencers (Applied Biosystems). Seventeen to twenty-two clones were sequenced per sample.

Sequence and data analysis

WNV sequences were compiled, edited, and aligned using the SeqMan module of the DNASTAR software package and a minimum of two-fold redundancy throughout each
clone was required for individual sequence data. Normalized Shannon entropy ($S_n$) was calculated based on frequency of genotypes in populations as follows: Shannon entropy ($S_n$) = $\sum P_i \ln P_i / \ln N$, where $P_i$ = frequency of individual genotype and $N$ = number of clones sequenced. $S_n$ values range from 0 (completely homogeneous) to 1 (completely heterogeneous). Mean hamming distance was calculated by averaging the number of base substitutions in individual clones relative to the consensus sequence. Tests for recombination including the RDP (Martin & Rybicki, 2000), GENECONV (Padidam et al., 1999), and MaxChi Smith 1992 (Smith, 1992) methods were performed using the Recombination Detection Program v.4.13 (Martin et al., 2010). Statistical analyses were performed using both Microsoft Excel 2003 and GraphPad Prism version 4.00.

Results

*Distribution of relative fitness values of clonal populations of mosquito cell-adapted WNV*

Previous studies have demonstrated that the accumulation of minority variants in the WNV mutant swarm is coupled with passage and adaptation to mosquito cell culture ([Ciota et al., 2007c]; Fig. 12). In order to assess the extent to which this genetic diversity corresponds to phenotypic diversity, 40 biological clones of WNV CP40 were isolated and the fitness of individual variants was quantified. All variants tested out-competed the WNV MARM strain, which was previously shown to have a fitness value equivalent to the parental strain from which WNV CP40 was derived (Ciota et al.,
Relative fitness values ranged from 3.2 to 45.9 (WNV CP40-12), with a mean value of 14.3 and a median value of 13.2 (Fig. 13). Despite this phenotypic variation, 50% of variants had relative fitness values between 10 and 20. Following mixing at equal proportions, the relative fitness value of the population was 18.6, which was higher than both the mean and median values and 72.5% of the individual variants.

**WNV infectivity in mosquito cells**

In order to assess if the level of co-infection was sufficient to allow the potential for interaction among viral genomes and/or proteins, qRT-PCR and flow cytometry were used to determine the mean number of viral particles per cell. Results demonstrated that following a 1-hour infection of mosquito cells with WNV CP40 at an MOI of 10, approximately $7.4 \log_{10}$ WNV RNA particles entered the cells, representing 29.2% of the virus inoculum ($8.0 \log_{10}$ particles/$7.3 \log_{10}$ pfu). Flow cytometry identified 68.2% of the cells in the monolayer as WNV positive, which is equivalent to approximately $6.1 \log_{10}$ mosquito cells. Taken together, these results demonstrate an approximate mean particle to cell ratio of 14.3:1, signifying widespread co-infection of individual mosquito cells at high MOIs (table 9).

**Relationship between levels of co-infection and relative fitness**

*In vitro* competition assays at MOIs 0.01, 0.1, 1.0, and 10 were performed to assess the relationship between co-infection and the capacity of high fitness WNV variants to out-
compete WNV MARM. In a population of variants which are independently competing, the highest fitness variant, in this case CP40-12 (circled value, Fig. 13), should experience a swift selective sweep. The goal of these studies was therefore to assess the role of cooperative interactions in inhibiting this selective sweep. The results demonstrate a significant negative correlation between MOI and relative fitness values, indicating that the capacity for the lower fitness variant to be retained in the population is increased with increased co-infection (Fig. 14a; linear regression analysis, \( r^2 = 0.68, \ p = 0.0035 \)). Conversely, this indicates that a strain (CP40-12) which in the absence of co-infection achieves titers over 30 times higher than its lower fitness competitor (WNV MARM) would be subject to an approximately 6-fold decrease in its capacity to out-compete under these conditions when levels of co-infection are high (MOI 10 relative fitness < 5, Fig. 14a). It should be noted that the output viral titers for these competition assays were similar at 72 hours, suggesting that the absolute fitness of the combined infection is likely comparable and that changes only reflect different proportions of the two strains. Measuring relative fitness at 72 hours post infection allowed such estimations of the overall replicative ability of the ‘population’, yet it is possible that differences in levels of replication could also be used to explain different relative fitness values (i.e. high MOI infections more quickly become saturated and may be less capable of out-competing). Thus, to confirm the role of co-infection in MOI-dependent fitness values, relative fitness values were also quantified at intermediate time points for the MOI 10 competition assay. Results indicate an inverse relationship between viral titers and relative fitness values, demonstrating replication and secondary infection cannot explain differences in relative fitness and providing further support for the negative
correlation between co-infection and the capacity for high fitness variants to out-compete (Fig. 14b; Pearson correlation, r=-0.88, p=0.0195).

Relationship between levels of co-infection and genetic diversity

To directly assess the role of co-infection in maintaining genetic diversity, mutant swarm breadth was quantified after 24 hours growth following infection at MOIs of 0.1, 1, and 10. Performing these experiments at 24 hours eliminated concerns about the role of secondary infection significantly altering the outcome. Although only approximately 20% of the genome was analyzed for these studies, results clearly demonstrate a significant positive correlation between MOI and multiple measures of genetic diversity (Fig. 15). Specifically, comparisons of normalized Shannon entropy on both the nucleotide and amino acid level significantly increased with increasing levels of co-infection (linear regression analysis, \( r^2 = 0.99 \) for both measures, \( p<0.05 \)). In addition, mean hamming distance of the MOI 10 mutant swarm (1.64) was significantly higher than both the MOI 1.0 (0.95) and MOI 0.01 (0.95) swarms (t-tests, \( p<0.05 \); Fig. 15), indicating that not only were more haplotypes present, but that individual variants were on average more different from the consensus sequence. If it is assumed that the region analyzed is a reliable representation of the level of diversity genome-wide, then the mean hamming distance for the full genome of individual variants is likely greater than 10 (WNV full genome ~ 11kb). Multiple sequence comparisons of variants from the MOI 10 group performed with the Recombination Detection Program v.4.13 (Martin et al., 2010) did not indicate that recombination contributed to the increased levels of genetic
diversity observed with increased co-infection. Although recombination cannot be completely ruled out without full genome comparisons, these results are consistent with previous studies demonstrating that recombination of WNV is extremely rare (Taucher et al., 2010).

*Relationships between co-infection and fitness distributions*

In order to equate changes in genetic diversity to phenotypic change and directly assess how levels of co-infection alter fitness landscapes, the relative fitness of 40 biological clones was assessed for WNV CP40 following an additional passage of 24 hours on mosquito cells at either an MOI of 10 or 0.01. Since the starting population was already characterized (Fig. 13), this allowed a direct assessment of how co-infection functions to alter the phenotypic diversity in the WNV swarm. Viral titers of individual variants following plaque isolation and 72 hours growth on mosquito cell culture were found to be modestly higher for the MOI 0.01 group (Fig. 16 inset; t-test, p<0.001). Competition assay results demonstrate that the level of co-infection in a single round of amplification can substantially alter the composition of variants maintained in the population (Fig. 16). The mean relative fitness value for the variants derived from the MOI 10 infection was 21.8, which, despite the relatively lower individual viral titers, was significantly higher than variants derived from the MOI 0.01 infection (mean =5.2; t-test, p<0.0001). In addition, more phenotypic diversity was maintained in the MOI 10 group relative to the MOI 0.01 group (standard deviations =14.7 and 4.8, respectively). Relative fitness values ranged from 5.6 to 68.5 in the MOI 10 group and 0.87 to 19.4 in the MOI 0.01
group. The relative fitness value of the combined population for the MOI 0.01 group was 16.9, which was higher than all but 3 (92%) of the variants. None of the individual variants in the MOI 10 group had relative fitness values that measured as high as the combined population fitness (84.3; Fig. 16).

Discussion

A more complete understanding of the phenotypic consequences of diverse arboviral swarms is needed in order to equate RNA swarm breadth and composition to specific adaptive and evolutionary outcomes. This requires both an ability to correlate genomic diversity to intrahost fitness landscapes and an understanding of the potential for interactions among viral genomes and proteins. We determined clonal fitness landscapes of a highly genetically diverse mosquito cell-adapted WNV strain and evaluated how altering the capacity for cooperative interactions among variants affects mutant swarm dynamics. The somewhat counterintuitive observation of accumulating genetic diversity of WNV in the face of adaptation ([Ciota et al., 2007c]; Fig. 12) can be explained in two ways; (i) A mutationally robust population has formed around a high fitness master sequence, or (ii) populations of variants of lower fitness values are maintained in the population via cooperative interactions with more highly adapted variants. Our results demonstrate that the production and maintenance of genetic diversity in mosquito cell-adapted WNV is likely a result of a combination of these factors. Specifically, although each clone is predicted to have an average of 12 mutations relative to the consensus sequence [WNV CP40 nt diversity = 0.11%; (Ciota et al., 2007c)], as predicted in a
robust viral population, there is substantial phenotypic redundancy in the population, with many clones displaying similar values for relative fitness (Fig. 13). Recent characterization of a mutagenized poliovirus swarm demonstrated the capacity for an RNA virus population to occupy such a neutral fitness landscape (Lauring et al., 2011). In contrast, results here also demonstrate significant fitness variation among outliers in the WNV CP40 population, with a greater than 15-fold disparity measured in relative fitness values of individual clones within the population. Similar to previous studies with VSV, the majority of variants were found to have fitness values lower than that of the combined population (Duarte et al., 1994). These data suggest that cooperative interactions may work to maintain lower fitness variants, yet also suggest that there may be a limit to the potential of complementation, as all variants possess at least modest adaptation to mosquito cells.

Historically, discussion of complementation among viral proteins has focused primarily on defective interfering particles. The capacity for these incomplete genomes to hijack proteins from viable viral particles for replication has been documented extensively in cell culture systems, particularly during persistent infections (Huang & Baltimore, 1970). In addition, previous studies with Dengue virus have demonstrated the maintenance of nonviable deletion mutants in both Aedes aegypti mosquitoes and acutely infected humans (Aaskov et al., 2006; Li et al., 2011), and recent work with WNV infected Culex quinquefasciatus also suggests that complementation may act to maintain deletion mutants in this system (Brackney et al., 2011). These studies together suggest that arboviruses in natural systems have the capacity to utilize complementation in co-infecting cells and that such interactions could readily occur among variants with more
intermediate fitness levels. Experimental studies with VSV have shown that relative fitness is MOI-dependent due to increased levels of strain complementation (Novella et al., 2004). Our results show a similar trend, i.e. an increased capacity for the lower fitness WNV MARM to remain in the population when levels of co-infection are high (Fig. 14a). Further, we demonstrate that this is not just dependent on initial MOI, but also on viral titer during a single infection (Fig. 14b). Specifically, as viral replication increases, increasing co-infection effectively decreases the pace at which lower fitness mutants are selected against and, therefore, the capacity for high fitness variants to experience selective sweeps. Related studies with the bacteriophage Φ6 also demonstrate that co-infection can weaken selection, yet interactions with variants of this virus are further complicated by its capacity to readily reassort (Froissart et al., 2004). With a single open-reading frame, WNV is not capable of reassortment, nor is it thought to readily recombine (Pickett & Lefkowitz, 2009; Taucher et al., 2010). Although the full genomes of output variants from MOI studies were not sequenced, the partial sequences again fail to identify evidence of recombination, suggesting interactions likely occur on the level of viral proteins. Decreased strength of selection during single mosquito infections could be particularly advantageous to arboviruses as it would increase the capacity to maintain vertebrate-specific variants which might otherwise be purged. Previous studies with WNV CP40 demonstrated that this strain accrued no obvious phenotypic cost as a result of mosquito cell adaptation in vertebrate hosts in vitro or in vivo (Ciota et al., 2007a; Ciota et al., 2008). It remains to be seen how widespread this phenomenon is in nature, but it is possible that it could contribute to the slower than expected pace of consensus level evolution of arboviruses, often attributed to the
conflicting selective pressures resulting from host cycling (Scott et al., 1994; Woolhouse et al., 2001). Our results confirm that increased co-infection results in both increased genetic and phenotypic variation (Figs. 15 and 16). Although one could argue that the genetic diversity differences could be attributed to variable bottleneck size, an MOI of 0.01 in these studies still requires approximately 1 million WNV particles. In addition, the differences in mean hamming distances of the MOI 10 variants, which represent an increased likelihood to retain more genetically distant haplotypes, cannot be explained by variation in bottleneck size (Fig. 15). The difference in distributions of relative fitness values among individual variants with or without co-infection demonstrates the profound effect that cooperative interactions can have in shaping fitness landscapes (Fig. 16).

Initial titrations of WNV clones isolated from these populations suggest only modest phenotypic differences, with mean viral titer ~ 0.5 log_{10} pfu/ml lower for the MOI 10 group (t-test; p<0.001; Fig. 16). This result is consistent with the prediction that when co-infection is common, variants that have moderately inferior replicative ability are more likely to be maintained in the population as a result of complementation.

Surprisingly, despite lower viral titers, relative fitness values were substantially higher for the MOI 10 variants (t-test, p <0.001, Fig. 16). This result highlights how individual replicative ability is very much distinct from fitness in competition and, therefore, how misleading using such measures as a surrogate for viral fitness can be. Here, high fitness variants do not grow to higher titers in the absence of competition, yet are exceedingly better at competing for cellular resources. The MOI 10 population clearly demonstrates variant complementation, as the population fitness in this case is greater than any of its individual components. These data are unique in that they suggest not just that individual
genotypes of lower fitness can be retained by complementation, but also that there may be a population benefit as a result of additive fitness. In this scenario, individual genotypes could act as specialists coding for proteins that vary in their capacity to carry out different functions in the WNV life cycle. For unique variants in the absence of co-infection, it is feasible to imagine that there could be some functional trade-offs via antagonistic pleiotropy, yet if products of variants are readily shared among co-infecting particles, such trade-offs could potentially be overcome. Although there are no specific examples of functional antagonistic pleiotropy for an arbovirus during infection of a single host, experimental evolution studies clearly demonstrate that there are many ways for an arbovirus to enhance fitness, and despite the pace and breadth at which these viruses can survey sequence space, highly fit genotypes often possess dissimilar changes in variable genomic regions (Greene et al., 2005; Novella & Ebendick-Corp, 2004; Weaver et al., 1999).

To determine the relevance of cooperative interactions in natural systems there remain many questions to answer. First, does co-infection occur frequently enough in nature that cooperation could significantly contribute to alterations in arbovirus mutant swarms and, under what conditions would cooperative interactions be selected for? Studies with Φ6 suggest intrahost competition could ultimately select for selfish genotypes if levels of co-infection are high, yet the constantly changing selective pressures arboviruses face may not permit this (Turner & Chao, 1999). Further, it is not clear if the potential costs of cooperative interactions, such as decreased strength of selection for fast replicating variants would outweigh the benefit in mosquito hosts, yet if in vitro results presented here also are observed in vivo, populations of cooperating
arbovirus genotypes may have the added benefits of both higher population fitness and increased phenotypic plasticity in the face of ever changing landscapes. Studies utilizing technology to track large populations, such as recent work with poliovirus, could provide a more comprehensive characterization of dynamic RNA arboviral swarms which could complement phenotypic characterization in vivo (Lauring et al., 2011). Although results presented here by no means depict the full complexity of the arboviral swarm, these data demonstrate the profound effects cooperative interactions could have on arbovirus evolution and adaptation and provide a baseline by which to study the impact of this phenomenon in natural systems.

Acknowledgments

Dylan J. Ehrbar, Greta V.S. Jerzak, Graham G. Willsey, and Laura D. Kramer are also authors on the originally published manuscript and all contributed to these studies. The authors thank Kiet Ngo for experimental assistance and Claus Wilke for critical review of this manuscript. In addition, the authors thank the Wadsworth Center Applied Genomics Technology Core for sequencing, and the Wadsworth Center Media and Tissue Culture Facility for providing cells and media for this work. This work was supported by federal funds from the National Institute of Health (grant number RO1-AI-077669).
**Chapter 4 tables**

**Table 9.** Infectivity and co-infection of WNV CP40 in mosquito cells.

<table>
<thead>
<tr>
<th>WNV particles inoculum</th>
<th>Remaining WNV particles -1 hr.</th>
<th>Total infectious WNV particles (%)</th>
<th>Cells infected -12 hrs (%)</th>
<th>Mean particle/cell ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1.0 \times 10^8$</td>
<td>$8.0 \times 10^7$</td>
<td>$2.0 \times 10^7 (29.2)$</td>
<td>$1.4 \times 10^6 (68.2)$</td>
<td>14.3</td>
</tr>
</tbody>
</table>

**Chapter 4 Figures**

**Figure 12.** Haplotype accumulation of *West Nile virus* during passage in mosquito cell culture. Unique shades and/or patterns represent unique haplotypes identified by high-fidelity molecular cloning and sequencing of 15-20 clones of nt 1311-3248 following passage in C6/36 mosquito cells (modified from data originally published in (Ciota et al., 2007c)).
Figure 13. Distribution of relative fitness values among 40 biological clones of WNV following passage on mosquito cell culture (WNV CP40). Relative fitness refers to the proportion of individual WNV CP40 clones relative to the control virus (WNV MARM) following 72 hours of competition on mosquito cell culture. The solid line represents the mean +/- standard deviation and the dotted line represents the fitness of the combined population. The circled data point represents the highest fitness clone isolated, WNV CP40-12.
Figure 14. Relationship between levels of co-infection and relative fitness of WNV CP40-12 on mosquito cell culture. a. Individual points represent relative fitness values +/- standard error of duplicate competition assays carried out for 72 hours at indicated MOIs. A significant negative correlation between MOI and relative fitness was measured (linear regression analysis, $r^2=0.68$, $p=0.0035$). b. Relationship between WNV titer and relative fitness during competition following infection at an MOI of 10. Individual points represent duplicate measurements of output virus following 24 (circles), 48(squares), or 72 (triangles) hours of competition. Results demonstrate a negative correlation between WNV titer and relative fitness (Pearson correlation, $r=-0.88$, $p=0.0195$).
Figure 15. Relationship between levels of co-infection and genetic diversity of WNV. Bars depict normalized Shannon entropy ($S_n$) on both the nucleotide (nt) and amino acid (aa) levels and circles depict relative differences in mean hamming distance for populations following a single passage of WNV CP40 at indicated MOIs. All measures of genetic diversity are for nt 1311-3248 and 17-22 clones were sequenced and analyzed for each measure. A significant positive correlation between MOI and $S_n$ for both nt and aa was measured (linear regression analyses, $r^2=0.99$, $p<0.05$).
**Figure 16.** Fitness distributions of WNV CP40 with or without co-infection on mosquito cell culture. Individual data points represent relative fitness values of 40 clones isolated following 24 hours growth at MOI 10 or 0.01. Solid lines represent means +/- standard deviations and dashed lines represent relative fitness values of combined populations. Inset graph depicts individual viral titers of clones following 72 hours growth on mosquito cell culture. Significantly higher relative fitness values and significantly lower viral titers were measured in the MOI 10 group (t-tests, p<0.05), as indicated by the *. 
References


CHAPTER 5: THE EVOLUTION OF VIRULENCE OF WEST NILE VIRUS IN A MOSQUITO VECTOR: IMPLICATIONS FOR ARBOVIRUS ADAPTATION AND EVOLUTION

Abstract

Virulence is often coupled with replicative fitness of viruses in vertebrate systems, yet the relationship between virulence and fitness of arthropod-borne viruses (arboviruses) in invertebrates has not been evaluated. Although the interactions between vector-borne pathogens and their invertebrate hosts have been characterized as being largely benign, some costs of arbovirus exposure have been identified for mosquitoes. The extent to which these costs are strain and host-specific and the subsequent consequences of these interactions on vector and virus evolution have not been adequately explored. Here, using West Nile virus (WNV) and Culex pipiens mosquitoes, we tested the hypothesis that intrahost fitness is correlated with virulence in mosquitoes by evaluating life history traits following exposure to either non-infectious bloodmeals or bloodmeals containing wildtype (WNV WT) or the high fitness, mosquito-adapted strain, WNV MP20. Our results demonstrate strain-specific effects on mosquito survival, fecundity, and blood feeding behavior. Specifically, exposure to MP20, but not WT decreased survival of Cx. pipiens and altered fecundity and bloodfeeding such that early egg output was enhanced at a later cost. As predicted by the trade-off hypothesis of virulence, costs of infection with MP20 in terms of survival were directly correlated to viral load, yet resistance to infection with this virulent strain was equally costly. Taken together, these results demonstrate that MP20 infection decreases the transmission potential of Cx. pipiens populations despite the increased intrahost fitness of this strain, indicating that a
virulence-transmission trade-off in invertebrates could contribute significantly to the adaptive and evolutionary constraint of arboviruses.

**Introduction**

Virulence, the fitness cost to a host resulting from pathogen infection, is a dynamic trait fluctuating with the co-evolution of both host and pathogen as well as with their interactions with changing environments. Although newly emergent pathogens are often more virulent and many pathogens have displayed decreased virulence over time (Read, 1994), the avirulence hypothesis, the idea that pathogens should always evolve away from virulent interactions with their hosts, has been largely disproven by epidemiological and experimental data demonstrating the persistence and/or evolution of highly virulent pathogen strains (Ewald, 2004; Alizon et al. 2009). Despite this, although vector-borne pathogens are often associated with high virulence in vertebrate hosts (Day, 2002), interactions between arthropod vectors and arthropod-borne viruses (arboviruses) have historically been characterized as benign (Burnet and White, 1972; Hardy et al. 1983; Dohm et al. 1991). Although the term *vector* implies a lack of significant biological interaction between arthropods and the pathogens they carry, it has become clear in recent years that such interactions are complex and are likely dominant forces shaping the evolution of arboviruses (Ciota 2010; Kuno and Chang, 2005; Fragkoudis et al. 2009; Weaver, 2006).

The alternative to the avirulence hypothesis is the trade-off hypothesis, which proposes that virulence and transmission are coupled and that the extent of virulence at equilibrium is subsequently limited by the trade-off that maximizes pathogen
transmissibility (Anderson and May, 1982). Although variability in modes of transmission, intrahost competition, and relationships between virulence and pathogen load for individual host-pathogen systems may argue against the broad applicability of this hypothesis to explain variations in virulence (Ebert and Bull, 2003), the trade-off hypothesis nevertheless provides a useful framework by which to evaluate the capacity for virulence evolution in individual systems. The coupling of virulence and transmission has indeed been noted in many systems (Alizon et al., 2009; Lipsitch and Moxon, 1997; Sacristan and Garcia-Arenal, 2008), yet to-date has not been evaluated for an arbovirus in an invertebrate host.

Arboviruses, which are almost exclusively mosquito-borne RNA viruses with inherently vast evolutionary potential, have been relatively slow to evolve (Cilnis et al., 1996; Weaver et al., 1992; Holmes and Twiddy, 2003; Pesko and Ebel, 2012). This evolutionary constraint has been attributed primarily to the obligate cycling between disparate vertebrate and invertebrate hosts (Weaver et al., 1999; Coffey et al., 2008; Deardorff et al., 2011), yet the effect of host cycling may at times be overstated (Ciota et al., 2007; Ciota et al., 2008; Ciota et al., 2009; Novella et al., 1999); and genetic bottlenecks both within and among hosts and seasons (Ciota et al., 2011; Ciota et al., 2012), as well as cooperative interactions among variants (Ciota et al., 2012) could also contribute to dampened rates of evolutionary change in nature. In addition, a coupling of viral fitness and vector virulence could add further to this evolutionary constraint.

Although mosquito-borne viruses which rely heavily on vertical transmission for maintenance are generally not thought to be highly pathogenic to their invertebrate hosts, significant effects on life history traits of mosquitoes have at times been associated with
infection of horizontally transmitted viruses generally associated with human disease (Lambrechts and Scott, 2009). Fitness costs in terms of decreased survival or fecundity, as well as tissue-specific pathology have been noted with both Alphaviruses (Mims et al., 1966; Weaver et al., 1992; Scott and Lorenz, 1998; Moncayo et al., 2000; Mahmood et al., 2004; Westbrook et al., 2010) and Flaviviruses (Styer et al., 2007; Maciel-de-Freita et al., 2011; Ciota et al., 2011b). West Nile virus (WNV; Flaviviridae: Flavivirus), the most geographically widespread arbovirus in the world, is vectored primarily by Culex mosquitoes. Previous studies have demonstrated species-specific differences in the costs of WNV resistance and infection in Culex mosquitoes which correlate with variation in WNV vector competence (Ciota et al., 2011b). Specifically, colonized Culex tarsalis exhibited fecundity costs associated with WNV infection but no cost for resistance (Styer, Meola, and Kramer, 2007), while Cx. pipiens demonstrated a cost for resistance but no cost associated with infection with wildtype WNV. Experimental evolution studies with WNV previously generated a mosquito-adapted strain with both increased replicative ability and infectivity in Cx. pipiens [WNV MP20, (Ciota et al., 2008)]. Here, by evaluating and contrasting life history traits of Cx. pipiens following exposure to either wildtype WNV or mosquito-adapted WNV MP20, we tested the hypothesis that virulence and viral fitness are coupled in vector-virus relationships, therefore limiting the capacity for arbovirus adaptation for higher levels of fitness in mosquito vectors. Our results provide a straightforward assessment of the relationships between intrahost viral fitness, virulence, and vectorial capacity which demonstrate that arbovirus adaptation and evolution could be profoundly influenced by strain-specific effects on life-history traits and transmission potential of mosquito vectors.
Materials and Methods

Virus strains and testing

Wildtype WNV (WT) was derived from WNV NY003356, a primary isolate from an American crow that was collected in 2000 in Staten Island (Ebel et al., 2001) by plaque purification and amplification on Vero cells (ATCC #CCL-81) as previously described (Ciota et al., 2007). Mosquito-passaged and adapted WNV (MP20) was derived by passage of WNV WT 20 times in Cx. pipiens using intrathoracic inoculation and subsequent collection of salivary secretions for each passage as previously described (Ciota et al., 2008). Mosquito bodies, legs, and larval pools were collected in 1ml mosquito diluent (20% heat-inactivated fetal bovine serum (FBS) in Dulbecco’s phosphate-buffered saline (PBS) plus 50 μg/ml penicillin/streptomycin, 50 μg/ml gentamicin, and 2.5 μg/ml Fungizone) and subsequently subject to homogenization and centrifugation as previously described (Ciota et al., 2009). All WNV screens and titrations for virus quantification were completed by plaque assay on Vero cell culture (Payne et al., 2006). Resistant and susceptible mosquitoes were defined following virus exposure as uninfected or WNV-infected, respectively, given these experimental conditions, and therefore do not necessarily imply competence of mosquitoes when exposed to other WNV strains and/or doses.

Mosquitoes
*Cx. pipiens* egg rafts were originally collected in Pennsylvania in 2004 (courtesy of M. Hutchinson) and colonized at the Arbovirus laboratory, Wadsworth Center. Mosquitoes were reared and maintained in 30.5 cm$^3$ cages in an environmental chamber at 27°C, 50-65% relative humidity with a photoperiod of 16:8 (light:dark) hours. 300 adult mosquitoes (100 male/200 female) were collected for each exposure group upon emergence, held in mesh top 3.8 L paper cartons, and provided cotton pads with 10% sucrose *ad libitum*. Mosquitoes were held for 5-7 days prior to blood feeding to allow for mating.

**Blood feeding**

Mosquitoes were deprived of sucrose for 24 hrs prior to blood feeding. Following starvation, females were distributed into three 0.6 L cups for experimental infections in the BSL-3 insectary. Mosquitoes were fed on defibrinated chicken blood (Rockland) with 2.5% sucrose together with either 20% BA-1 (unexposed; Hanks M-199 salts, 0.05 Tris pH 7.6, 1% bovine serum albumin, 0.035g/l sodium bicarbonate, 100 units/ml penicillin, 100mg/ml streptomycin, 1 mg/ml fungizone), WNV WT, or WNV MP20. Virus strains were diluted to 8.5 log$_{10}$ pfu/ml in BA-1 prior to bloodmeal mixing. Feeding was carried out for one hr using Hemotek membrane feeders (Discovery Workshops) heated to 37°C. Mosquitoes were then anesthetized using CO$_2$ and fully-engorged females were separated and housed individually in cups containing oviposition dishes with 15mls of distilled water and access to 10% sucrose.

Subsequent uninfected blood meals were offered to all groups once a week for the duration of the study. Specifically, mosquitoes were again starved for 24 hrs and then
offered pledgets soaked with chicken blood with 2.5% sucrose for one hour. Mosquitoes were monitored during these feedings and both numbers fed and degree of engorgement (1, small amount of blood in abdomen, no abdominal distention; 2, some distention, no pleural membrane observed; 3, significant abdominal blood, pleural membrane observed; 4, fully engorged, distended abdomen) were recorded.

*Mosquito fitness*

Mortality and egg production were monitored and recorded daily for all groups. Wings were removed from dead mosquitoes, individually mounted on slides with double-sided tape, and measured as previously described using a Zeiss microscope, Axiocam camera, and Axiovision software [Carl Zeiss; (Styer et al., 2007)]. Individual mosquito bodies and legs were stored separately at -80°C and subsequently processed and tested for WNV as described above. Egg rafts were photographed under 50X magnification using a digital camera (Nikon) and digital images were used to count individual eggs.

Oviposition cups containing egg rafts were removed and held for approximately 2 days at 27°C to allow for hatching. Hatched larvae were provided with food (ground koi food: ground rabbit pellets, 1:1) and allowed to develop to 1st–2nd instar to permit counting and subsequent calculation of hatch rates. Larvae from individual rafts were combined in pools of 20-25, stored in MD at -80°C, and processed and tested as described for mosquito bodies.

*Data analysis*
GraphPad Prism software version 4.0 was used for generation and analyses of survival curves. Statistical comparisons of curves were completed using a log-rank test. Both survival and reproductive data were used to construct life history tables for each group in separate replicates. Survival \((l_x)\) is equivalent to the proportion of mosquitoes surviving to day \(x\), and reproductive output \((m_x)\) is equivalent to the number of eggs produced on day \(x\). Data for \(m_x\) was smoothed by averaging individual daily egg output with the egg output on both previous and subsequent days. Reproductive output (total eggs produced in an average female’s lifetime; \(R_0=\sum l_x m_x\)), generation time (average age at which females produce eggs; \(T=\sum l_x m_x x / R_0\)), and intrinsic rate of increase (instantaneous population growth rate; \(r = \ln R_0/T\)) (Ricklefs, 1973; Carey, 1993) were subsequently calculated. GraphPad Prism 4.0 was used for Chi-squared tests, Fisher’s exact tests, F tests, and correlation analyses, and Microsoft Excel was used to perform t-tests. Vectorial capacity (VC) is defined as the number of new hosts exposed to a pathogen by a specified population of mosquitoes per infected host per day (Macdonald, 1957; Black and Moore, 1996). \(VC= mh^2p^N b/ -\ln(p)\), where \(m\) = the number of mosquitoes/host, \(h\) = host feeding rate, \(p\) = the probability of daily survival, \(N\) = the mean extrinsic incubation period, and \(b\) = vector competence (proportion of exposed mosquitoes with disseminated infections).

**Results**

*WNV infection*
Viral titers of infectious bloodmeals were $7.7$ and $7.9 \log_{10} \text{pfu WNV/ml}$ for WNV MP20 and WNV WT, respectively. Initial feeding rates were similar among groups, averaging ~55.0%. There were a total of 66, 67, and 70 fully engorged Cx. pipiens in the uninfected, MP20 and WT groups, respectively. WNV infection rates were significantly higher for mosquitoes exposed to MP20 relative to those exposed to WT ($74.6\%$ vs $55.7\%$; Fisher’s exact test, $p=0.031$; Fig. 17). Mean WNV loads at the time of death were $4.7$ and $4.4 \log_{10} \text{pfu/mosquito}$ for MP20 and WT groups, respectively. Although this approximately 2-fold difference does not equate to statistically significant higher overall viral loads for the MP20 group when comparing geometric means, the variation in time tested (time of death) between both groups and individuals does not permit an accurate comparison of overall differences in viral loads. A significantly greater variation in body titers was measured for the MP20 group ($F$ test, $F=2.38$, $p=0.009$; Fig. 18). Dissemination rate, i.e., the proportion of infected individuals with WNV positive legs, was higher in the MP20-susceptible group, although the difference was not statistically significant (Fisher’s exact test, $p=0.21$; Fig. 17) and again is confounded by variability in times of death. Significantly higher viral loads at the time of death were measured in MP20-susceptible mosquitoes relative to the WT-susceptible group when comparing individuals with disseminated infections ($5.8$ v. $4.9 \log_{10} \text{pfu/mosquito}$; $t$-test, $p=0.004$), whereas the opposite relationship (higher viral loads in the WT group) was observed when comparing viral loads of individuals with non-disseminated infections ($3.9$ v. $4.9 \log_{10} \text{pfu/mosquito}$; $t$-test, $p=0.006$).

A total of 5182 larvae, 2772 from the MP20-susceptible group and 2410 from the WT-susceptible group, were pooled and screened for WNV. A total of 7 pools were identified.
as WNV positive, 5 from WT-susceptible mosquitoes and 2 from MP20-susceptible mosquitoes. This equated to vertical transmission rates of 2.08 and 0.72 per 1000 larvae for WT and MP20-susceptible mosquitoes, respectively (table 10). The 2 MP20 positive pools came from a single egg raft, whereas the 5 WT positive pools came from 4 egg rafts derived from 3 individual mosquitoes. All egg rafts from which positive larvae were derived were from the 3rd or 4th oviposition of individual females. In fact, all 5 egg rafts producing WNV positive larvae represented the only egg rafts from 3rd or 4th oviposits, demonstrating that WNV infected Cx. pipiens with the capacity to produce at least 3 egg rafts vertically transmitted the virus 100% of the time.

Mosquito fitness and virulence

In order to assess if alterations to Cx. pipiens fitness are associated with exposure to Cx. pipiens-adapted WNV, survival, fecundity, and wing size were assessed for individual mosquitoes following blood feeding on MP20 and compared to both WT-exposed and unexposed groups.

Survival of both WNV-resistant and WNV-susceptible mosquitoes was similar within groups (log-rank, p>0.05), permitting further comparisons based on exposure rather than infection status (Fig. 19). Decreased survival was measured in the MP20-exposed group relative to both the WT-exposed and unexposed groups (log-rank, p<0.05), whereas no difference in survival was associated with WT exposure relative to the unexposed group (Fig. 19). Differences in survival are also significant when MP20-resistant and susceptible groups are compared individually to other groups. Mean survival was 16.9
days post feeding (dpf) for the MP20-exposed group, relative to 21.9 and 23.8 dpf for the WT-exposed and unexposed groups, respectively. Maximum survival, which was 35 dpf for the MP20-exposed group, was also lower than both the WNV WT-exposed group (50 dpf) and the unexposed group (41 dpf).

Although a WNV strain-specific difference in mosquito survival is clearly demonstrated here, the relationship between viral load and virulence (survival post exposure) was more difficult to assess since viral replication is a confounding factor when comparing time of death and levels of WNV in mosquito bodies (Fig. 18). Despite this, when viral loads are compared to days surviving following the plateau of viral replication [~ 14 days, (Ciota et al., 2008)] a negative correlation between days surviving and viral load is measured for the MP20-susceptible but not the WT-susceptible (correlation analysis, Pearson r= -0.36, p=0.038). Despite this relationship beyond 14 dpf, many MP20-exposed mosquitoes died prior to 14 dpf with relatively low viral loads or a complete lack of detectable infections (MP20-resistant group).

Differences in fecundity among groups were assessed by comparing eggs/female, hatch rates, and larvae/female following exposure. Overall egg production did not differ significantly among groups (Fig. 20A), yet patterns of reproductive output were highly variable depending on exposure and/or infection status (Fig. 20B). Specifically, egg production/female was significantly higher in the MP20-exposed groups relative to WT-exposed or unexposed mosquitoes in the first week of the study (Fig. 20A; t-test, p<0.05). Eggs/female was similar among groups in the 2nd week followed by significant decline for both WT-susceptible mosquitoes and, to larger extent, MP20-exposed groups (Fig. 20B).
As has been noted in previous studies (Styer et al., 2007), hatch rates declined with time (Fig. 5A; linear regression analysis, \( r^2 = 0.817, p = 0.035 \)). Overall, significantly higher hatch rates were measured in the MP20-exposed groups relative to WT-exposed and unexposed groups (Fig. 21B; Chi-squared, \( p < 0.0001 \)) and these differences are also significant when both MP20-susceptible and MP20-resistant groups are considered separately (Chi-squared, \( p < 0.0001 \)). Since no significant differences in hatch rates among groups were measured for individual weeks (Chi-squared, \( p > 0.05 \)), the increased hatch rate for MP20-exposed mosquitoes can be attributed wholly to differences in the timing of oviposition (i.e. more egg rafts produced early when hatch rates are high).

Larvae per female and probability of daily survival were used to produce life history tables to calculate net reproductive output (\( R_0 \)), generation time (T), and intrinsic rate of population increase (r) for experimental groups (table 11). Although increased early reproductive output and increased hatch rates resulted in modestly higher mean larvae/female in the MP20-susceptible group, no significant differences were identified among groups (t-test, \( p > 0.05 \)). Values for \( R_0 \) and r were also similar among groups, with the highest net reproductive output calculated for the unexposed group (52.8) and the lowest values for both \( R_0 \) (46.3) and r (0.23) measured in the WT-resistant group. Generation times were similar for unexposed and WT exposed mosquitoes yet, consistent with the observation of increased early egg production, substantially lower for MP20-exposed mosquitoes (Fig. 20). There were no differences in mean wing length among groups, indicating that differences in life history traits and/or WNV susceptibility could not be attributed to mosquito size (table 11).
Blood feeding

In order to assess if WNV exposure, infection, and/or virus strain altered blood feeding behavior in *Cx. pipiens*, weekly feeding rates and levels of engorgement were compared among groups. No differences in mean weekly feeding rates were measured, yet WNV infection significantly enhanced early (week 1) feeding rates relative to unexposed mosquitoes, particularly for MP20-infected mosquitoes (Chi-squared, p<0.05; table 12). Since there were more individuals early in the study, week 1 differences account for the modestly elevated overall feeding rates of infected individuals, yet rates of feeding beyond week 1 of the study increased for unexposed mosquitoes and decreased for WNV-exposed mosquitoes (table 12). These differences in late feeding were significant when comparing individual MP20-resistant mosquitoes or combined MP20-exposed mosquitoes to unexposed mosquitoes (Chi-squared, p<0.01). The proportion of females that imbibed at least one bloodmeal following the initial feeding to enter the study was significantly higher in the WT-exposed groups relative to all other groups (Chi-squared, p<0.05; table 12). This can be attributed to both slightly higher feeding rates in the WT-exposed groups relative to unexposed mosquitoes as well as significantly better survival relative to the MP20-exposed groups. Although not significant, this same trend, i.e. increased likelihood of feeding with WT-exposed mosquitoes, is evident when comparing the mean number of bloodmeals taken by individual mosquitoes in each group (table 11). Levels of engorgement were also statistically equivalent among groups, yet volumes of blood ingested by fed mosquitoes were on average lower for the MP20-susceptible group, a result which may be biologically significant.
Vectorial Capacity

VC, i.e. the WNV transmission potential of this population of *Cx. pipiens*, was calculated for the WNV-exposed groups, using experimentally determined parameters, in order to assess if strain-specific differences existed. Since infection and dissemination, but not transmission were evaluated in this study, the product of these two values alone was used as a measure of vector competence (b, table 13). Although not all mosquitoes with disseminated infections transmit virus, previous characterization of these strains failed to identify significant differences in WNV transmission once disseminated (Ciota et al., 2008). Additionally, using VC alone to compare the potential for population level expansion of individual strains ($R_0$) assumes equivalent avian host susceptibility and levels of viremia, which have also been demonstrated in previous work evaluating infection and replication of these strains in chicks (Ciota et al., 2008). Results of the current study confirmed increased intrahost fitness of MP20 relative to WT, both in terms of replicative ability (Fig. 2) and vector competence (b, table 13; Fig. 17). Regression analyses of survival curves of WNV susceptible mosquitoes demonstrates generally linear mortality with daily probabilities of survival (p) of 0.97 ($r^2=0.97$) for MP20-susceptible mosquitoes and 0.98 for WT-susceptible mosquitoes ($r^2=0.88$). Mean weekly blood feeding rates were used for the host feeding rate variable (h). Mean extrinsic incubation period (N) was estimated at 10 days based on previous studies (Ciota et al., 2008; Moudy et al., 2007; Reisen et al., 2006). Since the effect of population size was not evaluated here, mosquitoes/host (m) was not included in calculations of VC, resulting in
a value representing the average transmission potential for each individual mosquito/host. Taken together, calculation of VC demonstrates that, despite adaptation for increased replicative ability and individual transmission potential, the population level transmission potential of *Cx. pipiens* exposed to the mosquito-adapted MP20 strain is lower than that of the WT-exposed population (table 13).

**Discussion**

As is the case with many pathogen-host systems, the capacity for transmission of arboviruses generally increases with increases in pathogen load in mosquitoes. For this reason, one would predict that, in the absence of opposing selective forces, evolution would favor maximal replicative fitness of arboviruses in mosquito vectors. Conventional wisdom predicts interaction of vectors and the pathogens they carry should generally be benign (Ewald, 1994), implying that there should be little constraint on intrahost fitness in invertebrate hosts. Experimental evolution studies with *St. Louis encephalitis virus* (*Flaviviridae: Flavivirus*) demonstrate an inability for further adaptation to *Cx. pipiens*, suggesting this virus may indeed have achieved its evolutionary potential in this system (Ciota et al., 2009), yet similar studies with WNV have shown significant capacity for further adaptation of this virus to *Cx. pipiens* (Ciota et al., 2008). Although there have been modest adaptive and consensus-level genetic changes since its introduction to the U.S., WNV, like many arboviruses, has remained remarkably static over time (Armstrong et al., 2011; Ebel et al., 2004; Moudy et al., 2007; McMullen et al., 2011; Jenkins et al., 2002). As has been shown with some systems, this could be partially
attributed to differential selective pressures resulting from host cycling (Jerzak et al., 2008; Deardorff et al., 2011; Coffey et al., 2008; Cilnis et al., 1996), yet studies with mosquito-adapted WNV MP20 demonstrate that host-specific adaptations without significant fitness trade-offs in vertebrate hosts are attainable, suggesting other adaptive constraints may exist (Ciota et al., 2008). In this study, using WNV MP20, we show that arbovirus adaptation could be further constrained by the coupling of intrahost fitness and virulence in mosquitoes. Although a correlation between virulence and viral load, as well as strain-specific differences in vertebrate virulence for WNV and other arboviruses is well established (Griffin et al., 2004; Bowen and Nemeth, 2007; Ewald, 1998), virulence of arboviruses in invertebrate systems has not previously been considered a dynamic trait contributing significantly to pathogen evolution. The lack of studies in this field is likely a result of the historic assumption that the invertebrate immune response is relatively generic, yet recent advances in mosquito genetics and invertebrate immunity have revealed complex interactions between vector-borne pathogens and their invertebrate hosts (Iwanaga and Lee, 2005; Fragkoudis et al., 2009). In addition, a number of studies have documented highly variable levels of arbovirus vector competence among populations of individual mosquito species, demonstrating the specificity of arbovirus-mosquito interactions and host-virus genotype by genotype outcomes (Lambrechts et al., 2009; Kilpatrick et al., 2010). This work establishes that strain-specific interactions with invertebrate hosts have the potential to be substantial forces shaping both vector and arbovirus evolution and adaptation.

Our results clearly demonstrate virulence resulting in decreased survival for MP20-exposed Cx. pipiens relative to both unexposed and WT-exposed Cx. pipiens. The fact
that this decreased survival was not measured with WT-exposed mosquitoes demonstrates strain-specificity and establishes that virulence was a by-product of experimental evolution studies selecting for this high fitness strain (Ciota et al., 2008).

For exposed individuals with detectable infections, a clear cost of infection was measured. In addition, a direct correlation between viral load and virulence beyond 14 days exposure was measured, demonstrating that arbovirus intrahost replicative fitness may be coupled with virulence in an invertebrate host. As predicted with the higher fitness *Cx. pipiens*-adapted MP20 strain, overall viral loads, as well as infection and dissemination rates were higher. Yet, interestingly, decreased survival for MP20-exposed mosquitoes was also measured for individuals with both relatively low levels of infection as well as with the MP20-resistant group, for which there were no detectable WNV infections. Previous studies with both WNV (Ciota et al., 2011) and Dengue virus (Maciel-de-Freitas et al., 2011) have also demonstrated that fitness costs can be associated with resistance to infection, yet the cost for resistance measured in our study was measured only with MP20-exposed mosquitoes, demonstrating that the decreased fitness of resistant mosquitoes is not likely due simply to a coupling of WNV resistance and low fitness of mosquitoes, but instead a direct result of strain-specific exposure and, subsequently, defense against establishment of infection. These results suggest that the magnitude of invertebrate defense against establishment of arbovirus infection may be specific and that the cost for such defense may be directly correlated to strain virulence. Further studies investigating strain-specific immune response will be required to understand variation in the mechanisms and/or extent of immune gene activation that correlates with the costs of immune deployment. Regardless of mechanism, this result
implies that chronic exposure of mosquito populations to arboviruses could have measurable effects on mosquito fitness and, subsequently, selective pressures, which are independent of vector competence.

Previous studies with WNV have demonstrated a fecundity cost of infection for *Cx. tarsalis*, but not *Cx. pipiens* mosquitoes (Ciota et al., 2011). Although in the current study we again did not measure differences in overall fecundity in *Cx. pipiens*, our results do demonstrate an association between WNV infection and substantial alterations to reproductive patterns, particularly with MP20-exposed individuals. Specifically, MP20-exposed mosquitoes maximized early egg output at a later cost. This alteration in egg production could be viewed as an adaptation to maximize reproductive output in the face of decreased fitness, particularly since both bloodmeal digestion and egg production would be extremely costly for mosquitoes whose fitness is already compromised (Briegel, 2003). Indeed, increased early egg production resulted in increased hatch rates and similar overall reproductive output for WNV MP20-exposed individuals, despite both decreased survival and egg production beyond week 2 of the study. What is not considered in a controlled laboratory rearing setting is the uncertainty of successful feeding, oviposition and egg hatching, which in nature is largely dependent on fluctuating environmental conditions and likely maximized by producing multiple egg rafts over time (Vinogradova, 2000). For this reason, the reproductive strategy of MP20-exposed mosquitoes could be less productive in a natural setting. Another consequence of these altered reproductive patterns, at least from the perspective of the pathogen, is decreased probability of vertical transmission. Since *Culex* mosquitoes do not generally take a bloodmeal prior to overwintering (Mitchell and Briegel, 1989) and vertebrate hosts are
not known to be capable of developing significant persistent or recrudescent infective viremia (Wheeler et al., 2012; Nemeth et al., 2009), the capacity for vertical transmission of WNV in mosquitoes is likely critical for maintenance in temperate regions that experience significant seasonal breaks in transmission (Nasci, Savage et al. 2001; Dohm et al., 2002; Bugbee and Forte, 2004). Since the majority of egg rafts derived from MP20-infected mosquitoes were produced following the first week of infection, the number of females vertically transmitting was lower than that of WT-infected mosquitoes. Although our overall rates of vertical transmission were low for both groups, and generally comparable to what has been previously reported (Dohm et al., 2002; Goddard et al., 2003; Anderson et al., 2008; Anderson et al., 2012), what was remarkable is that all 4 WNV-infected mosquitoes that produced a 3rd or 4th egg raft vertically transmitted. Since previous studies evaluating vertical transmission have generally been done en masse, without knowledge of the reproductive history of individuals producing positive larvae, it is possible that the potential for vertical transmission of WNV and other flaviviruses in mosquitoes have been significantly underestimated.

As with fecundity, overall blood feeding rates among groups were similar, yet differences in timing of bloodmeal acquisition, which could be significant in WNV transmission, were associated with both exposure and infection status. WNV infection significantly increased the likelihood of bloodmeal acquisition in the first feeding following exposure, particularly with the WNV MP20-exposed mosquitoes. As these mosquitoes are anautogenous, bloodmeal acquisition is a requirement for egg maturation; thus these differences generally correlate to fecundity differences and, like fecundity,
early feeding enhancement results in a subsequent decrease in rates in the following weeks. Enhanced blood feeding with infection, as has been shown with WNV in Cx. tarsalis (Styer et al., 2007) as well as with Plasmodium infection of An. gambiae (Schwartz and Koella, 2001), could be viewed as a manipulation by the pathogen to increase transmission potential, yet since the likelihood of transmission increases with time, the decreased feeding beyond 7 days measured here, particularly with the MP20-exposed mosquitoes, is likely to instead decrease WNV transmission potential. This decreased probability of transmission could be further enhanced by the fact that, on average, MP20-infected mosquitoes consumed smaller bloodmeals, likely as a result of decreased feeding times.

The method of selection utilized in creation of the MP20 strain required only that a low proportion of surviving individuals transmit WNV to be used for subsequent passages (Ciota et al., 2008). These selection criteria, although sufficient to select strains with superior replicative fitness and, therefore, transmissibility, could also tolerate modest levels of virulence. Here, by calculating vectorial capacity, we have shown that this level of virulence would inhibit transmission potential on the population level, such that wildtype WNV would have an advantage in terms of invasion and population spread, despite the increased intrahost fitness of WNV MP20 in Cx. pipiens. This cost of virulence could contribute to the overall dampened rates of evolution and partially explain why similar ‘adaptive’ arbovirus strains do not emerge and persist as readily as would be predicted for pathogens with the capacity to so rapidly explore sequence space. The notion that individual strain fitness and vector virulence may be coupled, as predicted by the trade-off hypothesis and demonstrated here, could fundamentally change
our understanding of how vector-virus interactions work to shape the evolutionary trajectories of arboviruses and other vector-borne pathogens.

What remains unknown is how variation in natural populations of mosquitoes may affect susceptibility to virulence. It is well documented that colonization can be detrimental to population fitness and it is possible that field populations could tolerate more fit strains without significant costs (Mpho et al., 2002; Aguilar et al., 2005). There are also likely to be significant temporal and generational variations in fitness which may have profound effects on vector-virus interactions and subsequent outcomes of infections. Differences in terms of the cost of resistance and blood feeding behavior are evident when contrasting results presented here to previous evaluation of the effect of wildtype WNV exposure on Cx. pipiens (Ciota et al., 2011). Although these experiments were separated by ~4 years and it not surprising that significant changes to the colony population could occur over that time, this demonstrates that spatial and temporal variation precludes our capacity to make broad assumptions about the outcomes of vector-virus interactions. In whole, these results demonstrate that a greater comprehension of the complexity and specificity of interactions between vectors and pathogens will be required if we are to better characterize the evolution of these systems.

Acknowledgements
Amy C. Matachiero and Dylan J. Ehrbar also contributed significant technical assistance to this work. I thank all of the members of the Arbovirus laboratory insectary staff for assistance with these studies and, particularly, Pamela Chin for mosquito rearing. Thanks to the Wadsworth Center tissue and media facility for providing cells. This work was funded by the National Institutes of Health (NIH), grant number R01-AI-077669. The construction of the Wadsworth Center Insectary Facility was partially funded by NIH grant number C06-RR-17715.
Chapter 5 tables

Table 10. Vertical transmission of West Nile virus in Cx. pipiens following infection with WNV MP20 or WNV WT.

<table>
<thead>
<tr>
<th></th>
<th>WNV MP20</th>
<th>WNV WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>total larvae tested(^1)</td>
<td>2772</td>
<td>2410</td>
</tr>
<tr>
<td>WNV+ pools</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>infection rate/1000(^2)</td>
<td>0.72</td>
<td>2.08</td>
</tr>
<tr>
<td>WNV + rafts/total (%)</td>
<td>1/51 (1.96)</td>
<td>4/47 (8.51)</td>
</tr>
<tr>
<td>adults VT/ ovipositing (%)(^3)</td>
<td>1/39 (2.56)</td>
<td>3/30 (10.0)</td>
</tr>
<tr>
<td>WNV+ OV 3-4 rafts/total(^4)</td>
<td>1/1</td>
<td>4/4</td>
</tr>
</tbody>
</table>

\(^1\) larvae were processed and tested in pools of 20-25
\(^2\) WNV+ pools were assumed to have a single positive individual
\(^3\) Refers to the proportion of ovipositing females vertically transmitting
\(^4\) Refers to the proportion of 3\(^{rd}\) or 4\(^{th}\) oviposition rafts which were WNV+
<table>
<thead>
<tr>
<th></th>
<th>UNEXP</th>
<th>WNV WT S</th>
<th>WNV WT R</th>
<th>WNV MP20 S</th>
<th>WNV MP20 R</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean wing length (mm)</td>
<td>3.57</td>
<td>3.59</td>
<td>3.56</td>
<td>3.57</td>
<td>3.58</td>
</tr>
<tr>
<td>mean survival time (d)</td>
<td>16.4</td>
<td>16.3</td>
<td>17.3</td>
<td>16.6</td>
<td>13.7</td>
</tr>
<tr>
<td>mean larvae/female (+/−) SE</td>
<td>0.31</td>
<td>0.29</td>
<td>0.23</td>
<td>0.28</td>
<td>0.32</td>
</tr>
<tr>
<td>mean generation time (d)</td>
<td>5.9</td>
<td>5.9</td>
<td>5.8</td>
<td>5.6</td>
<td>5.1</td>
</tr>
<tr>
<td>net reproductive output (p)</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>intrinsic rate of population increase (r)</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>mean larva/female (+/−) SE</td>
<td>0.26</td>
<td>0.26</td>
<td>0.23</td>
<td>0.28</td>
<td>0.32</td>
</tr>
</tbody>
</table>

Table 11. Summary of mosquito fitness of West Nile virus susceptible (S), resistant (R), or non-infectious (UNEXP) exposed Cx. pipiens following feeding on WNV WT, WNV MP20, or non-infectious (UNEXP).
<table>
<thead>
<tr>
<th></th>
<th>UNEXP</th>
<th>WNV WT S</th>
<th>WNV WT R</th>
<th>WNV MP20 S</th>
<th>WNV</th>
</tr>
</thead>
<tbody>
<tr>
<td>MP20 R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>overall rate&lt;sup&gt;1&lt;/sup&gt;</td>
<td>0.29</td>
<td>0.39</td>
<td>0.29</td>
<td>0.37</td>
<td>0.24</td>
</tr>
<tr>
<td>wk. 1 rate&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.20</td>
<td>0.46&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.32</td>
<td>0.63&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.43</td>
</tr>
<tr>
<td>wks. 2-6 rate&lt;sup&gt;3&lt;/sup&gt;</td>
<td>0.41</td>
<td>0.31</td>
<td>0.28</td>
<td>0.23</td>
<td>0.10&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>mean wk. rate&lt;sup&gt;4&lt;/sup&gt;</td>
<td>0.23</td>
<td>0.32</td>
<td>0.27</td>
<td>0.28</td>
<td>0.15</td>
</tr>
<tr>
<td>w/o bm&lt;sup&gt;5&lt;/sup&gt;</td>
<td>0.58</td>
<td>0.41&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.42&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.61</td>
<td>0.58</td>
</tr>
<tr>
<td>bm/female&lt;sup&gt;6&lt;/sup&gt;</td>
<td>0.66</td>
<td>0.87</td>
<td>0.77</td>
<td>0.59</td>
<td>0.41</td>
</tr>
<tr>
<td>engorgement&lt;sup&gt;7&lt;/sup&gt;</td>
<td>2.61</td>
<td>2.27</td>
<td>2.46</td>
<td>2.08</td>
<td>2.29</td>
</tr>
</tbody>
</table>

<sup>1</sup> proportion feeding throughout the study (total fed/total offered) p<0.05
<sup>2</sup> proportion feeding at 7 days post infection p<0.05
<sup>3</sup> combined proportion feeding in weeks 2-6 of the study p<0.05
<sup>4</sup> mean weekly feeding rates
<sup>5</sup> proportion of females not taking a bloodmeal throughout the study
<sup>6</sup> mean number of bloodmeals imbibed by individual mosquitoes
<sup>7</sup> mean level of engorgement (1-4)

<sup>a</sup> significantly higher than unexposed (Chi-squared,
p<0.05)
<sup>b</sup> significantly lower than unexposed (Chi-squared,
p<0.05)
<sup>c</sup> significantly lower than all other groups (Chi-squared,
p<0.05)
Table 13. Vectorial capacity (VC) of experimental populations of *Cx. pipiens* following exposure to WNV WT and WNV MP20.

<table>
<thead>
<tr>
<th></th>
<th>(h_1)</th>
<th>(b_2)</th>
<th>(p_3)</th>
<th>(N_4)</th>
<th>VC = (h^2p^N_4b/\ln(p))</th>
</tr>
</thead>
<tbody>
<tr>
<td>WNV WT</td>
<td>0.32</td>
<td>0.32</td>
<td>0.98</td>
<td>10</td>
<td>1.00</td>
</tr>
<tr>
<td>WNV MP20</td>
<td>0.28</td>
<td>0.44</td>
<td>0.97</td>
<td>10</td>
<td>0.83</td>
</tr>
</tbody>
</table>

1 mean blood feeding rate
2 vector competence (b) = infection rate * dissemination rate
3 probability of daily survival
4 extrinsic incubation period (N) = mean time in days from infection to transmission

Chapter 5 figures

![Bar chart](image)

Figure 17. Vector competence of *Cx. pipiens* following feeding on infectious bloodmeals containing either WNV WT or WNV MP20. Infection refers to the percent of individuals with WNV positive bodies at the time of death and dissemination refers to the proportion of infected individuals with WNV positive legs at the time of death. *Fisher’s exact test, p<0.05.*
Figure 18. WNV loads in *Cx. pipiens* at the time of death. (A) Individual WNV loads for all infected mosquitoes and best-fit nonlinear relationship between survival and geometric viral titers. (B) Relationship between WNV titers and survival beginning at 14 days post-infection and best fit linear relationship between survival and geometric viral titers. Slopes of lines differed significantly (linear regression analysis, \( p=0.009 \)) and a negative correlation between days surviving and viral load was measured for MP20-susceptible mosquitoes (Correlation analysis, Pearson \( r= -0.36, p=0.038 \)).
Figures 19A-C. Survival of Cx. pipiens following blood feeding. (A) Survival of WT-susceptible and WT-resistant Cx. pipiens following ingestion of WNV-exposed, MP20-exposed, and unexposed bloodmeals. (B) Survival of MP20-susceptible and MP20-resistant Cx. pipiens following ingestion of WNV-exposed and WT-exposed bloodmeals. (C) Survival following blood feeding for WT-exposed, MP20-exposed, and unexposed Cx. pipiens. Significant differences in survival were identified between MP20-exposed and unexposed Cx. pipiens. Significant differences in survival were identified between MP20-exposed and both WT-exposed and unexposed Cx. pipiens. Significant differences in survival were identified between MP20-exposed and both WT-exposed and unexposed Cx. pipiens.
Figure 20. Fecundity of *Cx. pipiens* following bloodfeeding. (A) Mean eggs/female (B) Daily eggs/female smoothed by averaging daily output with output on both previous and subsequent days.

Figure 21. *Cx. pipiens* egg hatch rates. (A) Combined weekly hatch rates for all groups. (B) Total hatch rates for individual groups. * Chi-squared, p<0.05.
References


